

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO

PUC - SP

Cassio Cristiano Giordano

**Concepções sobre Estatística: um estudo com alunos do ensino
médio**

DOUTORADO EM EDUCAÇÃO MATEMÁTICA

SÃO PAULO

2020

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO

PUC - SP

Cassio Cristiano Giordano

**Concepções sobre Estatística: um estudo com alunos do ensino
médio**

DOUTORADO EM EDUCAÇÃO MATEMÁTICA

Tese apresentada à Banca Examinadora da Pontifícia Universidade Católica de São Paulo, como exigência parcial para obtenção do título de Doutor em Educação Matemática, sob a orientação da Profa. Dra. Cileda de Queiroz e Silva Coutinho.

SÃO PAULO

2020

Banca Examinadora

Autorizo, exclusivamente para fins acadêmicos e científicos, a reprodução total ou parcial desta tese por processos de fotocopiadoras ou eletrônicos.

Assinatura: _____ **Local e Data:** _____

Dedico este trabalho ao meu pai, Sebastião Giordano (in memoriam), que nos deixou em 2019. Devo a ele tudo o que sou.

O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 88887.149050/2017-00.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 88887.149050/2017-00.

AGRADECIMENTOS

Agradeço a todos os que contribuíram, direta ou indiretamente, para a conclusão deste trabalho.

À minha orientadora, Profa. Dra. Cileda de Queiroz e Silva Coutinho, por todo o seu apoio e incentivo à pesquisa em Educação Estatística.

Aos demais professores do Mestrado e Doutorado em Educação Matemática da PUC-SP, pelo excelente ensino e pelas contribuições no desenvolvimento desta pesquisa.

Aos Drs. Celso Ribeiro Campos, Jean-Claude Régnier, Mauren Porciúncula Moreira da Silva e Saddo Ag Almouloud, por aceitarem prontamente nosso convite para compor a banca examinadora na qualificação, e a Gabriel Loureiro de Lima, que integrou a banca de defesa. Suas críticas e sugestões foram muito importantes para a realização deste trabalho.

Ao Dr. Michel Henry, por suas orientações durante meu estágio na *Université de Franche-Comté*, Besançon, em 2019.

À CAPES, pelo apoio financeiro, e à coordenação do Programa de Pós-graduação da PUC-SP, pela concessão da bolsa, que possibilitaram a realização deste trabalho.

Aos meus alunos, aos professores e à equipe de gestão da Escola Estadual Professor José Calvitti Filho, sempre prestativos.

E, finalmente, à minha família que, com amor e compreensão, me deu forças nos momentos mais difíceis.

RESUMO

Este estudo analisou os conhecimentos e as concepções mobilizadas por alunos do último ano do ensino médio de uma escola pública brasileira, dos 16 aos 19 anos de idade, antes e depois do desenvolvimento de projetos envolvendo elementos de Estatística. A fim de investigar as concepções estatísticas mobilizadas por eles, uma pesquisa de natureza qualitativa, do tipo estudo de caso, foi realizada. Dessa forma, buscou-se responder que concepções são mobilizadas por alunos do ensino médio na resolução de problemas após o desenvolvimento de projetos, utilizados como abordagem para a aprendizagem de conceitos da Estatística. Com esse objetivo, por meio da Análise Estatística Implicativa e Análise de Similaridades, foram identificados os conhecimentos prévios de 86 alunos, com base nos questionários respondidos por eles antes da participação nos projetos de pesquisa. Assim, este trabalho teve o auxílio de recursos tecnológicos oferecidos pelo *software* CHIC – Classificação Hierárquica Implicativa e Coesitiva. Além disso, foram analisadas as concepções mobilizadas por quatro grupos de alunos (duas duplas e dois trios) após o desenvolvimento dos projetos, por meio de registros escritos e gravações da sua produção bem como das interações com os colegas dos respectivos grupos, durante a resolução de problemas estatísticos semelhantes àqueles resolvidos em suas pesquisas. No que se refere aos conhecimentos, considerou-se que ainda existe um longo caminho para que os alunos cheguem ao término do ensino médio com um nível de letramento estatístico satisfatório para uma maior compreensão de mundo e exercício pleno da cidadania assim como para a transição para o ensino superior, e que a implantação da BNCC pode ajudar nesse desafio. Foi possível identificar onze concepções estatísticas mobilizadas pelos alunos, algumas delas deslocadas de seu domínio de validade, refletindo dificuldades de aprendizagem, mas também avanços, falhas estruturais e possíveis caminhos no ensino e aprendizagem de Estatística no Brasil.

Palavras-chave: Educação Estatística, Concepções, Análise Estatística Implicativa e de Similaridades, Projetos, BNCC.

ABSTRACT

This research has analyzed the knowledge and the conceptions mobilized by students on the last grade of the Brazilian public high school, aged from 16 to 19 years old, before and after the development of projects involving elements of Descriptive Statistics. In order to investigate the statistical conceptions mobilized by them, a qualitative research (case study kind), has been conducted. That way, it has sought to answer which conceptions are mobilized by high school students in solving problems after the development of projects, used as an approach to learning concepts from Descriptive Statistics. For this purpose, through Implicative Statistical Analysis and Similarity Analysis, during methodological procedures, the previous knowledge of 86 students has been identified, based on the questionnaires filled by them before their participation in the research projects. Thus, this work has had the assistance of the technological resources offered by CHIC software – Hierarchical Classification Implicative and Cohesive. Besides that, it has been performed an analysis of the conceptions mobilized by four groups of students (two pairs and two trios) after the development of the projects, using their annotations and the recordings of their production as well as of their interactions with their respective groups' colleagues, when solving statistical problems similar to those solved in their research. With regard to knowledge, it is believed that there is still a long way for students to reach high school with a satisfactory level of statistical literacy for a better understanding of the world and full exercise of citizenship, as well as for the transition to higher education, and the implementation of the National Common Curricular Base – BNCC might help in this challenge. It was possible to identify eleven statistical conceptions mobilized by the students, some of them displaced from their domain of validity, reflecting learning difficulties, but also advances, structural failures and possible paths in the teaching and learning of Statistics in Brazil.

Keywords: Statistical Education, Conceptions, Implicative Statistical Analysis and Similarity Analysis, Projects, BNCC.

LISTA DE ILUSTRAÇÕES

Figura 1 – Modelo de letramento estatístico.....	43
Figura 2 – Modelização da ASI	50
Figura 3 – Modelização da hierarquia de similaridade.....	51
Figura 4 – Esquema das possíveis relações entre concepções, conhecimentos e conceitos, propostas por Balacheff	56
Figura 5 – Elementos da Estatística associados ao Pensamento Estatístico	57
Figura 6 – Retroação no sistema sujeito-meio	58
Figura 7 – Sistema didático na Teoria das Concepções.....	58
Figura 8 – Medidas estatística numéricas	62
Figura 9 – Construção gráfica da moda de Czuber	66
Figura 10 – Pesos de varones y hembras	69
Figura 11 – Pesos de varones y hembras	70
Figura 12 – Organização matemática envolvendo variabilidade nos livros didáticos	88
Figura 13 – Questão 25 do questionário aplicado	96
Figura 14 – Questão 26 do questionário aplicado	97
Figura 15 – Questão 27 do questionário aplicado	99
Figura 16 – Questão 28 do questionário aplicado	101
Figura 17 – Questão 29 do questionário aplicado	103
Figura 18 – Árvore coesitiva para as questões relativas à Estatística	106
Figura 19 – Trecho da árvore coesitiva I	107
Figura 20 – Trecho da árvore coesitiva II	107
Figura 21 – Trecho da árvore coesitiva III	108
Figura 22 – Trecho da árvore coesitiva IV	109
Figura 23 – Trecho da árvore coesitiva V	110
Figura 24 – Trecho da árvore coesitiva VI	110
Figura 25 – Trecho da árvore coesitiva VII	111
Figura 26 – Trecho da árvore coesitiva VIII	112
Figura 27 – Trecho da árvore coesitiva IX	113
Figura 28 – Árvore de similaridades	114
Figura 29 – Árvore de similaridades: relações significativas I	115
Figura 30 – Árvore de similaridades: relações significativas II	116
Figura 31 – Árvore de similaridades: relações significativas III	117
Figura 32 – Árvore de similaridades: relações significativas IV	119
Figura 33 – Árvore de similaridades: relações significativas V	120
Figura 34 – Árvore de similaridades: relações significativas VI	121
Figura 35 – Árvore de similaridades: relações significativas VII	122

Figura 36 – Árvore de similaridades: relações significativas VIII	122
Figura 37 – Grafo implicativo	124
Figura 38 – Questão 1 dos exercícios aplicados	136
Figura 39 – Excerto I do Apêndice 2, Protocolo 1 – Grupo I.....	138
Figura 40 – Excerto II do Apêndice 2, Protocolo 1 – Grupo I.....	141
Figura 41 – Excerto III do Apêndice 2, Protocolo 1 – Grupo I.....	142
Figura 42 – Excerto IV do Apêndice 2, Protocolo 1 – Grupo I	146
Figura 43 – Excerto V do Apêndice 2, Protocolo 1 – Grupo I	146
Figura 44 – Questão c dos exercícios aplicados	148
Figura 45 – Excerto VI do Apêndice 2, Protocolo 1 – Grupo	148
Figura 46 – Excerto VII do Apêndice 2, Protocolo 1 – Grupo I	149
Figura 47 – Excerto do Apêndice 2, Protocolo 2 – Grupo II.....	154
Figura 48 – Excerto I do Apêndice 2, Protocolo 3 – Grupo III.....	156
Figura 49 – Excerto II do Apêndice 2, Protocolo 3 – Grupo III.....	157
Figura 50 – Questão b dos exercícios aplicados	158
Figura 51 – Excerto III do Apêndice 2, Protocolo 3 – Grupo III.....	159
Figura 52 – Excerto do Apêndice 2, Protocolo 4 – Grupo IV	162

LISTA DE QUADROS

Quadro 1 – Síntese das concepções identificadas na análise dos protocolos de observação referentes ao conhecimento específico de Estatística.....	37
Quadro 2 – Principais contribuições das pesquisas analisadas em nossa revisão da literatura.....	38
Quadro 3 – Siglas utilizadas neste capítulo.....	93
Quadro 4 – Caminhos implicativos determinados na análise das variáveis em jogo....	126
Quadro 5 – Resumo dos projetos desenvolvidos pelos alunos	130
Quadro 6 – Número e duração das sessões de resolução de problemas dos quatro grupos de alunos	135
Quadro 7 – Concepções mobilizadas por alunos do último ano do ensino médio ..	164

SUMÁRIO

1 INTRODUÇÃO	13
1.1 Justificativa	15
1.2 Objetivos gerais	17
1.3 Objetivos específicos	17
1.4 Metodologia de pesquisa	17
1.5 Procedimentos metodológicos	20
2 REVISÃO DA LITERATURA	25
2.1 Estatística e a abordagem por meio de projetos no ensino médio	25
2.2 Pesquisas desenvolvidas na perspectiva da Teoria das Concepções	32
3 FUNDAMENTAÇÃO TEÓRICA E METODOLÓGICA	40
3.1 Análise Exploratória de Dados	40
3.2 Análise Estatística Implicativa (ASI) e Análise de Similaridades	47
3.3 Teoria das Concepções	52
4 ESTUDO DO OBJETO	62
4.1 Determinação de medidas de tendência central e de dispersão	62
4.1.1 Medidas de tendência central	63
4.1.1.1 Moda	64
4.1.1.2 Mediana	67
4.1.1.3 Média aritmética simples	71
4.1.1.4 Articulando média, mediana e moda	73
4.1.2 Medidas de dispersão	74
4.1.2.1 Variância	75
4.1.2.2 Desvio padrão	76
4.1.2.3 Desvio médio absoluto	76
4.1.2.4 Amplitude total	77
4.2 Registros de representação: tabelas de frequência e gráficos estatísticos	78
4.2.1 Gráficos estatísticos	81
4.2.2 Tabelas de distribuição de frequência	85
4.3 Variabilidade	86
5 O QUE OS ALUNOS SABEM SOBRE ESTATÍSTICA	91
5.1 Análise Coesitiva	105
5.2 Análise de Similaridades	113
5.3 Análise Implicativa	123
6 ANÁLISE DAS CONCEPÇÕES DOS ALUNOS	128
6.1 Concepções mobilizadas pelos alunos	133
6.2 Análise das possíveis concepções mobilizadas pelos alunos na resolução de problemas após a realização de projetos	134
7 CONSIDERAÇÕES FINAIS	169
REFERÊNCIAS	174
APÊNDICE 1 – Questionário	185
APÊNDICE 2 – Protocolos de quatro grupos de alunos	190
ANEXO 1 – Termo de Consentimento Livre e Esclarecido	236
ANEXO 2 – Folhas de respostas dos quatro grupos de alunos	237
ANEXO 3 – Carta de esclarecimento sobre o projeto de pesquisa	265
ANEXO 4 – Atestado de aprovação do projeto	267

1 INTRODUÇÃO

Meu interesse pela Educação Estatística surgiu de minhas experiências docentes e da receptividade dos alunos quanto às aulas de Matemática, ao explorar assuntos diversos mediados por essa ciência.

Desde 1993 lecionando para turmas de ensino médio — boa parte empregando a abordagem por meio de projetos — observei que havia interesse e dedicação maiores às aulas de Matemática quando envolvia tanto a Matemática Financeira quanto a Estatística, possivelmente devido ao amplo leque de aplicações imediatas em seu cotidiano. Por isso, no momento de definir minha pesquisa de Mestrado, em 2014, optei por trabalhar com a segunda.

Assim, em minha dissertação, defendida em 2016 e desenvolvida no Programa de Estudos Pós-Graduados em Educação Matemática da Pontifícia Universidade Católica de São Paulo — PEPG em Educação Matemática da PUC-SP, investiguei o letramento estatístico — nos termos de Iddo Gal (2002) — no desenvolvimento de projetos de aprendizagem, como definidos por Carmen Batanero e Carmen Díaz (2004, 2011), à luz da Teoria das Situações Didáticas – TSD de Guy Brousseau (1986, 2007).

Uma das questões levantadas tinha potencial para futuras pesquisas e é justamente a que norteia este estudo. Embora, naquele momento, as concepções mobilizadas pelos alunos não fossem meu foco de análise, chamaram a minha atenção.

Pensamos, então, em um quadro teórico favorável aos nossos objetivos e que fosse mais próximo possível daquele com o qual eu já trabalhava. Para tanto, adotamos a Análise Exploratória de Dados — AED, a fim de fundamentar nossa abordagem de resolução de problemas e desenvolvimento de projetos, além de incluirmos a Teoria das Concepções.

Julgamos relevante direcionar esforços para investigar as concepções estatísticas mobilizadas pelos alunos envolvidos na resolução de problemas após o desenvolvimento de projetos, sobretudo considerando a perspectiva da presença dessa metodologia de ensino nos currículos nacionais, com base na Base Nacional Comum Curricular — BNCC (BRASIL, 2018).

Optamos pela abordagem metodológica da pesquisa qualitativa, na perspectiva de John Creswell (2010), do tipo estudo de caso, na perspectiva de João Pedro da

Ponte (2006). Coletamos dados por meio de questionários; de produção escrita de alunos durante realização de tarefas de Estatística; e de gravações das discussões entre estudantes.

Para a análise das respostas dos alunos, obtidas por meio de um questionário, contamos com a Análise Estatística Implicativa — ASI — e Análise de Similaridades com o recurso tecnológico do software CHIC, sigla francesa para “*Classification Hiérarchique Implicative et Cohésitive*” (Classificação Hierárquica Implicativa e Coesitiva). Dessa forma, o objetivo deste trabalho é identificar conhecimentos estatísticos e concepções de estudantes de quatro turmas do 3º ano do ensino médio em uma escola pública brasileira, com o intuito de responder à nossa questão de pesquisa: “Quais concepções são mobilizadas por alunos do ensino médio ao resolverem problemas relacionados à Estatística, após o desenvolvimento de projetos nessa área?”

Em 2018, por intermédio da minha orientadora, a Dra. Cileda Coutinho, surgiu a oportunidade de realizar um estágio no *Institut de Recherche sur l'Enseignement des Mathématiques* (IREM) de *Franche-Comté*, na *Université de Franche-Comté* (UFC), em Besançon — França. Fui recebido pelo Dr. Michel Henry, que foi coorientador da tese de doutorado da própria Dra. Cileda Coutinho e pelo Dr. Philippe Le Borgne, diretor do IREM de Besançon. Na UFC pude apresentar a minha pesquisa em andamento para professores franceses e pesquisadores do IREM, por meio de dois seminários.

Durante esse estágio, que aconteceu nos meses de janeiro e fevereiro de 2019, recebi inestimáveis contribuições do Dr. Michel Henry, na fase de análise das concepções mobilizadas pelos alunos. Além disso, ele me apresentou ao Dr. Jean-Claude Régnier, do *Institut des Sciences et Pratiques d'Éducation et de Formation* (ISPEF), da *Université Lumière Lyon 2*. Fui convidado por este para apresentar a minha pesquisa nesta universidade, por meio de outro seminário.

O Dr. Jean-Claude Régnier não somente deu importantes sugestões para o desenvolvimento da minha tese durante o seminário que realizei como também integrou a banca de qualificação do meu doutorado. Ele me auxiliou na análise implicativa dos dados coletados minha pesquisa por meio de um questionário, com o auxílio do software CHIC.

Ainda como resultado dessa parceria, recebemos um convite para participar do *10ème Colloque International sur l'Analyse Statistique Implicative* (ASI 10), em Belfort

— França, em outubro de 2019. Um recorte de nossa pesquisa foi apresentado neste evento pela minha orientadora de tese.

Consideramos relevante deixar aqui o registro dessas contribuições.

A seguir, será apresentada a problematização, definindo a questão de pesquisa, objetivos, método de pesquisa utilizado e procedimentos metodológicos. Nos capítulos 2 e 3, abordamos nossa fundamentação teórica, a qual inclui a revisão da literatura e o quadro teórico, ou seja, o suporte para nossas análises. No capítulo 4, fizemos um breve estudo panorâmico sobre o objeto de investigação. A discussão sobre os resultados de nossa pesquisa é realizada nos capítulos 5 e 6. E, finalmente, no capítulo 7, tecemos nossas considerações finais.

1.1 Justificativa

Pesquisas recentes sugerem que os alunos do ensino médio, embora apresentem determinados conhecimentos sobre a Estatística, demonstram dificuldade em relacionar seus conceitos básicos aos problemas cotidianos de sua vida pessoal, escolar e profissional. Porém, essa realidade parece mudar após a participação ativa no planejamento e na execução de uma pesquisa estatística, como indicam os resultados observados nos trabalhos que serviram de fundamentação teórica para este trabalho, apresentados no próximo capítulo.

O desenvolvimento de projetos no ensino e aprendizagem de Estatística permite a aplicação, em contextos significativos para os alunos, de conhecimentos estatísticos, técnicas e estratégias de grande relevância para a formação do cidadão, em uma abordagem mais rica e complexa do que aquela oferecida pelo livro didático. Carmen Batanero e Carmen Díaz (2011, p. 21) entendem que:

[...] precisamos diferenciar entre saber e poder aplicar conhecimento. A capacidade de aplicar conhecimento matemático geralmente é muito mais difícil do que se supõe, porque requer não apenas conhecimento técnico (como preparar um gráfico ou calcular uma média), mas também conhecimento estratégico (saiba quando usar um determinado conceito ou gráfico). Os problemas e exercícios de livros didáticos geralmente se concentram apenas no conhecimento técnico. Ao trabalhar com projetos, os alunos são colocados na posição de ter que pensar em perguntas como as seguintes (Graham, 1987)¹: Qual é o meu problema? Preciso de dados? Quais? Como posso obtê-los? O que esse resultado significa na prática?

¹ Graham, A. (1987). Statistical investigations in the secondary school. Cambridge: The Open University Centre for Mathematics Education.

No entanto, a abordagem por meio de projetos demandaria um registro tão extenso que inviabilizaria a análise, nos moldes definidos por nosso quadro teórico. Assim, optamos pela resolução de problemas posterior ao desenvolvimento de projetos. Celi Espasandin Lopes e Elaine Meirelles (2005) asseveram que tanto a Estatística quanto a Probabilidade se originaram da resolução de problemas: a primeira, associada à gestão de questões governamentais, no comércio e agricultura, ao passo que a segunda, aos jogos e ao processamento de dados estatísticos. Essas autoras ressaltam que:

as raízes da Estatística estão centradas nas diferentes áreas do conhecimento e essa percepção nos remete à interdisciplinaridade. Para adotar uma abordagem interdisciplinar, talvez seja necessária uma revisão da prática docente, pois não bastará ao professor o domínio do conteúdo de sua disciplina; será necessário investigar os assuntos de outras áreas e integrar conceitos, procedimentos e metodologias, uma vez que o trabalho interdisciplinar não deve se limitar à integração de conteúdos programáticos das disciplinas [...] Para a eficácia do trabalho pedagógico interdisciplinar, acreditamos ser necessário o desenvolvimento de um projeto educacional mais abrangente, centralizado no trabalho em equipe. (LOPES; MEIRELLES, 2005, p.3).

Tais ideias vão ao encontro da nossa proposta de associar resolução de problemas na Estatística à aprendizagem baseada em projetos interdisciplinares, por meio de práticas cooperativas em contextos realistas, ao mesmo tempo em que contempla muitos dos pressupostos da emergente nova base curricular brasileira.

Durante a realização desta pesquisa, acompanhamos atentamente as discussões que resultaram na elaboração e homologação da Base Nacional Comum Curricular – BNCC (BRASIL, 2018). Tanto nas versões preliminares quanto na final, a abordagem por meio de projetos sempre esteve presente. O documento dedicou à Probabilidade e à Estatística uma das cinco unidades de conhecimento, na tentativa de assegurar, graças a seu caráter normativo, sua presença em todos os bimestres letivos, desde a educação infantil até o último ano do ensino médio. Além disso, buscou redistribuir melhor seus conteúdos, o que, ao nosso ver, está de acordo com os apontamentos teóricos deste trabalho.

Dessa forma, consideramos relevante direcionar nossos esforços para a investigação das concepções estatísticas mobilizadas pelos alunos envolvidos na resolução de problemas após o desenvolvimento de projetos, considerando a perspectiva da presença dessa metodologia de ensino nos currículos nacionais, com base na homologação da BNCC (BRASIL, 2018).

1.2 Objetivo geral

O objetivo deste estudo é identificar concepções mobilizadas por alunos do ensino médio ao resolverem problemas relacionados à Estatística, após o desenvolvimento de projetos nessa área.

1.3 Objetivos específicos

A fim de alcançar o objetivo geral, estabelecemos os seguintes objetivos específicos:

- a) identificar o que os alunos do ensino médio sabem sobre medidas de tendência central, medidas de dispersão, tabelas de distribuição de frequência e gráficos estatísticos;
- b) identificar os conhecimentos em ação de alunos do ensino médio, mobilizados na resolução de problemas, envolvendo tabelas estatísticas, gráficos estatísticos, medidas de tendência central e medidas de dispersão;
- c) evidenciar os operadores mobilizados pelos alunos na resolução de problemas estatísticos;
- d) analisar as estruturas de controle mobilizadas pelos alunos na resolução de problemas estatísticos.

A partir disso, pretendemos analisar as concepções dos alunos participantes da pesquisa, oriundas de diferentes conhecimentos mobilizados.

1.4 Metodologia de pesquisa

Neste trabalho, foi realizada uma pesquisa qualitativa com base em Creswell (2010), a qual é, mais precisamente, um estudo de caso, sob a perspectiva de Ponte (2006).

Creswell (2010) elenca algumas características fundamentais para definir esse método de pesquisa, reconhecidas em nossa investigação. Segundo ele, os dados são geralmente coletados em ambiente natural, ou seja, no campo e no local em que os participantes convivem. Em nosso caso, na escola em que os sujeitos de pesquisa estudam há anos, durante o período regular de aula.

O pesquisador os recolheu pessoalmente, por meio de análise de documentos, observação de comportamento e/ou entrevistas com os participantes. Nós coletamos

diretamente de nossos alunos — que concordaram em participar —, por meio de registro escrito e audiogravado, respondendo a um questionário individualmente e resolvendo problemas em duplas ou trios. Realizamos, ainda, entrevistas com o recurso da filmagem, as quais foram descartadas, devido ao grande volume de informações e a inviabilidade de sua análise neste momento, mas que serão objeto de estudos futuros.

Além disso, de acordo com Creswell (2010), a análise de dados é realizada em um processo indutivo e, por isso, os pesquisadores criam seus próprios padrões, categorias e temas, respeitando os limites estabelecidos pelo seu quadro teórico, muitos dos quais não são previamente definidos e emergem do processo analítico. Essa característica pode ser observada em nossa pesquisa sobretudo na análise das respostas dos alunos, segundo o modelo cK\$, devidamente apresentado no capítulo 3 desta tese.

Na pesquisa qualitativa, segundo esse autor, os significados manifestos pelos participantes são valorizados, em detrimento aos significados previamente trazidos pelo pesquisador. Nós mantivemos o foco nos significados peculiares que os participantes conferiram às suas experiências de aprendizagem.

O projeto é emergente na pesquisa qualitativa. O plano inicial da pesquisa não deve ser rigidamente prescrito, de forma que a pesquisa é continuamente revista e modificada de acordo com as situações que emergem do contexto. Portanto, é necessário que o pesquisador seja flexível e adapte-se às condições de pesquisa oferecidas. No caso deste estudo, prevíamos a filmagem dos alunos em diversos momentos, mas não obtivemos autorização unânime. Também registramos entrevistas com alguns estudantes voluntários, porém nos demos conta de que não atendiam nossos objetivos principais. Assim, os revimos e redirecionamos nossa pesquisa, ação comum nas pesquisas qualitativas, de acordo com Creswell (2010).

Outra característica desse método é o reconhecimento de uma lente teórica direcionando a análise. Isto é, os pesquisadores qualitativos enxergam os problemas estudados com base em uma perspectiva metodológica e teórica bem fundamentada, o que, em nosso caso, consiste na Análise Estatística Implicativa — ASI² e Análise de Similaridades, além da Teoria das Concepções — modelo cK\$.

² Sigla francesa para *Analyse Statistique Implicative*.

Ademais, a pesquisa qualitativa é uma forma de investigação predominantemente interpretativa, a qual não pode ser totalmente dissociada de suas origens, história, contextos e entendimentos prévios. Portanto, esta pesquisa de Doutorado se volta para o passado, estando intimamente ligada àquela desenvolvida durante o Mestrado — cujo foco foi a abordagem por meio de projetos —, e, ao mesmo tempo, para o futuro, na medida em que considera as potencialidades desse método de ensino e a possível disseminação da abordagem da Estatística através de projetos, por estar previsto na BNCC (BRASIL, 2018), que já impacta diversas revisões curriculares em todo o Brasil.

O relato holístico, segundo Creswell (2010), também se faz presente na pesquisa qualitativa, uma vez que ela reflete um quadro amplo, complexo e multifacetado, o qual não se esgota em uma dissertação ou tese, propiciando, geralmente, desdobramentos pelo pesquisador inicial ou por outros que acessaram seus resultados. Desse modo, por reunir todas essas características, consideramos nossa pesquisa como qualitativa, na perspectiva desse autor, muito embora concordemos com o autor de que tais características não são exclusivas desse tipo de pesquisa, podendo, também, ser encontradas em quantitativas e mistas.

Dentro do grande rol de pesquisas qualitativas, enquadramos a nossa na categoria de estudo de caso, que, de acordo com Ponte (2006), pode seguir uma, de duas perspectivas: a pragmática, que visa apresentar um sentido global do objeto; ou a interpretativa, que busca um maior entendimento a respeito do objeto na visão dos sujeitos da pesquisa. Tratando de concepções, na perspectiva de Nicolas Balacheff (1995, 2001, 2002), optamos por esse segundo caminho, uma vez que se trata de fenômenos estritamente individuais.

Na Educação Matemática, segundo Ponte (2006, p. 3), essa opção metodológica vem sendo utilizada para “investigar questões de aprendizagem dos alunos bem como conhecimento das práticas profissionais de professores”. O autor ainda acrescenta como objetivos do estudo de caso “programas de formação inicial e contínua de professores, projetos de inovação curricular, novos currículos, etc.” (ibid.).

Para este estudo, a entidade em questão é uma escola pública de Santo André - SP, cidade da região metropolitana de São Paulo. O perfil dos alunos e de seus professores, o material didático utilizado e o currículo comum nos permitem considerá-la uma unidade escolar representativa da rede pública de ensino do Estado de São Paulo.

Em relação ao estudo de caso, Robert Yin (2015) afirma que não se pergunta “O quê?”, mas sim “Como?” e “Por quê?”. Ponte (2006) assevera que os estudos de caso buscam compreender o como e os porquês com base em uma orientação teórica bem definida, a qual, aqui, foi a Teoria das Concepções. Este teórico também enfatiza a necessidade de se definir com clareza o objeto de estudo, representado neste trabalho pelas concepções dos alunos sobre conteúdos estatísticos abordados ao longo do desenvolvimento de um projeto multidisciplinar.

Segundo Yin (2015), esses estudos podem ser classificados como exploratórios, descritivos ou analíticos. Neste último, o estudo problematiza seu objeto, construindo e desenvolvendo uma teoria ou confrontando uma já existente. Ainda, Ponte (2006) acrescenta que são os estudos de cunho analítico que proporcionam maiores avanços à Educação Matemática. Nesse sentido, acreditamos que nossa investigação se aproxima desta última categoria.

Além disso, Ponte (2006) destaca a evidência de características essenciais do caso, apresentadas em nossa análise, bem como uma indicação clara de que a pesquisa acrescenta algo significativo às noções já existentes na área. Quanto a isso, acreditamos que nossas investigações trazem contribuições para a Educação Estatística por configurar uma das poucas pesquisas realizadas nesse campo à luz da Teoria das Concepções. Uma busca realizada no Banco de Teses da Capes detectou apenas quatro pesquisas com as palavras-chave “Teoria das Concepções”, “Teoria cK Φ ” ou “Modelo cK Φ ”. Dessas quatro, apenas três relacionadas à Educação Matemática: Melo (2010), Oliveira (2010) e Novaes (2011). Na Biblioteca Digital Brasileira de Teses e Dissertações – BDTD, encontramos 22 trabalhos, mas apenas três, os mesmos, a respeito da teoria de Balacheff (1995, 2001, 2002). O Sistema de Publicação Eletrônica de Teses e Dissertações – TEDE da PUC-SP, na mesma busca, detectou três trabalhos: Oliveira (2010), Novaes (2011) e um terceiro, que tratava de concepções de professores sobre violência nas escolas, sem relação direta com o nosso marco teórico. Internacionalmente, só encontramos outras teses afins na França e vamos apresentá-las em nossa revisão da literatura.

1.5 Procedimentos metodológicos

Em nossa pesquisa, analisamos os conhecimentos estatísticos prévios dos alunos antes de participarem da realização de investigações em uma abordagem por meio de projetos. Além disso, averiguamos as concepções sobre Estatística

mobilizadas por estudantes na resolução de problemas após a realização de projetos estatísticos — desde a escolha do tema da pesquisa, planejamento e coleta de dados até a divulgação dos resultados. Nossos dados foram coletados em duas etapas, descritas a seguir.

Na primeira, durante o mês de junho de 2018, todos os 114 alunos dos anos finais do Ensino Médio de uma escola pública paulista, com idades entre 16 e 19 anos, responderam a um questionário composto por 29 questões sobre Probabilidade e Estatística. Descartando os parcialmente respondidos ou com respostas anuladas, ficamos com 86 questionários válidos (ver Apêndice 1).

Vale ressaltar que o Currículo do Estado de São Paulo (SÃO PAULO, 2012) prevê, para o ensino médio, o ensino de Estatística em apenas um bimestre do segundo semestre do terceiro e último ano do ensino médio. Portanto, os alunos responderam às questões antes de estudarem o assunto na escola, mesmo que alguns, eventualmente, tenham tido contato com a Estatística em outras disciplinas no ensino fundamental, como Ciências, Educação Física ou Geografia, ou em disciplinas como Biologia, Física ou Sociologia no ensino médio.

Também não podemos desconsiderar seu contato com a Estatística através das mais diversas mídias. No ano em que nossa coleta foi realizada, houve eleições majoritárias no Brasil e as primeiras pesquisas eleitorais já eram regularmente apresentadas em jornais, programas de rádio e de televisão e internet.

Além disso, devemos esclarecer que o referido currículo só considera a Estatística Descritiva, ou seja, não há Estatística Inferencial no ensino médio nessa rede de ensino. Nosso objetivo foi, justamente, avaliar os seus conhecimentos estatísticos.

Para atingir o objetivo de avaliar os conhecimentos estatísticos dos alunos, optamos pela análise estatística de dados multidimensionais, a qual nos permite organizar e sintetizá-los, evidenciando as relações entre os agentes envolvidos nos fenômenos estudados (ORÚS; ZAMORA; GREGORI, 2009). Saddy Ag Almouloud (2005) ressalta que, nas pesquisas qualitativas da área da educação, são comuns as análises estatísticas, uma vez que permitem sintetizar e organizar os dados multidimensionais no reconhecimento das variáveis estatísticas e didáticas, os fatores envolvidos, suas relações e sua hierarquia bem como evidenciar as relações interpessoais de estudantes e professores em situação de resolução de problemas.

Assim, os 86 questionários foram analisados com o auxílio do *software* CHIC — o qual atendeu bem nossas necessidades —, por possibilitar extrair informações de um conjunto de dados, cruzando sujeitos e atributos, regras de associação entre variáveis, indicando o índice de qualidade de associação além de representar uma estruturação dessas variáveis, como afirmam Raphaël Couturier e Régis Gras (2005) e Gras et al. (2013). Por ser capaz de revelar elementos significativos sobre as reflexões tanto de professores quanto de alunos acerca das questões relacionadas à aprendizagem e prática pedagógica, Maria Elisabette Prado (2002, p. 14) considera a utilização desse *software* fundamental. Segundo a autora, o *software* CHIC “propiciou identificar relações inesperadas, as quais desvelaram novas compreensões sobre o processo de aprendizagem” (ibid.).

O *software* gerou grafos de similaridade, de implicação e de coesão das relações entre as variáveis identificadas pelo instrumento de coleta de dados, com base no cruzamento das mesmas. A multidimensionalidade permitiu, dessa forma, que várias características (variáveis) relativas a um mesmo grupo de sujeitos fossem consideradas, simultaneamente, na análise desses sujeitos quanto a um tema específico — em nosso caso, conhecimentos sobre Estatística.

Em relação ao questionário, o aplicamos em sala de aula, durante 50 minutos, e o organizamos em 3 partes. A primeira delas, composta por 1 questão aberta e 10 questões objetivas (Q1 a Q11, ver Apêndice 1), buscou identificar os conhecimentos dos alunos sobre Estatística, os quais contemplam o modelo de letramento estatístico proposto por Gal (2002), apresentado no capítulo 3 deste trabalho.

Desse modo, tendo como base os modelos de letramento estatístico (ibid.), a questão Q1 procurou identificar conhecimentos e concepções dos alunos sobre a Estatística — seu objetivo, seu método e sua relação com as demais ciências —; avaliou habilidades de letramento referentes ao domínio da língua natural, conhecimentos estatísticos e questionamento crítico, acerca do elemento de conhecimento de Gal (2002); e, quanto aos elementos disposicionais, permitiu que examinássemos crenças, atitudes e postura crítica dos estudantes. As questões Q2, Q3, Q4 e Q5 avaliaram o entendimento do aluno sobre sua própria aprendizagem em Estatística, as questões Q6 a Q9 consideram conhecimento de contexto e questionamento crítico bem como suas crenças, e as questões Q10 e Q11 abordaram crenças e atitudes dos estudantes a respeito da produção do conhecimento científico

em Estatística, das etapas do ciclo investigativo e da confiabilidade das investigações estatísticas.

Destacamos, ainda, que foi inclusa a variável suplementar 'sexo', a qual, embora não fizesse parte de nenhuma das 29 questões de nosso instrumento de coleta de dados, pôde ser apurada no cabeçalho dos questionários, no campo de identificação do aluno.

A segunda parte, constituída por 13 questões objetivas — Q12 a Q24, ver Apêndice 1 —, teve o intuito de detectar os conhecimentos e indícios de letramento dos alunos a respeito da Probabilidade. No entanto, no desenvolvimento da pesquisa, optamos, por focar nas perguntas que abordassem apenas a Estatística, deixando essa análise para futuros estudos.

Como esta parte do questionário deu ênfase ao letramento probabilístico, as questões Q12 a Q15 consideraram o entendimento que o aluno possuía sobre sua própria aprendizagem na área. As 9 questões seguintes, de Q16 a Q24, investigaram os conhecimentos sobre ideias básicas de Probabilidade, tais quais variação, aleatoriedade, independência, previsibilidade e incerteza, bem como o questionamento crítico. Em termos de elementos disposicionais, destacamos crenças, atitudes e sentimentos quanto à incerteza e risco.

Por fim, a terceira parte trazia 5 questões objetivas — Q25 a Q29, ver Apêndice 1 — extraídas das avaliações do Exame Nacional do Ensino Médio – ENEM — o qual possibilita o ingresso, bolsa e/ou financiamento de estudos no ensino superior brasileiro — dos últimos anos, de modo que permitiu investigarmos os conhecimentos estatísticos dos alunos referentes aos conteúdos específicos de Estatística, previstos no Currículo do Estado de São Paulo (SÃO PAULO, 2012). Isto é, pretendeu indicar conhecimentos matemáticos e estatísticos, com foco na leitura e interpretação de tabelas e gráficos estatísticos. No entanto, o nível das habilidades de letramento exigidas foi o básico.

A primeira pergunta, única questão aberta do nosso questionário, não foi respondida por cerca de 16,3% dos alunos. As demais foram respondidas pelos 86 estudantes. Para nossa variável suplementar, sexo, obtivemos como resposta 51,2% de homens e 48,8% de mulheres. Suas idades variam de 16 a 19 anos, prevalecendo 17 anos (82,6%), idade esperada para um aluno concluinte do ensino médio que nunca tenha sido reprovado em seu percurso escolar.

Em relação à segunda etapa da coleta de dados, 4 grupos de alunos (2 trios e 2 duplas) resolveram, em novembro de 2018, 3 problemas de Estatística, semelhantes — quanto à forma e conteúdo — àqueles solucionados por eles mesmos no desenvolvimento de seus projetos de pesquisa, já mencionados por nós.

A resolução de tais problemas durou de 1 a 3 sessões de 100 minutos cada, no máximo, em 3 dias distintos da mesma semana. Essas sessões foram registradas por meio de produção escrita e de gravações de áudio das interações entre os elementos dos grupos, tendo sido analisadas à luz da Teoria das Concepções (ver Anexo 2 e Apêndice 2).

2 REVISÃO DA LITERATURA

Para a composição desta seção, consultamos o Banco de Teses e Dissertações da CAPES³, o Sistema de Publicação Eletrônica de Teses e Dissertações – TEDE da PUC-SP⁴, a Biblioteca Digital Brasileira de Teses e Dissertações – BDTD⁵, bancos de teses e dissertações de outras universidades brasileiras bem como de universidades espanholas, americanas, inglesas e francesas, além de sites de pesquisa internacionais — como *Publimath*⁶ e *Recherche en D.D.M.*⁷ em busca de pesquisas envolvendo a abordagem por meio de projetos em Educação Estatística.

Também procuramos por dissertações e teses que utilizaram, em seus quadros teóricos, a Teoria das Concepções. Acreditamos que nossa tese se encontre na intersecção dessas duas áreas de investigação. Contudo, constatamos que existem poucas pesquisas realizadas utilizando a Teoria das Concepções e nenhuma delas tratou da Estatística na abordagem por meio de projetos.

Em nossa revisão de literatura, apresentaremos dissertações e teses, nacionais e internacionais, que investigaram o ensino e a aprendizagem de Estatística por meio de projetos, situadas no período entre a publicação dos Parâmetros Curriculares Nacionais – PCN (BRASIL, 1997, 1998, 2000) — os quais introduziram a área de Tratamento da Informação no currículo nacional — e a homologação da versão final da Base Nacional Comum Curricular – BNCC (Id., 2018) — que desencadeou o início da implantação da Estatística e Probabilidade (um dos cinco grandes campos de estudo da Matemática) nos currículos regionais, desde a educação infantil até os anos finais do ensino médio.

2.1 Estatística e a abordagem por meio de projetos no ensino médio

Nesta seção, apresentaremos pesquisas envolvendo a Educação Estatística, realizadas na mesma perspectiva da nossa investigação, a resolução de problemas após o desenvolvimento de projetos, com alunos do mesmo segmento de ensino de nosso trabalho.

³ Disponível em: <https://catalogodeteses.capes.gov.br/catalogo-teses/#/>.

⁴ Disponível em: <https://tede.pucsp.br/>.

⁵ Disponível em: <http://bdtd.ibict.br/vufind/>.

⁶ Disponível em: <http://publimath.irem.univ-mrs.fr/>.

⁷ Disponível em: <https://ardm.eu/recherche-en-d-d-m/laboratoires-ed-masters/>.

Uma dissertação que atraiu nossa atenção foi “A Educação Estatística em um ambiente de modelagem matemática no ensino médio”, de Luzinete Mendonça (2008), por se tratar da primeira pesquisa brasileira feita na perspectiva da abordagem por meio de projetos no ensino médio. Embora fale sobre modelagem matemática, a autora assumiu como filosofia norteadora de seu estudo o trabalho por projetos:

No presente estudo também se acredita que o trabalho de projeto deve ser considerado uma filosofia ou perspectiva pedagógica, por se entender o processo educativo como abrangente, globalizado e formador, o que está de acordo com a visão freiriana, a qual compreende a educação como um processo que se realiza no contato do homem com o mundo. (MENDONÇA, 2008, p. 47)

Na tentativa de responder à sua questão de pesquisa, “A Modelagem Matemática pode contribuir para o ensino da Estatística no Ensino Médio?”, realizou uma pesquisa qualitativa do tipo estudo de caso. Apesar das aparentes vantagens dessa metodologia de ensino, a autora indica que parece existir uma lacuna entre o ensino de Estatística de forma contextualizada — como preconiza boa parte das pesquisas nacionais e internacionais — e as práticas realizadas em sala de aula.

Mendonça (ibid.) afirma que, apesar de a Pedagogia de Projetos ser constantemente citada nos documentos oficiais de orientação curricular como uma metodologia de ensino favorável à aprendizagem em ambientes cooperativos — desenvolvendo autonomia, consciência crítica e responsabilidade nos alunos —, as pesquisas nacionais apontam não ter sido observada sua prática efetiva.

Ainda, uma das dificuldades identificadas pela autora refere-se às habilidades de linguagem dos alunos, o que comprometeu a formulação de suas questões de pesquisa e afetou a qualidade do produto de seus projetos. Outro problema encontrado diz respeito à organização curricular.

Mauren Porciúncula e Suzi Samá (2014, p. 185), em um estudo sobre aprendizagem estatística com projetos, destacam que a “falta de tempo e de cooperação entre os alunos, eventualmente, leva alguns grupos a optar por uma divisão de tarefas”. Tal erro conduziria a uma abordagem empobrecedora de pesquisa estatística, e, segundo as autoras, o emprego de tecnologias digitais poderia otimizar tempo e espaços, embora não descarte “a necessidade de interação pessoal no processo educacional”.

Nesse sentido, ressaltamos a discussão de Mendonça (2008) sobre o problema de tempo versus cumprimento do programa curricular. Para ela, um único bimestre —

como apresentado na proposta curricular paulista — não foi o suficiente para o desenvolvimento de uma investigação estatística em um ambiente de modelagem, sobretudo considerando apenas as aulas da disciplina de Matemática no último bimestre.

Buscando uma solução para esse problema, propusemos a parceria entre professores, mobilizando diversas disciplinas escolares, muitas vezes até mesmo em aulas compartilhadas (GIORDANO, 2016). Em nossa pesquisa, apontamos para a possibilidade de ganho cronológico em trabalhos multidisciplinares bem como o auxílio das aulas compartilhadas de Língua Portuguesa e Matemática aos alunos em suas dificuldades de escrita, seja na elaboração dos instrumentos de coleta de dados — dificuldades essas já apontadas por Mendonça (2008) —, seja no relatório final de seus trabalhos.

O segundo estudo que chamou nossa atenção foi a dissertação de Mario Santana (2011), em que foram observadas as características do letramento estatístico manifestadas em alunos do 3º ano do ensino médio ao realizarem o ciclo investigativo de pesquisa como base para a promoção da Educação Estatística.

Nesse trabalho, Santana (ibid.) encontrou convergências entre o ciclo investigativo e o trabalho com projetos em ambientes de modelagem matemática e, de acordo com suas considerações, tal proposta viabiliza o letramento estatístico, na concepção de Gal (2002). O mesmo referencial de letramento estatístico do qual o autor partiu foi utilizado em nossas pesquisas — tanto na de 2016 quanto nesta.

Em seu estudo de caso, Santana (2011) partiu da perspectiva da Educação Matemática Crítica, sob o enfoque da Ciência, Tecnologia e Sociedade – CTS, para identificar as possíveis contribuições do ciclo investigativo/modelagem/projeto para o letramento estatístico. Embora tenha utilizado os termos “modelagem” — assim como Mendonça (2008) — e “ciclo investigativo”, o autor tratou da metodologia de ensino por projetos. Dentre os principais resultados de sua pesquisa, destacou o grande envolvimento dos alunos e a importância de sua participação ativa em todas as fases do ciclo para seu letramento estatístico bem como para a compreensão dos procedimentos da investigação científica.

Tanto Santana (2011) quanto Mendonça (2008) tiveram, dentre os seus referenciais de Educação Estatística, os trabalhos de Carmen Batanero⁸. Seguindo a mesma base teórica, nossa terceira bibliografia, também caracterizada como estudo de caso, é uma tese defendida por María Quirós (2012) na Universidad de Granada, na Espanha, sede do grupo de pesquisa de Batanero e colaboradores, os quais também figuraram entre as suas referências.

Intitulada “El aprendizaje estadístico en la educación secundaria obligatoria a través de una metodología por proyectos: estudio de caso em una aula inclusiva”⁹, a tese teve o objetivo de avaliar o desenvolvimento de competências relacionadas à Educação Estatística de estudantes em situação de inclusão da secundária — equivalente ao ensino médio brasileiro. Para tanto, Quirós (ibid., p. 550) destacou as competências estatísticas no âmbito dos conceitos estatísticos, das informações estatísticas e dos gráficos estatísticos e concluiu que “o trabalho desenvolvido em torno de projetos, mesmo reconhecendo o pouco tempo dedicado a eles, incide positivamente na evolução das competências dos alunos”.

Uma das grandes dificuldades dos estudantes, observada pela autora, foi o que chamou de “baixo nível na competência linguística” — também mencionada por Santana (2011) e Mendonça (2008) —, e, segundo ela, a abordagem por projetos auxiliou no aprimoramento da competência comunicativa entre os alunos. Quirós (2012) considerou os resultados dessa abordagem bastante satisfatórios no que diz respeito ao desenvolvimento das habilidades estatísticas em um trabalho colaborativo, tanto nos aspectos cognitivos quanto nos relacionais.

A autora destacou, também, que a metodologia de ensino baseada em projetos é um tema relevante, embora, de modo geral, exista pouca investigação nessa área. Assim, dentre os resultados que apontam na mesma direção do que os seus, Quirós (ibid.) ressalta os estudos de Lynley Hicks Anderman e Carol Midgley (1998) e de Linda Lumsden (1994) sobre a motivação dos estudantes do ensino médio, os quais concordam ser necessário um enfoque interdisciplinar que incentive o trabalho colaborativo para aumentar o estímulo e participação dos estudantes.

Como pontos favoráveis dessa abordagem, aponta quatro elementos:

⁸ Maria del Carmen Batanero Bernabeu é matemática, estatística e professora do Departamento de Ensino de Matemática da Universidad de Granada. Para saber mais sobre a autora e sua obra, acesse: <https://www.ugr.es/~batanero/>.

⁹ Aprendizagem estatística no ensino médio regular por meio de uma metodologia por projetos - estudo de caso em uma sala de aula inclusiva (tradução nossa).

- a divisão do trabalho;
- a ajuda imediata, recebida por parte do companheiro;
- a facilidade de entendimento do companheiro, devido especialmente à igualdade do vocabulário utilizado e ao conhecimento por parte do companheiro sobre como resolver o problema do entendimento;
- a comodidade do trabalho colaborativo em relação ao individual. (QUIRÓS, 2012, p. 557)

Diferentemente das pesquisas apresentadas até agora, a de Eric Dangiό (2014), “O Ensino de Estatística no Ensino Médio através de projetos”, que também utilizamos como base para esta tese, empregou pressupostos metodológicos de Engenharia Didática. Segundo o autor, por meio do trabalho com projetos, o estudante passou de mero expectador a protagonista de sua aprendizagem, contribuindo para a elevação de sua autoestima, graças à relação de confiança e cooperação estabelecida com os pares e com o professor. Nessa perspectiva, “habilidades e conhecimentos estatísticos permitem refletir e formar opiniões críticas, efetivando uma ação reflexiva e garantindo uma educação para a formação de alunos cidadãos” (ibid., p. 80-81).

O autor aponta que suas hipóteses foram validadas ao final da pesquisa através de ferramentas da Engenharia Didática, dentre elas as de que a abordagem por meio de projetos promove o trabalho em grupo, contribui para construir coletivamente o conhecimento e estimula a participação efetiva bem como a frequência dos alunos. É importante lembrar que, de acordo com o Currículo do Estado de São Paulo (SÃO PAULO, 2012), os conteúdos da Estatística são abordados no 4º bimestre do 3º ano do ensino médio, momento em que se verifica a baixa frequência nas escolas, como detectado por vários pesquisadores, dentre eles Mendonça (2008).

Outra autora na qual nos fundamentamos foi Patricia Albani (2015), que conduziu seu estudo de caso no Paraná, estado que não limita o estudo da Estatística ao 3º ano do ensino médio. Assim, trabalhou diretamente com 25 alunos da 1ª e 2ª ano do ensino médio e, indiretamente, com seus pais, sujeitos que raramente participam de pesquisas em nossa área. Em sua pesquisa, o banco de dados utilizado pelos estudantes era constituído dos questionários respondidos por seus pais. Com o uso de planilhas eletrônicas, os estudantes trabalharam os conceitos estatísticos construindo tabelas de distribuição de frequências e gráficos, além de calculando medidas de tendência central e de dispersão.

Albani (ibid.) observou a mudança de postura dos alunos de 1º e 2º ano do ensino médio frente ao saber estatístico e suas aplicações, sobretudo no que se refere

à produção do conhecimento científico através da pesquisa. Ela assevera que seus resultados confirmaram que o ensino de Estatística por meio de projetos motivou o interesse dos estudantes, aprimorando o raciocínio estatístico, bem como propiciou condições para que eles conhecessem uma matemática diferente daquela a qual estavam acostumados.

A pesquisa realizada por Marcio Freitas (2016) com metodologia mista — empregando recursos qualitativos e quantitativos — teve a finalidade de caracterizar elementos que permitissem consolidar uma aprendizagem de Estatística significativa, próxima da realidade do estudante. Para o autor, o papel fundamental do professor, nesse contexto, consiste em estimular o estudante a se tornar um pesquisador e a desenvolver autonomia crítica, construindo e reconstruindo seu conhecimento. Uma de suas conclusões foi a de que “estudar Estatística a partir dos dados coletados pelos próprios alunos [...] traz uma significativa contribuição para o aprendizado do aluno” (ibid., p. 49). Isso vai ao encontro dos resultados obtidos por Batanero e Díaz (2011), nossa principal referência ao falar de projetos.

Outra literatura utilizada por nós foi Mariana Barberino (2016), que, em sua dissertação “Ensino de Estatística através de Projetos”, investigou o desenvolvimento de conceitos estatísticos a partir do trabalho com projetos em uma escola da rede estadual paulista. A pesquisa foi um estudo de caso com alunos concluintes do ensino médio como sujeitos, e sua análise teve por referencial teórico a pedagogia freiriana e a psicologia histórico-social de Lev Vygotsky, para analisar as contribuições do uso de projetos no ensino de Estatística. A autora concluiu que, ao coletar os dados e analisá-los, os alunos participam efetivamente da produção do conhecimento científico, particularmente do estatístico, elaborando gráficos e tabelas, familiarizando-se com a leitura de informações e aprimorando a percepção e o questionamento crítico.

Assim, julgamos relevante para nossa pesquisa esse trabalho, por ter utilizado o mesmo referencial da Estatística por projetos, de Batanero e Díaz (2011), e de letramento estatístico de Gal (2002). Além disso, ainda que sua análise tenha partido de um quadro teórico diferente de todos os citados até aqui, chegou a conclusões muito próximas. Dentre suas inferências, destacamos a importância da apresentação dos resultados de pesquisa pelos alunos bem como sua autocrítica durante todo o processo, etapa considerada fundamental para o letramento estatístico, no que diz respeito às habilidades de linguagem e postura crítica.

Ademais, em 2016, realizamos um estudo de caso no qual buscamos responder à questão de pesquisa “Que contribuições de uma abordagem da Estatística por meio de projetos podem ser identificadas no desenvolvimento do letramento estatístico de alunos do Ensino Médio?”. O objetivo principal de nossa pesquisa foi estudar as possíveis contribuições da abordagem da Estatística por meio de projetos de pesquisa, empreendidos por estudantes do 3º ano do ensino médio, para seu letramento estatístico, muito próximo do estudo de Barberino (2016). Entretanto, nosso marco teórico — a Teoria das Situações Didáticas – TSD — foi diferente, assim como seus objetivos específicos de analisar as contribuições do trabalho através de projetos para o desenvolvimento e aprendizagem de conceitos estatísticos; analisar os tipos de quebra de contrato didático no desenvolvimento do projeto bem como seus efeitos sobre a construção do letramento estatístico; e avaliar os níveis de letramento, segundo Gal (2002), alcançados pelos estudantes do 3º ano do ensino médio.

Em nosso trabalho, destacamos as conexões entre Língua Portuguesa e a Matemática por meio de aulas compartilhadas. Tal abordagem, norteada pelos pressupostos da Análise Exploratória de Dados – AED, mudou, de forma notável, as relações entre professor, aluno e saber, típicas do contrato didático, como é caracterizado na TSD, promovendo, assim, maior autonomia por parte dos estudantes no desenvolvimento de suas pesquisas. Os resultados revelaram que a abordagem favorece o desenvolvimento do letramento estatístico e gera condições para uma quebra de contrato didático, preparando-os para os desafios futuros de suas vidas, além da apropriação de fundamentos do ciclo investigativo de pesquisa.

A pesquisa mais recente que encontramos sobre a Educação Estatística e a abordagem por meio de projetos foi a dissertação de Alyson Oliveira (2019), “Práticas pedagógicas no Ensino Médio: por uma Estatística crítica e contextualizada”, a qual teve o objetivo de analisar como os conteúdos de Estatística são trabalhados em sala de aula nesse nível de ensino e como são estabelecidas as relações desses conteúdos com as informações presentes no cotidiano dos alunos.

Através de uma pesquisa etnográfica educacional de abordagem qualitativa sob o enfoque da Educação Estatística Crítica, Alyson Oliveira (ibid.) concluiu — assim como os pesquisadores anteriormente citados — que, embora os estudantes consigam estabelecer algumas relações entre a Estatística e o cotidiano, a forma como ela vem sendo abordada em sala de aula não possibilita que compreendam os

conceitos de forma crítica. Assim, o autor defendeu “a instituição de uma Educação Estatística Crítica nos currículos educacionais” para propiciar aos alunos

maior engajamento quanto às questões sociais, culturais e políticas presentes em seus cotidianos, podendo enxergá-las de modo global, e dando atenção às suas especificidades, relações e variações, podendo interferir de forma crítica. (ibid., p. 200)

Apesar de todas essas pesquisas terem objetivos diversos, chegaram a conclusões muito próximas quanto à ineficácia do material didático utilizado pelos alunos; ao seu baixo nível de letramento estatístico; à quase ausência de experiências prévias dos estudantes relacionadas à aprendizagem de Estatística; e ao potencial para o desenvolvimento do letramento e do pensamento estatísticos da abordagem por meio de projetos.

Dessa forma, acreditamos que a abordagem com projetos modifica as concepções que os alunos mobilizam na resolução de problemas de Estatística após o desenvolvimento desses. Nessa perspectiva, consideramos adequada a adoção, em nosso quadro teórico, da Teoria das Concepções e realizamos um levantamento sobre as pesquisas realizadas sob esse marco teórico.

2.2 Pesquisas desenvolvidas na perspectiva da Teoria das Concepções

Como já mencionamos, em nossa revisão da literatura, buscamos pesquisas no Brasil, nos EUA, na França e na Espanha que abordassem o ensino e a aprendizagem estatística por meio de projetos sob o ponto de vista da Teoria das Concepções, mas nenhuma foi encontrada.

Sendo assim, apresentamos neste capítulo algumas pesquisas nacionais e internacionais que utilizaram o marco teórico da Teoria das Concepções, ainda que tenham investigado objetos matemáticos de outras áreas, sobretudo da Geometria.

Uma das primeiras teses a utilizar o modelo cK Φ foi “La rotation: Approche cognitive et didactique – une étude de cas au Mali”, de Mamadou Sangaré, em 2000. Muito embora tenha sido publicada pela *Université du Mali*, foi conduzida sob orientação de Nicolas Balacheff, na perspectiva da didática francesa. O autor, como grande parte dos investigadores que utilizam a Teoria das Concepções, focou em um objeto matemático da Geometria — mais especificamente no tema “ensino e aprendizagem da rotação com alunos do 10^{ème} au lycée”. Sua pesquisa, caracterizada

como um estudo de caso, foi realizada no contexto de uma reforma curricular nacional de Mali, iniciada em 1998.

Assim como nós, Sangaré (2000) buscou identificar as concepções de alunos em um momento de transição curricular, perspectiva teórica de nosso interesse. O estudo teve quatro objetivos: definir o nível de conhecimento sobre transformações geométricas de estudantes do 2º ciclo do ensino fundamental do país; determinar as relações entre figura e transformação no ensino; avaliar o desenvolvimento das concepções sobre rotações na transição do 9º para o 10º ano do ensino fundamental; e determinar quais os meios didáticos disponíveis para promover tal desenvolvimento.

O motivo de essa tese ter atraído nossa atenção se deve à análise das mudanças de concepções realizadas pelo autor, o qual observou que uma concepção pode ser considerada como uma das primeiras noções da didática francesa. No entanto, até Balacheff abordá-lo, tal conceito não era apresentado de forma clara e explícita, por exemplo, dentro da Engenharia Didática, problema levantado por Michèle Artigue (1990).

Sangaré (2000) destacou, ainda, a diversidade e as contradições que permeiam a concepção a partir do filtro da análise mediadora entre o problema e seus respectivos operadores. Tais contradições resultam da complexidade dos problemas que promovem o desequilíbrio da relação “sujeito-*milieu*”. Em português, essa palavra tem sido traduzida, em alguns casos, como “meio”, mas nós optamos por utilizar a palavra original, em francês, uma vez que não encontramos um equivalente que a caracterize com precisão. José Luiz Freitas (2012, p. 79) designa *milieu* como:

[...] onde ocorrem as interações do sujeito, é o sistema antagonista no qual ele age. É no *milieu* que se provocam mudanças visando desestabilizar o sistema didático e o surgimento de conflitos, contradições e possibilidades de aprendizagem de novos conhecimentos. Num dado *milieu*, em cada momento, assituações didáticas são regidas por um conjunto de informações recíprocas, explícitas ou implícitas, envolvendo alunos, professores e um conteúdo em jogo, denominado contrato didático.

Almouloud (2007) acrescenta que, para desenvolver novas aprendizagens, esse *milieu* deve ser munido de intencionalidade didática. Assim, cabe ao professor, na condição de mediador, criar e organizar um *milieu* no qual estão engajados saberes matemáticos propícios ao ensino e à aprendizagem.

Uma tese que apresenta muitas semelhanças com a de Sangaré (2000) é a de Lucile Vadcard, que também foi realizada em 2000 sob a orientação de Balacheff, com um objeto matemático oriundo da Geometria: a noção de ângulo. A autora foi uma das

precursoras da pesquisa em Educação Matemática sob o enfoque da Teoria das Concepções, definindo concepção como uma instância do conhecimento de um sujeito dentro de uma esfera de práticas. Uma concepção é definida por um domínio de validade, pelos sistemas de representação, pelos operadores e pelas estruturas de controle. Segundo Vadcard (2000), as estruturas de controle, que estavam implícitas na Teoria dos Campos Conceituais de Gerard Vergnaud, ganham um destaque especial na releitura de Nicolas Balacheff.

A autora ainda propõe que as concepções são um modelo coerente com as incoerências e cita as ideias de Efraim Fischbein (1982) ao admitir a coexistência de conhecimentos contraditórios, que contraria a ideia, até então vigente, da construção do conhecimento por adaptação que sugere a não contradição. O modelo das concepções associa cada manifestação do conhecimento a uma esfera de práticas, permitindo explicar a contradição aos olhos de um dado observador. Tal modelo possibilita, então, a preservação do caráter racional dos conhecimentos operatórios.

No Brasil, uma das pesquisadoras pioneiras no campo da Teoria das Concepções foi Iranete Lima (2006), que, em sua tese de Doutorado — também orientada por Balacheff —, investigou a modelização da aprendizagem de noções de simetria de reflexão dos alunos do *quatrième*, equivalente francês ao 8º ano do ensino fundamental brasileiro.

Um de seus objetivos foi o de identificar controles susceptíveis de serem mobilizados por alunos na resolução de problemas de construção e de reconhecimento de figuras simétricas. Por isso, incluímos seu trabalho em nossa revisão da literatura e buscamos identificar operadores e estruturas de controle nas concepções mobilizadas por nossos sujeitos de pesquisa.

Lima (ibid.) optou por inserir a modelagem das concepções dos alunos sobre a noção de simetria ortogonal, com base na identificação da estrutura de controle, considerando a formalização proposta pelo modelo cK Φ dentro da Teoria das Situações Didáticas. Sua premissa foi de que os controles refletem os critérios relacionados à escolha, decisão, adequação e validade de uma ação, realizando um estudo teórico dessa noção de um ponto de vista matemático e didático. Assim, a autora analisou o processo de tomada de decisão dos professores, a fim de promover o avanço dos estudantes na aprendizagem, utilizando os modelos de Níveis de Atividade — desenvolvido por Claire Margolinas — e cK Φ — desenvolvido por Nicolas

Balacheff — como ferramentas de análise. Segundo Lima (2008, p. 15), seu principal interesse foi:

[...] estudar a forma como os professores tomam suas decisões didáticas com o objetivo de levar os alunos a avançar na aprendizagem de um determinado conhecimento, e os elementos que influenciam estas decisões. Para isso, se fez necessário, em um primeiro momento, modelizar conhecimentos de alunos relativos a um objeto matemático, a simetria de reflexão.

Além disso, analisou as concepções dos estudantes, sendo um de seus objetivos: “identificar, a priori, controles susceptíveis de serem mobilizados por alunos na resolução de problemas” (LIMA, 2008, p. 15).

Sua pesquisa corrobora com a hipótese de que a ação racional dos sujeitos em interação com o *milieu* é localmente lógica na perspectiva do observador bem como constata a aprendizagem na passagem de uma concepção para outra. Por fim, Lima (2006) sugere novas investigações, dada a escassez de produção acadêmica nacional nessa área.

Isso levou à pesquisa de Diógenes Melo (2010), seu orientando de Mestrado que, em sua dissertação sobre simetria de reflexão, investigou alguns elementos de concepções mobilizadas por alunos brasileiros ao resolverem questões envolvendo figuras simétricas. Esse estudo foi realizado com 50 estudantes do 9º ano do ensino fundamental de escolas da rede pública — municipal e estadual — de ensino da região do agreste de Pernambuco e utilizou elementos de Engenharia Didática. O autor confirmou resultados apresentados por Lima (2006), como os referentes às estruturas de controle mobilizadas pelos alunos.

Em relação à Estocástica, o trabalho que garantiu nosso interesse e sua consequente inclusão em nosso estudo pelo ineditismo temático foi a dissertação de Priscila Oliveira (2010), “Probabilidade: Concepções construídas e mobilizadas por alunos do Ensino Médio à luz da Teoria das Concepções (cK Φ)”.

Em sua pesquisa de mestrado, utilizando a metodologia do estudo de caso, a autora investigou as concepções probabilísticas construídas e mobilizadas por estudantes do ensino médio quando submetidos a uma aprendizagem baseada na então nova proposta curricular estadual paulista, tendo em seu quadro teórico a Teoria das Concepções (modelo cK Φ).

Para isso, a análise das concepções feita por Priscila Oliveira (ibid.) levou em conta as características probabilísticas apresentadas por Pilar Azcárate (1995, 1996):

não probabilística da realidade, intuitiva, emergente e normativa, em escala crescente de complexidade, nessa ordem.

Ainda, a autora considerou que os livros didáticos adotados naquele momento nas escolas estaduais paulistas, de modo geral, não respeitavam a proposta dos PCN — dentre outras falhas, não estabeleciam relações entre a Probabilidade e a Estatística; propunham exercícios repetitivos e distantes da realidade do aluno; apresentavam uma perspectiva puramente determinística de probabilidade; não exploravam adequadamente a noção de aleatoriedade; e não propiciavam a compreensão de probabilidade condicional.

Para tanto, o foco da investigação de Priscila Oliveira (2010) recaiu sobre a resolução de problemas, por parte dos estudantes, a partir do material apresentado no Caderno do Professor e no Caderno do Aluno, elaborados pela Secretaria da Educação do Estado de São Paulo, mesma rede de ensino em que estudam os sujeitos da nossa pesquisa. A autora também considerou as características das questões, presentes no Exame Nacional do Ensino Médio – ENEM, que versavam sobre Probabilidade, estudadas por Amari Goulart (2007).

Além disso, Priscila Oliveira (2010) elaborou 11 atividades para alunos do 2º e do 3º ano do ensino médio — únicos anos desse nível em que a proposta curricular paulista previa o estudo de Probabilidade e de Estatística, respectivamente —, organizados em duplas ou trios. Dessa forma, dentre os principais resultados observados, confirmando pesquisas anteriores — nacionais e internacionais —, destacou-se o viés da equiprobabilidade e a concepção de probabilidade caracterizada pelo invariante “razão entre parte e todo”.

Enquanto o ENEM exigia do aluno uma concepção normativa de probabilidade, na expectativa de que fosse capaz de avaliar a probabilidade em diferentes contextos, os estudantes apresentaram concepções não probabilísticas da realidade, “não demonstrando conhecimento claro do azar e do sucesso aleatórios”. Ao final, a autora considerou a Teoria das Concepções bastante eficaz no reconhecimento e caracterização das concepções dos seus sujeitos de pesquisa.

Por fim, nos fundamentamos na tese de Diva Novaes (2011), a única das encontradas dedicada, especificamente, à Estatística sob o enfoque da Teoria das Concepções, o que a tornou importante para nosso estudo. Entretanto, o foco de sua pesquisa recaiu sobre as concepções mobilizadas pelos professores, ao passo que a nossa se dedica às concepções mobilizadas pelos alunos.

Novaes (ibid.) encontrou raras pesquisas envolvendo concepções de professores e alunos à luz do modelo cK Φ . Assim, em sua tese de Doutorado, investigou as concepções de docentes da educação básica sobre a variabilidade estatística, mobilizadas durante a execução e gestão de sequências didáticas nessa área. Caracterizada como um estudo de caso, seu trabalho nasceu da constatação de "entraves" na aprendizagem dos estudantes, as quais poderiam estar associadas às dificuldades dos professores em compreender conceitos essenciais à Estatística, como variabilidade.

Desse modo, identificou 16 concepções, sendo 13 referentes ao conhecimento didático e 3 ao conhecimento específico de Estatística, assim sintetizadas pela autora:

Quadro 1 – Síntese das concepções identificadas na análise dos protocolos de observação referentes ao conhecimento específico de Estatística

Concepção Identificada	Invariantes operatórios	
	Operador associado	Controle associado
C _{E1} : Considerar a frequência de uma variável qualitativa como variável quantitativa discreta	Existe a possibilidade de contagem do observado	Se houver número na distribuição, então a variável é quantitativa.
C _{E2} : Identificar a frequência como os valores assumidos por essa variável	A frequência representa o valor da variável	As medidas-resumo podem ser calculadas utilizando-se os valores da frequência.
C _{E3} : A variabilidade nos dados pode ser caracterizada apenas com o valor mais frequente na distribuição.	O mais importante são os dados que apresentam maior frequência.	Olhar dirigido para o valor mais frequente e para os valores centrais.

Fonte: NOVAES, 2011, p. 157.

Segundo Novaes (ibid.), o modelo cK Φ viabilizou uma análise consistente dos dados coletados, contribuindo para a compreensão sobre a natureza das concepções dos docentes e da maneira como eram mobilizadas. A autora sugeriu a revisão curricular nas formações iniciais e continuadas desses profissionais e novas investigações sobre o desenvolvimento das competências dos alunos em Estatística, direção para a qual se encaminha este estudo.

Resumimos as principais contribuições desta revisão da literatura para a nossa pesquisa no quadro a seguir:

Quadro 2 – Principais contribuições das pesquisas analisadas em nossa revisão da literatura

Título	Ano	Autor	Universidade	Principais contribuições
Tese: La Rotation: Approche Cognitive et Didactique – une étude de cas au Mali	2000	Mamadou Sangaré	Université du Mali	Destaque dado às mudanças de concepções dos alunos no processo de aprendizagem de Matemática.
Tese: Étude de la notion d'angle sous le point de vue des conceptions	2000	Lucile Vadcard	Université Joseph Fourier – Grenoble I	Caracterização de conhecimentos contraditórios e incoerências nos conhecimentos dos alunos.
Tese: De la modélisation de connaissances des élèves aux décisions didactiques des professeurs: étude didactique dans le cas de la symétrie orthogonale	2006	Iranete Lima	Université Joseph Fourier – Grenoble I	Análise das concepções mobilizadas pelos alunos na aprendizagem de Matemática bem como os mecanismos de controle utilizados pelos professores.
Dissertação: A Educação Estatística em um ambiente de Modelagem Matemática no ensino médio	2008	Luzinete de Oliveira Mendonça	Universidade Cruzeiro do Sul	Reflexão sobre as dificuldades e limitações de uma abordagem da Estatística por meio de projetos.
Dissertação: A simetria de reflexão: elementos de concepções mobilizadas por alunos do Ensino Fundamental	2010	Diógenes Maclyne Bezerra de Melo	Universidade Federal de Pernambuco	Reflexão sobre as estruturas de controle mobilizadas pelos alunos.
Dissertação: Probabilidade: Concepções construídas e mobilizadas por alunos do Ensino Médio à luz da Teoria das Concepções (cK Φ)	2010	Priscila Glaucete de Oliveira	Pontifícia Universidade Católica de São Paulo	Pioneirismo na abordagem da Estocástica no modelo cK Φ no Brasil e análise da resolução de problemas do livro didático e Caderno do Aluno.
Tese: Concepções de professores da educação básica sobre variabilidade estatística	2011	Diva Valério Novaes	Pontifícia Universidade Católica de São Paulo	Identificação de concepções referentes aos conhecimentos específicos de Estatística dos alunos e professores.
Dissertação: A Educação Estatística com base num ciclo investigativo: um estudo do desenvolvimento do letramento estatístico de estudantes de uma turma do 3.º ano (Ens. Médio)	2011	Mario de Souza Santana	Universidade Federal de Ouro Preto	Apresentação de vantagens e de desvantagens da opção de abordagem da Estatística por meio de projetos.
Tese: El aprendizaje estadístico en la educación secundaria obligatoria a través de una metodología por proyectos: estudio de caso en una aula inclusiva	2012	María Vega Quirós	Universidad de Granada	Análise dos níveis de competência linguística dos alunos e da importância da comunicação entre eles na aprendizagem da Estatística por meio de projetos.
Dissertação: O Ensino de Estatística no Ensino Médio através de projetos	2014	Eric Giovanni Zenatti Dangiό	Universidade Federal de São Carlos	Discussão sobre o protagonismo dos alunos no desenvolvimento de projetos bem como da confiança e autoestima.

Dissertação: Estatística com projetos: uma alternativa de Ensino e de Aprendizagem	2015	Patricia Albani	Universidade Tecnológica Federal do Paraná	Discussão sobre o uso de tecnologias digitais no desenvolvimento de projetos no ensino de Estatística bem como o envolvimento dos pais nesse processo.
Dissertação: Praticando Estatística no Ensino Médio	2016	Marcio Melo Freitas	Universidade de Brasília	Reflexão sobre o papel do professor na abordagem por projetos no ensino da Estatística.
Dissertação: Ensino de Estatística através de Projetos	2016	Mariana Ribeiro Busatta Barberino	Universidade de São Paulo	Valorização da participação do aluno na coleta de dados bem como da divulgação de seus resultados.
Dissertação: O desenvolvimento do letramento estatístico por meio de projetos: um estudo com alunos do Ensino Médio	2016	Cassio Cristiano Giordano	Pontifícia Universidade Católica de São Paulo	Análise das contribuições de uma abordagem da Estatística Descritiva por meio de projetos para o letramento estatístico dos alunos.
Dissertação: Práticas Pedagógicas no Ensino Médio: Por uma Estatística Crítica e Contextualizada	2019	Alyson Fernandes de Oliveira	Universidade Federal de Goiás	Discussão sobre o aprimoramento da criticidade do aluno e do papel do contexto ao desenvolver projetos.

Fonte: o autor, 2020.

Diante desses resultados, julgamos necessário investigar as concepções dos alunos envolvidos no processo de aprendizagem de Estatística, procurando por possíveis mudanças nelas após seu envolvimento na abordagem por projetos.

3 FUNDAMENTAÇÃO TEÓRICA E METODOLÓGICA

O objetivo desta seção é apresentar os aportes teóricos e metodológicos que fundamentam nosso trabalho. Buscamos identificar as concepções estatísticas de alunos após o desenvolvimento de pesquisa por meio de projetos e suas possíveis mudanças. Assim, nos pareceu adequado a adoção, em nosso quadro teórico, do modelo cK Φ .

A mobilização das concepções aqui estudadas se deu no contexto da resolução de problemas de Estatística, após a realização de projetos pelos alunos participantes de nossa pesquisa. A abordagem da Estatística por projetos foi desenvolvida na perspectiva da Análise Exploratória de Dados – AED, marco teórico e, ao mesmo tempo, filosofia norteadora do trabalho investigativo.

Além disso, para analisar os questionários respondidos, em nossa sondagem sobre os conhecimentos prévios dos alunos, utilizamos a Análise Estatística Implicativa – ASI e a Análise de Similaridades.

Esse é, portanto, o tripé de sustentação desta tese.

3.1 Análise Exploratória de Dados

A opção pela AED nos pareceu uma escolha natural, uma vez que sua abordagem da Estatística valoriza a postura investigativa e crítica do aluno e pressupõe uma proposta didático-pedagógica centrada na pesquisa.

Essa abordagem surge com o estatístico americano John Wilder Tukey (1962, 1970), o qual propunha a exploração dos dados e a formulação de hipóteses que pudessem levar a novas coletas de dados e experimentos, sem depender de um modelo estatístico preestabelecido, extrapolando a modelagem e os testes estatísticos tradicionais, com a finalidade de obter dos dados a maior quantidade possível de informações, apontando para modelos plausíveis a serem utilizados em etapas posteriores. Assim, caberia aos investigadores iniciarem sua análise pelo exame dos dados disponíveis, para, somente então, definirem as técnicas mais adequadas na resolução dos problemas.

A AED nos fornece um extenso repertório de métodos para um estudo detalhado dos dados antes de adaptá-los. Tukey (1970) sugeriu a utilização das seguintes medidas-resumo representativas de uma dada amostra: os valores máximo e mínimo, a mediana e os quartis. O estatístico considerava a mediana e os quartis

funções da distribuição empírica mais robustas em relação a outras medidas-resumo tradicionais, como a média e o desvio padrão. Tais medidas-resumo são empregadas na organização e apresentação dos dados, enfatizando uma apreensão mais eficaz das informações subjacentes ao trabalho de pesquisa censitária ou por amostragem.

Essa proposta se destaca no contexto de transição “de um enfoque tecnicista, no qual o aspecto matemático é predominante e os modelos são pré-estabelecidos, para um analítico, no qual se busca construir modelos a partir do estudo de dados observados” (COUTINHO e MIGUEL, 2007, p. 1). Sobre a abordagem pela AED, Carolina Carvalho (2003) ressalta que ela é mais fácil, motivadora e criativa e, acima de tudo, está imbuída do espírito investigativo que caracteriza toda e qualquer produção científica.

Batanero (2001, p. 27, tradução nossa) ressalta que

Anteriormente a este enfoque, a análise dos dados era baseada principalmente no cálculo estatístico (médias, variância, coeficientes de correlação), levando a duas consequências. Em primeiro lugar, diminuiu a importância visual da representação dos dados, atribuindo-a exclusivamente aos cálculos e, em segundo lugar, pensou-se que, para obter conclusões dos dados, era necessário recorrer à inferência (modelo confirmatório).

Segundo a autora, os dados são compostos de duas partes intrinsecamente relacionadas: a “regularidade”, ou tendência, e “desvios”, ou variabilidade. Até o surgimento da AED, o foco dos estudos estatísticos repousava sobre a busca de um modelo capaz de representar fielmente a regularidade das observações. A AED, por sua vez, não tenta impor modelos iniciais, e sim partir dos dados para poder criá-lo.

Como características básicas dessa abordagem, Batanero, Antonio Estepa e Juan Godino (1991) destacam a possibilidade da criação de situações de aprendizagem sobre temas de interesse dos alunos, apoiando-se em representações gráficas que favoreçam a compreensão da variabilidade, a valorização das medidas de ordem — que minimizem eventuais casos atípicos —, o uso de diferentes escalas, além da falta de necessidade de uma teoria matemática complexa, com ferramentas desnecessárias nesse momento.

Batanero (2001) assevera que, a fim de escolher um assunto a ser abordado nas aulas de Estatística, o professor deve levar em conta seu potencial para gerar situações de aprendizagem relacionadas a tópicos de interesse do aluno, o forte apoio nas representações visuais, valorizando o uso de múltiplas representações dos dados,

sobretudo os gráficos, priorizando a discussão entre os estudantes, a argumentação científica, em detrimento do uso de uma teoria matemática complexa.

Para Batanero et al. (1994), a Estatística, tradicionalmente, recebe menos atenção nos currículos escolares do que os outros campos da Matemática. Os autores acrescentam, ainda, que boa parte das pesquisas nessa área foi realizada em situações experimentais em vez de explorar situações escolares, e as realizadas em contextos escolares focalizaram, em sua maioria, o ensino superior. Ao realizar pesquisa estatística em ambiente escolar na educação básica, tentamos responder a essa demanda.

A AED se disseminou como uma filosofia de aplicação da Estatística graças à rápida popularização dos computadores e ao desenvolvimento de *softwares* estatísticos acessíveis, com amplas possibilidades para o tratamento de dados e para suas representações gráficas, segundo Batanero (1999). Para a autora, simplificou a vida de professores e alunos, permitindo a aproximação com outras disciplinas escolares, viabilizando o trabalho em equipe e possibilitando o desenvolvimento de projetos interdisciplinares.

No Brasil, a Estatística está presente no currículo das disciplinas de Matemática na educação básica desde a publicação dos PCN (BRASIL, 1997, 1998, 2000). No entanto, ela apresenta peculiaridades que a distinguem claramente da Matemática, como discutiremos adiante. Até mesmo os comportamentos e as atitudes dos alunos ao desenvolverem projetos que envolvem conhecimentos estatísticos reforçam essa ideia. Embora a Estatística venha conquistando *status* de ciência autônoma, paradoxalmente, vem desaparecendo do currículo oficial dos cursos superiores, como observam Samá, Porciúncula e Silva (2011, p. 2):

São raros os cursos que não contemplam, pelo menos, uma disciplina envolvendo Estatística. No entanto, temos observado a redução da carga horária dessas disciplinas quando não simplesmente eliminadas do currículo. O que pode ser explicado pela dificuldade dos estudantes na compreensão dos conceitos estatísticos.

Como competências estatísticas básicas para que os professores desenvolvam trabalhos com base na Análise Exploratória de Dados, Gal (2002) destaca:

- reconhecer a necessidade de manipular dados;
- saber como produzir dados significativos;

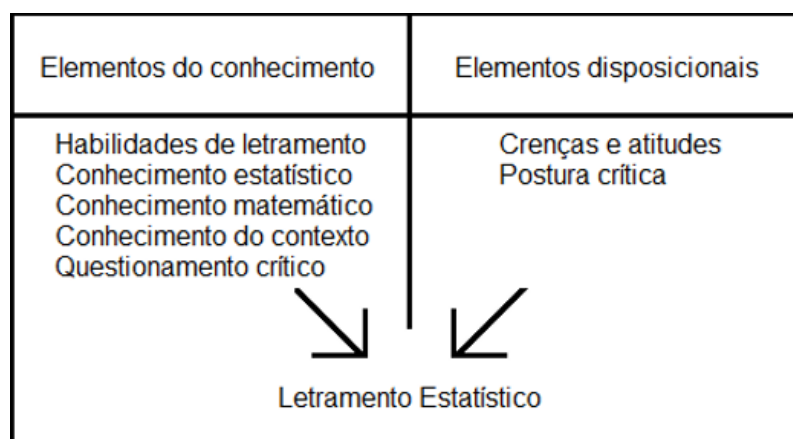
- apresentar familiaridade com os termos e ideias mais elementares da Estatística;
- apresentar familiaridade com seus registros de representação tabulares e gráficas;
- dominar conhecimentos de probabilidade e conhecer métodos de elaboração de Análise Estatística Inferencial.

De acordo com o autor, os conhecimentos estatísticos a serem desenvolvidos pelos alunos serão fruto de suas habilidades quanto a:

- conhecimento estatístico;
- domínio de sua língua natural;
- conhecimento matemático;
- conhecimento do contexto e do mundo;
- sua capacidade de elaborar perguntas frente aos saberes, associados a elementos de disposição, que envolvem sua postura crítica, bem como suas crenças e atitudes.

Gal (2002) ainda aponta os elementos que considera constituintes do letramento estatístico, os quais serão levados em conta na análise dos nossos 86 questionários:

Figura 1 – Modelo de letramento estatístico



Fonte: GAL, 2002, p. 4.

Para ele, o letramento estatístico implica no desenvolvimento de habilidades que permitam ao indivíduo ler, interpretar, produzir e comunicar informações

estatísticas de diferentes maneiras, como por meio de medidas de tendências central, gráficos e tabelas. O indivíduo letrado estatisticamente deve ser capaz de reconhecer a necessidade de obter dados estatísticos; fazê-los; organizar e apresentar informações de natureza estatística, lidando adequadamente com a transnumeração de dados; compreender a variabilidade em determinados contextos; atribuir-lhes significado; e agir de forma crítica e objetiva, tomando decisões acertadas, fundamentadas em dados estatísticos.

Quanto à mobilização do conhecimento de contexto e à criticidade, Batanero e Godino (2001) apontam para as vantagens da abordagem estatística por meio de projetos. Os autores destacam que os professores que ensinam disciplinas como Física, Química ou Biologia tanto no ensino médio quanto no ensino superior, em boa parte dos casos, praticam a alternância das aulas teóricas e resolução de problemas com práticas de laboratório, dando um espaço menor à Estatística. Isso gera dificuldades de aprendizagem e desinteresse dos alunos, que não veem aplicações práticas naquilo que aprendem.

A abordagem da AED, impulsionada pela crescente facilidade de acesso aos recursos tecnológicos nas escolas, oferece alternativas ao que os autores chamam de “ensino de laboratório”. O computador e a internet permitem agilidade no acesso e manipulação de um grande volume de dados, o que auxilia na investigação de problemas das mais diferentes áreas.

Resumindo, Batanero e Godino (Ibid., p. 9) afirmam que o uso de computadores no ensino de Estatística possibilitou

Estudar dados de estudos de caso reais, incorporando a "abordagem por projetos"; adquirir habilidade no manuseio da ferramenta de computador; o entendimento de conceitos e técnicas estatísticas através de simulações e do processo de análise de dados.

Para Batanero e Díaz (2004), os projetos estatísticos motivam os alunos, ao contrário da resolução de exercícios descontextualizados, tão comuns em boa parte dos livros didáticos. As autoras lembram da definição de David Moore (1995), a qual diz que a Estatística é a ciência dos dados, e estes não são apenas números, mas sim números em contexto. Segundo elas, no trabalho com projetos, a ênfase é dada a tarefas que devem ser realistas, como foi exemplificado por Giordano (2016).

Celso Campos (2007) destaca que a Estatística trabalha com conteúdos e com valores geralmente distintos daqueles da Matemática, empregando princípios de

aleatoriedade e incerteza diferentes dos aspectos matemáticos tradicionais lógicos e determinísticos, explorando o espaço da intersubjetividade. Dessa forma, é justo supor que, no ensino da Estatística, o professor emprega estratégias diferenciadas daquelas utilizadas regularmente nas aulas de Matemática.

Celi Lopes, Cileda Coutinho e Saddo Ag Almouloud (2010, p. 48) alertam que:

[...] embora a inserção de Estatística e Probabilidade seja recomendada nas propostas curriculares de Matemática, na maioria dos países do mundo, ainda não tem sido uma prioridade na escola.

Levantamos a hipótese de que tal situação se mantém ainda nos dias de hoje, com base na observação de resultados de pesquisas apresentadas em congressos nacionais e internacionais, como ENEM, SIPEM e ICOTS.

Sobre as vantagens da opção do ensino da Estatística por meio de projetos, Batanero e Díaz (2011) ressaltam a motivação, o desenvolvimento da criticidade e da autonomia. Segundo elas, enquanto os livros didáticos, geralmente, enfatizam as noções técnicas, os projetos permitem contextualizar a Estatística em situações com potencial para despertar um maior interesse dos alunos. Os projetos podem reforçar esse interesse, especialmente se são os estudantes mesmos que escolhem o tema.

Ainda, para Batanero e Díaz (ibid.), essa abordagem permite a aquisição de competências fundamentais para o aluno do ensino médio, tais como:

- competência comunicativo-linguística;
- competência matemática;
- competência de reconhecimento e interação com o mundo físico;
- competência para organizar e apresentar dados;
- competência digital;
- competência social, para o exercício da cidadania;
- competência para “aprender a aprender”;
- competência para questionar, identificar e gerenciar as diversas técnicas e estratégias para lidar com uma mesma situação-problema;
- competência para a conquista de autonomia.

Segundo Campos, Maria Lúcia Wodewotzki e Otávio Jacobini (2013), os professores de Estatística, sobretudo no ensino superior, costumam dar mais ênfase aos aspectos técnicos e operacionais da disciplina e, em contraposição, propõem o

trabalho com projetos. Sobre os hábitos mentais e as habilidades de resolução de problemas necessários ao pensamento estatístico, os autores acrescentam que o trabalho com projetos pode viabilizar o desenvolvimento desses hábitos, pois, nele, o aluno assume a responsabilidade de coletar os dados brutos, analisá-los, interpretá-los e divulgá-los.

Os autores asseveram que os alunos aprendem Estatística fazendo Estatística. Destacam, assim, a importância do trabalho com projetos, em cujo desenvolvimento eles participam de todas as etapas: coleta, organização, apresentação e interpretação de resultados. Durante esse processo, os estudantes elaboram textos estatísticos em diversos formatos, produzindo relatórios, tabelas, gráficos, pareceres etc. Para Campos, Wodewotzki e Jacobini (ibid.), o trabalho colaborativo deve ser incentivado. Nessa proposta, o foco da abordagem é redirecionado do produto para o processo, com valorização da análise e interpretação.

Os alunos devem ser devidamente estimulados à investigação, de modo a ler e escrever textos estatísticos, argumentar, interpretar e analisar, mais do que calcular e desenhar. Ademais, é necessário que os recursos tecnológicos disponíveis sejam mobilizados e as avaliações, orientadas, para o cumprimento de metas previamente estabelecidas — de forma aberta e transparente —, e não para a mera correção de resultados quantitativos obtidos por cálculos e aplicação de fórmulas.

Enfatizando a necessidade de promover tais investigações, em particular, com trabalho por meio de projetos, Gislaine Costa (2012) afirma ser de extrema importância para o professor valorizar os conhecimentos estatísticos dos alunos, pois, só assim, estes assumirão o protagonismo na pesquisa. Porciúncula e Samá (2015) encaram a opção de trabalhar com projetos de aprendizagem não apenas como uma metodologia, mas uma forma de refletir sobre o papel da Estatística na sociedade.

A mudança de postura do docente diante da Estatística vai ao encontro da proposta de trabalho com projetos. Como destacam Porciúncula e Samá (2014), essa proposta requer uma mudança no papel do professor, passando de transmissor de informação a orientador dos processos de aprendizagem dos alunos.

Vanessa Tomaz e Maria Manuela David (2012) definem a interdisciplinaridade como uma combinação entre duas ou mais disciplinas, com vista à compreensão de um objeto comum. Assim, projetos interdisciplinares possibilitam criar pontes entre diferentes áreas do conhecimento humano, além de permitirem uma ampla leitura da realidade.

Consideramos, então, que o trabalho em grupos possa contribuir para o desenvolvimento de projetos de Educação Estatística. Para Joan Garfield (1993), uma forma de o professor motivar o aprendizado ativo em suas aulas é estruturar oportunidades para que os alunos aprendam juntos em pequenos grupos. Em seu artigo sobre grupos cooperativos de aprendizagem, a autora destaca a importância dessa organização no ensino e aprendizagem da Estatística.

Ela ainda ressalta que, embora centenas de estudos, literalmente, tenham sido publicados sobre aprendizagem cooperativa, apenas alguns poucos são especificamente voltados ao ensino da Estatística. Nesse modelo, o papel do professor deve mudar de uma mera fonte de informação para um facilitador de aprendizagem. Cabe evidenciar que, quando Garfield publicou esse artigo em 1993, a internet não era acessível aos alunos como é atualmente, com a mobilidade de um *smartphone* no bolso de vários estudantes de ensino médio.

Em 2013, Garfield apontou que, dentre os muitos benefícios que essa estratégia de ensino e de aprendizagem pode trazer, podemos destacar: uma maior motivação e interesse do aluno; o desenvolvimento de atitudes positivas sobre sua capacidade; o fortalecimento do espírito de equipe; uma melhor comunicação; uma maior responsabilidade do estudante; a otimização do tempo; e o dinamismo nas aulas.

Como destaca a autora, “duas cabeças pensam melhor do que uma” e, ainda que todos no grupo encontrem uma mesma solução, o fazem de formas diferentes. O que enriquece a aprendizagem é a troca de opiniões, não somente sobre o resultado final estar certo ou errado, mas sobre os processos que conduzem a ele. Garfield (2013) sugere a formação de pequenos grupos, com aproximadamente quatro alunos, cumprindo funções distintas — incentivador, moderador, redator das discussões do grupo, redator de relatórios e/ou resumos.

Nesse aspecto, a proposta de Garfield (1993, 2013) se aproxima bastante daquela apresentada por Batanero e Díaz (2004, 2011), desenvolvida sob a perspectiva da AED.

3.2 Análise Estatística Implicativa (ASI) e Análise de Similaridades.

Nesta subseção, apresentamos nosso segundo referencial teórico-metodológico: a Análise Estatística Implicativa e Análise de Similaridades.

Adotamos a ASI em nossa investigação como ferramenta no tratamento de dados por ela nos permitir visualizar, organizar, construir modelos representativos bem como analisar os fenômenos a eles associados (GRAS; ALMOULLOUD, 2003).

O desenvolvimento da ASI foi impulsionado nos anos 1970 pelos trabalhos realizados e orientados por Régis Gras — professor emérito da Escola Politécnica da Universidade de Nantes —, ao passo que sua operacionalização se deu na década de 1990, com a criação e o aprimoramento do auxílio do *software* CHIC – Classificação Hierárquica Implicativa e Coesiva, utilizado para a análise de dados multivariados.

Originalmente concebido por Gras, o *software* foi desenvolvido para o computador por dois de seus orientandos: Saddo Ag Almouloud e Harrison Ratsimba-Rajohn. Atualmente, o CHIC tem sido atualizado e aprimorado por Raphaël Couturier. Em nossas investigações, utilizamos a versão 7.0 de 2014¹⁰.

A importância do CHIC para a ASI é tamanha a ponto de Gras (2015, p. 9) afirmar que

Sem o CHIC, ASI seria meramente objeto matemático cultural especulativo, desprovido de aplicações porque os cálculos seriam intratáveis; sem a ASI, a ferramenta informática de processamentos de cálculos e de gráficos, CHIC não poderia dar sentido aos resultados numéricos e às representações que produz, ele os deixaria empobrecidos e inacessíveis para o controle e a crítica.

A ASI é um método de análise de dados não simétricos que possibilita, com base em um conjunto de dados que cruza uma população de sujeitos ou objetos com um conjunto de variáveis, a extração e estruturação do conhecimento na forma de regras e regulamentos generalizados e, a partir da contingência dessas regras, a explicação e, conseqüentemente, uma certa previsão em diferentes ramos do conhecimento humano. Sua origem é a Modelagem Estatística da quase-implicação: quando a variável, ou a conjunção das variáveis, "X" é observada na população, geralmente a variável "Y" também é (ORÚS; ZAMORA; GREGORI, 2009).

Almouloud, Gras e Jean-Claude Régnier (2014) asseveram que a Análise Estatística Implicativa é um campo teórico que tem como objeto matemático o conceito de quase-implicação, distinguindo-se da implicação lógica da Matemática e da própria Lógica. A ASI busca conhecimentos, invariantes, e atributos das variáveis estudadas.

¹⁰ C.H.I.C.: Classificação Hierárquica Implicativa e Coesitiva, versão 7.0, *copywrite* 2014, licenciado para uso pelo grupo de pesquisa PEA-MAT, na PUC-SP.

O método estatístico implicativo tem sido amplamente utilizado em estudos qualitativos de regras de associação, apresentando grandes contribuições à pesquisa em Educação Matemática, permitindo evidenciar a dinâmica dos comportamentos dos alunos em situação de resolução de problemas. Ao contrário do que é legítimo na Matemática — em que todas as regras geralmente não sofrem exceção e o determinismo se faz sempre presente —, nas Ciências Humanas e Sociais, as exceções não somente estão quase sempre presentes como também são desejadas (ORÚS; ZAMORA; GREGORI, 2009). Segundo esses autores,

Ao contrário dos métodos de análise simétrica baseados, por exemplo, na distância na correlação, os conjuntos de regras obtidos podem levar à hipótese de causalidade. Esses conjuntos são estruturados de acordo com diferentes características comuns complementares (grafo implicativo, hierarquia orientada), criando conceitos de intensidade, do envolvimento, da coesão de classes, dos níveis significativos, etc. A determinação quantitativa dos sujeitos ou descritores responsáveis por essas estruturas são dadas por sua contribuição e sua tipicidade. (ibid., p. 66, tradução nossa)

Segundo Couturier, Bodin e Gras (2004), o *software* CHIC possibilita a extração de um conjunto de dados, cruzando sujeitos e variáveis (ou atributos), e regras de associação entre variáveis, fornecendo índices de qualidade das associações e representando a estruturação das variáveis obtida por meio de tais regras. Os indicadores com os quais trabalhamos em nossa pesquisa foram inseridos nesse *software* como variáveis binárias, assumindo os valores 0 ou 1, representando, respectivamente, a ausência e a presença da variável para cada indivíduo pesquisado.

No primeiro semestre de 2018, utilizamos a ASI e a Análise de Similaridades no levantamento dos conhecimentos prévios dos alunos sujeitos de nossa pesquisa. Até então, esses estudantes não haviam tido quase nenhum contato com a Estatística em ambiente escolar, como poderemos ver no capítulo 5 desta tese. Então, foi de nosso interesse avaliar o que eles já sabiam sobre a Estatística, antes de iniciar sua aprendizagem em uma abordagem por meio de projetos.

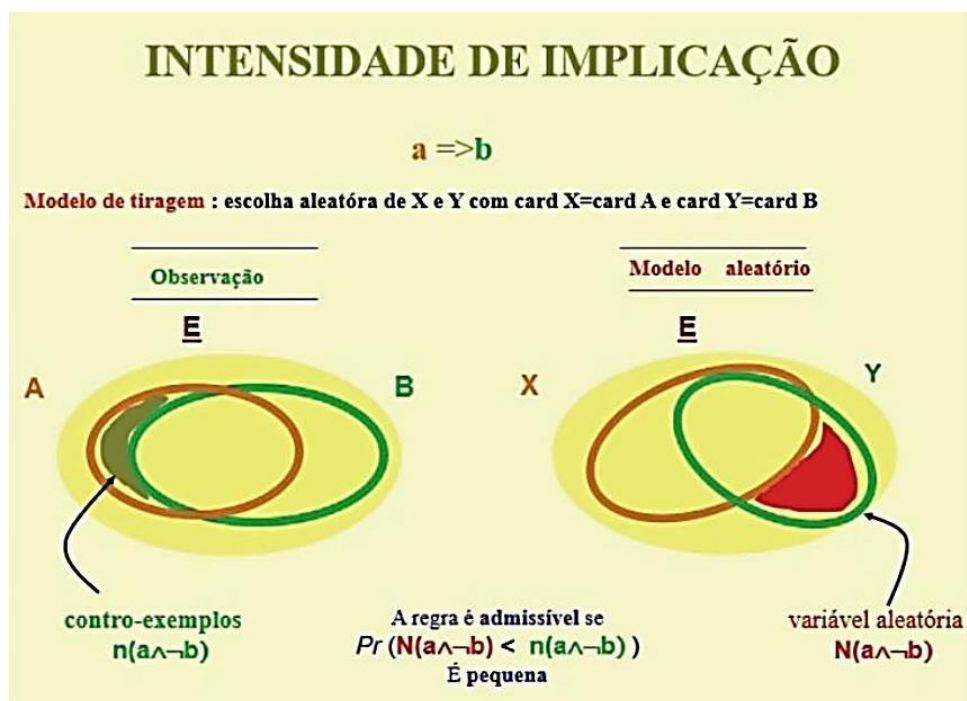
O *software* CHIC nos permitiu desenvolver a análise implicativa, a análise de similaridades bem como a análise coesitiva.

Por meio do grafo implicativo, tentamos compreender as relações de quase-implicação estabelecidas entre as variáveis observadas, utilizando um intervalo de confiança adequado para cada uma das setas que compõem os seus caminhos

implicativos. O valor do índice de confiança estabelecido por seus criadores é igual ou maior que 0,50. Segundo Almouloud (2015, p. 59-60):

Partindo dos subconjuntos (Figura 2) A e B, suportes respectivos de a e b, objetivamos estudar o subconjunto dos contraexemplos da implicação, isto é, as ocorrências da propriedade $(a \wedge \neg b)$ do suporte de $A \wedge \neg B$ ($\neg B$ sendo o complementar de B em E). O número K de contraexemplos é considerado como a realização de uma variável aleatória de um modelo de Poisson ou de um modelo binomial, um e outro, aproximados pela lei de Gauss quando é legitimado pelos parâmetros. [...] diremos que a implicação é admissível no índice de confiança a se a probabilidade que esta variável aleatória seja superior a k, e ela mesma superior a 1-a.

Figura 2 – Modelização da ASI

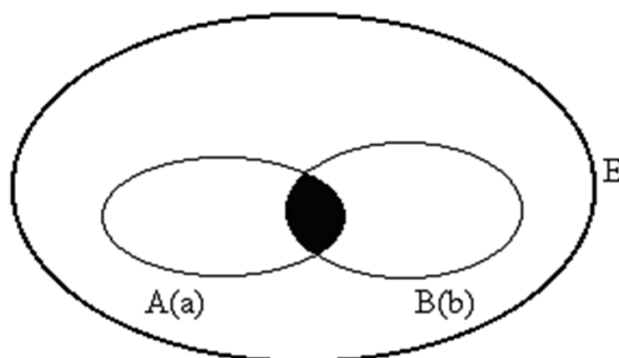


Fonte: ALMOULLOUD, 2015, p. 60.

Essa análise considera o critério de similaridade para variáveis binárias nos casos de presença/ausência, verdadeiro/falso e sim/não. Considere um conjunto E , composto pelos sujeitos da pesquisa, e duas variáveis, a e b . Sejam A e B subconjuntos de E , formados pelos sujeitos com as características a e b , respectivamente. Dizemos que as variáveis a e b são muito semelhantes, quando o número de sujeitos de C é suficientemente grande: em relação ao número de elementos dos conjuntos E , A e B ; e em relação ao número de elementos E que estão em A ou em B . Mede-se essa semelhança pela probabilidade de sua inverossimilhança. Sejam n_a e n_b os números de A e de B , respectivamente. O número de elementos de $A \cap B = A \wedge B$ é um indicador da similaridade de A e de B , considerando

o tamanho de **A** e de **B**. Tome uma parte **Y** ao acaso, extraída do subconjunto **P(E)** das partes de **E**, de cardinal n_b . Israël-César Lerman, Gras e Rostman (1981) consideram a variável aleatória $S_a = \text{card}(A \cap Y)$, sob a hipótese de ausência de relação, a priori, entre **A** e **Y**. Desta forma, a similaridade entre **a** e **b** é tão significativa quanto o valor observado $s = \text{card}(A \cap B)$ é maior do que os valores esperados. Assim, o índice de similaridade inicial será $\text{Prob}[S_a < s]$.

Figura 3 – Modelização da hierarquia de similaridade



Fonte: ALMOULOU, 2015, p. 57.

A similaridade é obtida pelo cruzamento do conjunto **V** das variáveis com um conjunto **E** de sujeitos (ou de objetos). Assim, é possível analisar — em termos de tipologia e de semelhança (e dessemelhança) decrescente — classes de variáveis, constituídas, significativamente, a certos níveis da árvore de similaridade, opondo-se a outros, nestes mesmos níveis. Podemos extrair daí três informações úteis para nossa interpretação: os níveis significativos, a tipicidade e a contribuição.

Um critério estatístico permite saber quais são os níveis significativos das árvores de similaridade e da hierarquia implicativa entre os demais níveis. Uma partição e as classes são formadas de acordo com os indícios de similaridade iniciais (ou da hierarquia implicativa). Cada nó significativo está associado à classe obtida nesse nível. A cada novo nível, forma-se uma classe de regras, cuja coesão é pior do que a das classes já formadas e melhor do que as que se formarão. Um nível é dito significativo (um nó é dito significativo) se corresponde a um máximo local desse índice estatístico, estabelecido em vermelho na árvore pelo *software* CHIC.

Quanto à tipicidade, dizemos que certos sujeitos são típicos do comportamento do conjunto da população se, no estudo da similaridade (ou da hierarquia implicativa), eles atribuem, ao conjunto das variáveis, valores compatíveis com as similaridades (ou a hierarquia implicativa) constituídas sobre essas variáveis

pela população. Sendo as variáveis suplementares previamente definidas pelo pesquisador, podemos obter a tipicidade dessas variáveis a partir das tipicidades dos indivíduos que as satisfazem. É possível conhecer a contribuição de cada um dos sujeitos e, então, das variáveis suplementares, conforme suas similaridades (ou hierarquia implicativa), por exemplo: no decorrer de uma pesquisa educacional, os professores de Matemática são típicos do comportamento de uma população de docentes que tem maior afinidade com as novas tecnologias digitais.

Por fim, a análise das classes representadas na árvore coesitiva nos permite visualizar aquelas que têm o maior número de nós significativos — assinalados em vermelho nos termos do *software* — bem como o conjunto de variáveis que ela envolve, explicitando as meta-regras identificadas, tal qual “se **A** ocorre, então provavelmente **B** ocorre”.

A ASI tem potencial para auxiliar na compreensão das relações que envolvem dados multivariados em contextos educacionais. Por exemplo, na perspectiva da ASI, uma relação $A \rightarrow B$, identificada na análise coesitiva, pode indicar que, se o aluno conhece Estatística, então ele também sabe calcular a média aritmética entre variáveis determinadas. Já na perspectiva da Teoria das Concepções, poderíamos, ainda nesse exemplo, considerar, como operador, “Se é variável, então calcula-se a média aritmética” e, como estrutura de controle, “Se é Estatística, então existe média aritmética”. Essa ideia ficará mais clara na próxima subseção.

Voltaremos a tratar da ASI e da Análise de Similaridades no capítulo 5 desta tese, quando apresentarmos os resultados de nossa pesquisa.

3.3 Teoria das Concepções

Nesta subseção, apresentamos nosso último referencial teórico-metodológico, a Teoria das Concepções.

Fortemente influenciada pela Teoria das Situações Didáticas – TSD e pela Teoria dos Campos Conceituais, segundo Oliveira e Coutinho (2011), a Teoria das Concepções visa estudar as relações entre os conceitos, os conhecimentos e as concepções que, porventura, os alunos possam vir a apresentar sobre determinado conteúdo matemático. Para Balacheff (1995), resgatar a noção de concepção é mobilizar ações no sujeito, mediante a realização de uma atividade. Este, em situações específicas, pode agir racional e coerentemente para resolver a questão.

Segundo Balacheff e Nathalie Gaudin (2002), os únicos indicadores sobre o sucesso do ensino são o comportamento dos alunos e as suas produções, consequências do conhecimento por eles construído e de seu relacionamento com o conteúdo ensinado. Oliveira e Coutinho (2011) destacam que a Teoria das Concepções se diferencia da Teoria dos Campos Conceituais, sobretudo, pelo duplo papel dos invariantes operatórios. Para Balacheff (1995), os operadores são os instrumentos necessários à resolução do problema, e as estruturas de controle fazem a validação ou não da concepção.

Ainda de acordo com Novaes (2011), Balacheff desenvolveu o modelo cK Φ considerando uma concepção como um objeto de estudo e ampliando a terna (S, I, L), estabelecida por Gerard Vergnaud, sendo: S – Situações que dão sentido ao conceito; I – Invariantes operatórios (como teoremas em ação); e L – Significantes, que permitem representar o conceito.

Para Takeshi Miyakawa (2005), o modelo cK Φ repousa sobre a problemática da TSD, desenvolvida por Guy Brousseau, na qual a questão da relação entre os comportamentos de um sujeito e os conhecimentos é fundamental. Essa teoria modela a situação de aprendizagem e de ensino, na qual têm lugar as diversas interações entre os alunos, o *milieu* e o professor, posição essencial para a modelização dos conhecimentos dos estudantes.

Lima (2008) lembra que no modelo cK Φ , uma concepção é definida como uma estrutura mental atribuída a um sujeito por um observador de seu comportamento enquanto a aprendizagem é compreendida como a passagem de uma concepção para a outra. A caracterização de uma concepção sobre um objeto matemático, segundo Balacheff (2001), não deve ser separada da caracterização da situação-problema que permite evidenciá-la.

De acordo com Melo e Lima (2011), no modelo cK Φ , uma concepção é um estado de equilíbrio de um sistema, sujeito-meio, considerando as limitações e imposições, ou seja, é algo que influencia ou interfere em seu funcionamento. Uma vez que a concepção é referente ao sujeito, pode ser adequada ou não, do ponto de vista do conhecimento de referência.

Outro aspecto importante deste modelo é que a concepção, muitas vezes, é local, no sentido de que ela funciona para resolver um determinado problema, e não outro. Isso aponta para um domínio de validade, que se refere a um problema ou conjunto de problemas específicos para os quais uma concepção funciona ou parece

funcionar bem, levando o sujeito a acreditar que a mesma possa ser generalizadamente aplicada a problemas de outra natureza.

Além disso, para compreender uma concepção, é necessário conhecer seus elementos constituintes, a qual envolve uma quádrupla (P, R, L e Σ):

- P – conjunto de problemas, sobre o qual a concepção Φ é operatória;
- R – conjunto de operadores (ferramentas cognitivas para ação);
- L – sistema de representação, que permite exprimir os elementos de P e R;
- Σ – estrutura de controle, que assegura a não contradição da concepção Φ .

Oliveira e Coutinho (2011) esclarecem, resumidamente, os elementos dessa quádrupla. Para tanto, de acordo com os autores:

- um problema (P) pode ser considerado resultado de uma perturbação do equilíbrio do sistema sujeito-milieu. Assim, um conjunto de problemas, formulado de forma empírica, com base na caracterização das situações, permite identificar as concepções dos alunos, dando significado à concepção;
- um operador (R) é uma ferramenta cognitiva mobilizada por um sujeito para resolver problemas, sendo explicitada na ação desse sujeito ao tentar resolvê-los. Essa ação pode ser considerada concreta, quando permite uma ação material, ou abstrata, quando permite transformações linguísticas, simbólicas ou representações gráficas;
- um sistema de representações (L) corresponde a um conjunto de significantes compreendidos como a representação de um objeto ou pensamento, tais como a linguagem algébrica, o desenho geométrico e a representação gráfica. Permite expressar os elementos de P e de R;
- uma estrutura de controle (Σ) permite exprimir os meios do aluno de fazer escolhas, tomar decisões e validar uma ação bem como tem a função de julgar a validade e a adequação da ação realizada pelo sujeito que resolve um problema. Assume, ainda, a não contradição da concepção Φ . Além disso, é o novo elemento apresentado por Balacheff (1995) e no

modelo em pauta e uma ampliação do modelo proposto por Vergnaud na Teoria dos Campos Conceituais.

Para o conjunto R, adotamos por operadores os conceitos e teoremas em ato/ação segundo Vergnaud (2005, p. 132, tradução nossa):

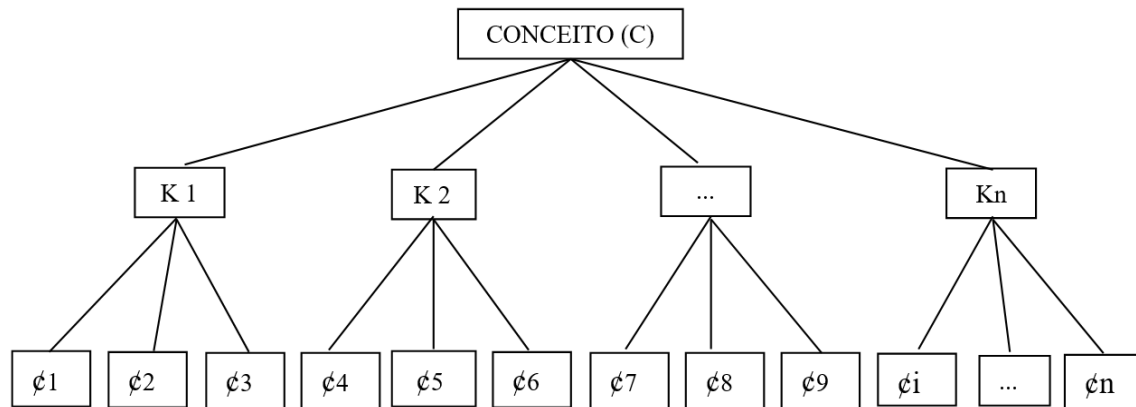
A distinção entre conceito em ação e teorema em ação é essencial [...]: não há nenhum teorema sem conceito ou conceito sem teorema. [...] Teoremas são propostas (declarações) e são suscetíveis à verdade e falsidade, e não conceitos! Os conceitos só podem ser relevantes ou irrelevantes para coletar informações. Estas não são propostas, mas funções proposicionais (predicados) ou argumentos (objetos), que não são nem verdadeiros, nem falsos. Não são enunciados. Mas precisamos de enunciados que são verdadeiros para determinar metas e regras. [...] Mas os conceitos e teoremas subjacentes à ação não são explícitos, ou mesmo explicitáveis, pelo sujeito no momento em que são implementados.

Para Balacheff e Gaudin (2002), o conhecimento não pode ser totalmente reduzido a comportamentos, mas também não pode ser ensinado na ausência desses. Toda ação mobiliza uma considerável quantidade de conhecimentos e, a fim de desenvolver novos conhecimentos bem como de aprofundar os anteriores, se faz necessária a mobilização de concepções diretamente relacionadas aos problemas enfrentados pelos alunos. Balacheff (2001) afirma que uma concepção não pode nem deve ser separada do contexto do qual emerge o problema, que a evidencia e que lhe dá sentido. Almouloud (2007) nos lembra de que essas concepções permitem interpretar, prever e construir modelos e, sobretudo, descrever uma parte da estrutura cognitiva, em nosso caso, do aluno.

Em nossa pesquisa, adotaremos as definições de concepção, conhecimento e conceito do modelo cK Φ proposto por Balacheff (1995) na Teoria das Concepções, o qual aponta que uma concepção é uma estrutura mental, característica de um dado sujeito (neste caso, o aluno) e constituída por um observador de seu comportamento (aqui, o pesquisador), enquanto a aprendizagem consiste na passagem de uma concepção para uma nova, mais complexa e abrangente.

Um conceito é constituído por um conjunto de conhecimentos, e um conhecimento, por sua vez, é constituído por um conjunto de concepções. Assim, nesta tese, nosso foco recai sobre o nível mais básico dessa estrutura, ou seja, o nível das concepções, como vemos representado a seguir:

Figura 4 – Esquema das possíveis relações entre concepções, conhecimentos e conceitos, propostas por Balacheff



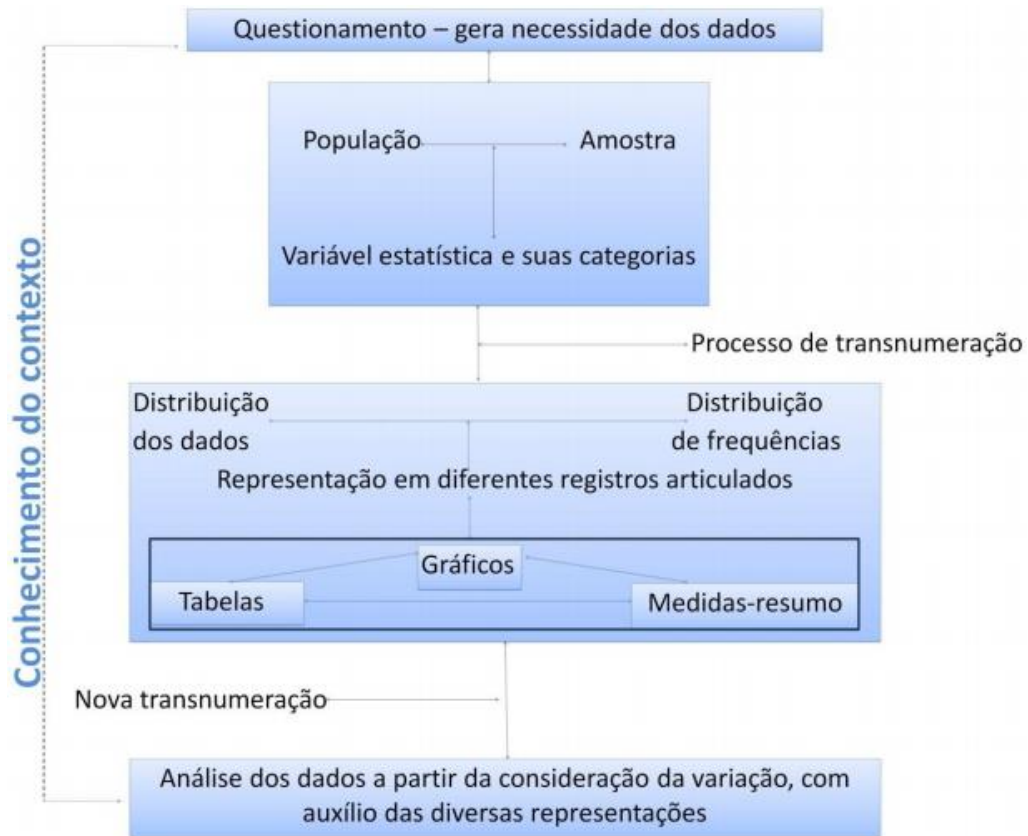
Fonte: OLIVEIRA, 2010, p. 43.

Na Figura 4, $c_1, c_2, c_3, \dots$ e c_n constituem o conjunto de concepções que forma o de conhecimentos, $K_1, K_2, \dots, K_n$, que, por sua vez, forma o conceito C .

Como já dissemos, em nossa pesquisa, os alunos poderão mobilizar determinadas concepções ao tentarem resolver problemas de Estatística similares àqueles com os quais se defrontaram no desenvolvimento de seus projetos. É esperado que mobilizem concepções ao realizarem suas próprias pesquisas, as quais serão úteis na resolução dos problemas discutidos no capítulo 5 desta tese. As pesquisas desses estudantes envolvem os mais variados contextos, em situações que não estão acostumados a enfrentar nas aulas de Matemática, por questões curriculares já apresentadas por nós.

Ao analisar o pensamento do aluno e ao se deparar com situações como essas, Novaes (2011) propôs o seguinte modelo:

Figura 5 – Elementos da Estatística associados ao Pensamento Estatístico



Fonte: NOVAES, 2011, p. 81.

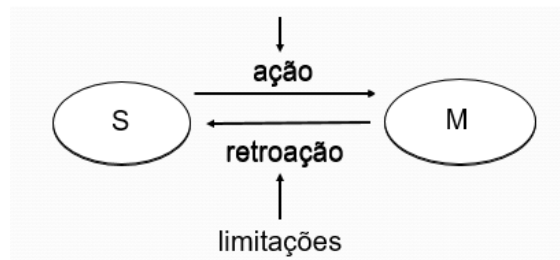
Não vamos nos ater, agora, a analisar cada elemento desse esquema, pois isso será feito no capítulo 4 deste trabalho, no estudo do objeto, bem como no capítulo 5, ao tratar da resolução de problemas de Estatística por parte dos alunos. Mas, enfim, o que são problemas para a Teoria das Concepções?

Cada problema desperta no aluno, comportamentos que são indicadores de conhecimento, evidenciando indícios de estratégias por ele elaboradas na busca de soluções. E esse conhecimento não pode ser atribuído só ao estudante, como uma propriedade que lhe será intrínseca, tampouco somente ao meio, uma vez que é uma propriedade do sujeito em interação com um sistema antagônico, com o *milieu*. Tal interação só será significativa se conseguir satisfazer as limitações que condicionam a viabilidade de relação sujeito-meio. A solução de problema — resultado da perturbação mais ou menos severa do equilíbrio dessa relação —, ou seja, o reequilíbrio nessa relação, é a fonte e o critério do saber.

Assim, o conhecimento é caracterizado por um estado de equilíbrio dinâmico na interação ação-retroação, entre um sujeito S e um determinado meio específico M,

de acordo com as limitações impostas pela situação, as quais determinam ou não a sua viabilidade, como ilustrado na figura a seguir.

Figura 6 – Retroação no sistema sujeito-meio



Fonte: BALACHEFF; MARGOLINAS, 2005, p. 81, tradução nossa.

O sistema didático é composto pelo agente didático, humano ou artificial, pelo aluno e pelo meio. O conhecimento envolvido em seu funcionamento é uma propriedade inerente ao sistema aluno-meio sob as restrições sustentadas pelo agente didático, o qual atua como representante da instituição produtora do sistema didático. O saber é um atributo do conhecimento, cujo valor se reconstrói do julgamento do agente didático, como vemos na figura a seguir:

Figura 7 – Sistema didático na Teoria das Concepções



Fonte: BALACHEFF, 1995, p. 230, tradução nossa.

Em suas tentativas de resolver um dado problema, o aluno pode se deparar com diversos obstáculos, os quais não implicam em uma total ausência de conhecimento, e sim em um conhecimento mobilizado fora de seu domínio de validade. Esse obstáculo pode resistir às contradições e continuar a se manifestar mesmo depois que o indivíduo perceba sua inexatidão e não está ligado apenas a

fatores estritamente cognitivos, mas também a fatores emocionais. Além disso, tem a ver com a sensação de confiança que esse “conhecimento” deslocado traz, pois ele, possivelmente, foi útil em outras ocasiões para resolver muitos problemas. Desse modo, cabe ao pesquisador identificar esses “erros” resistentes e mostrar que eles se organizam em torno de concepções.

O erro, nesse quadro teórico, é compreendido como sinal de uma perturbação, resultado da diferença reconhecida pelo sujeito entre o resultado esperado de uma ação e a retroalimentação do meio. O aluno consegue reconhecer essa diferença como inaceitável, reagindo a ela e buscando o reequilíbrio, o que o conduz a um estágio superior de conhecimento.

Para Balacheff (1995), uma concepção errônea requer a noção de uma conclusão aparentemente bem-sucedida de uma diversidade de problemas em que o ciclo de formulação do problema (expectativa), resolução do problema (ação) e reconstrução do problema (revisão), é efetivamente realizado. Tal concepção errônea tem seu próprio campo de validade, pois, como já observamos, o erro não é somente o efeito da ignorância, da incerteza e do azar que encontramos nas teorias empiristas ou behavioristas da aprendizagem, mas sim a consequência de um conhecimento anterior.

O que distingue, essencialmente, o paradigma das concepções errôneas do paradigma do erro é a significação epistemológica de cada um deles. Para Balacheff (1995), o estatuto do conhecimento é diferente, já que as concepções errôneas subentendem a existência de um conhecimento de referência global e legítimo, enquanto o paradigma do erro, particularmente na problemática dos obstáculos epistemológicos, só requer aquilo que relaciona, explicitamente, dois conhecimentos em uma perspectiva de desenvolvimento. Assim, o conhecimento é efêmero e provisório bem como pode ser o objeto de mudanças e ampliação de seu campo de validade após perturbações.

O erro persistente e os conhecimentos contraditórios são objetos de análise de muitas investigações à luz da Teoria das Concepções de pesquisadores como Sangaré (2000) e Vadcard (2000). Sobre tais contradições — discutidas com mais profundidade no capítulo 5 desta tese —, esta autora assevera que

Fischbein (1982) mostrou que, no mesmo sujeito, a coexistência de conhecimentos contraditórios pode estar presente em aspectos tão fundamentais quanto os meios de validação. Este fenômeno parece, à primeira vista, estar em contradição com a ideia de uma construção do

conhecimento por adaptação que sugere não-contradição: "o conhecimento racional é operacional ou não é" (VERGNAUD, 1991, p.136). A contradição removeria esse caráter operativo. O modelo das concepções, por associar cada manifestação do conhecimento a uma esfera prática, possibilita explicar a contradição que existe aos olhos do observador: a não coerência, atribuída por uma disjunção das esferas de práticas. Portanto, esse modelo permite preservar a natureza racional do conhecimento operacional, ao mesmo tempo em que fornece uma explicação para a coexistência de estados contraditórios no mesmo sujeito. (VADCARD, 2000, p. 16, tradução nossa)

Essas contradições observadas por Sangaré (2000) e Vadcarg (2000) serão discutidas, no próximo capítulo. Em resumo, neste capítulo, apresentamos nosso quadro teórico-metodológico. A escolha pela Análise Exploratória de Dados como filosofia norteadora é justificada por sua influência sobre a abordagem por meio de projetos, a qual adotamos na orientação das pesquisas dos alunos, (BATANERO, 2001; BATANERO; DÍAZ, 2004, 2011; BATANERO; GODINO, 2001). A maior contribuição da AED é sua perspectiva mais criativa e exploratória do que técnico-procedimental (TUKEY, 1962, 1970; BATANERO; ESTEPA; GODINO, 1991), tão necessária à análise na Educação Estatística.

Da mesma forma, a escolha pela Análise Estatística Implicativa é explicada pela nossa necessidade de analisar dados multivariados, coletados por meio de um questionário respondido pelos alunos sujeitos desta pesquisa. A maior contribuição da ASI, através do *software* CHIC, foi evidenciar a dinâmica dos comportamentos dos estudantes em situação de resolução de problemas. Esse *software* permitiu a extração de um conjunto de dados relevantes, cruzando sujeitos e atributos, regras de associação entre variáveis, fornecendo índices de qualidade das associações, representando a estruturação das variáveis (ALMOULOU; GRAS; RÉGNIER, 2014; COUTURIER; BODIN; GRAS, 2004; GRAS; ALMOULOU, 2003; ORÚS; ZAMORA; GREGORI, 2009).

Por fim, a Teoria das Concepções esteve no cerne de nossas investigações porque aborda exatamente o que desejávamos para tentar responder à nossa questão de pesquisa de identificar as concepções dos alunos — levantada ainda em nossa dissertação —, a qual nos motivou a, por meio do modelo cK ϕ , identificar os elementos essenciais à compreensão da mobilização das concepções por parte dos estudantes, em especial, os operadores e as estruturas de controle (BALACHEFF 1995, 2001, 2002; BALACHEFF; GAUDIN, 2002; BALACHEFF; MARGOLINAS, 2005).

A Teoria das Concepções nos permitiu estudar as relações entre os conceitos, os conhecimentos e as concepções apresentadas sobre determinados conteúdos estatísticos. Em nosso caso, o foco do estudo foram as próprias concepções.

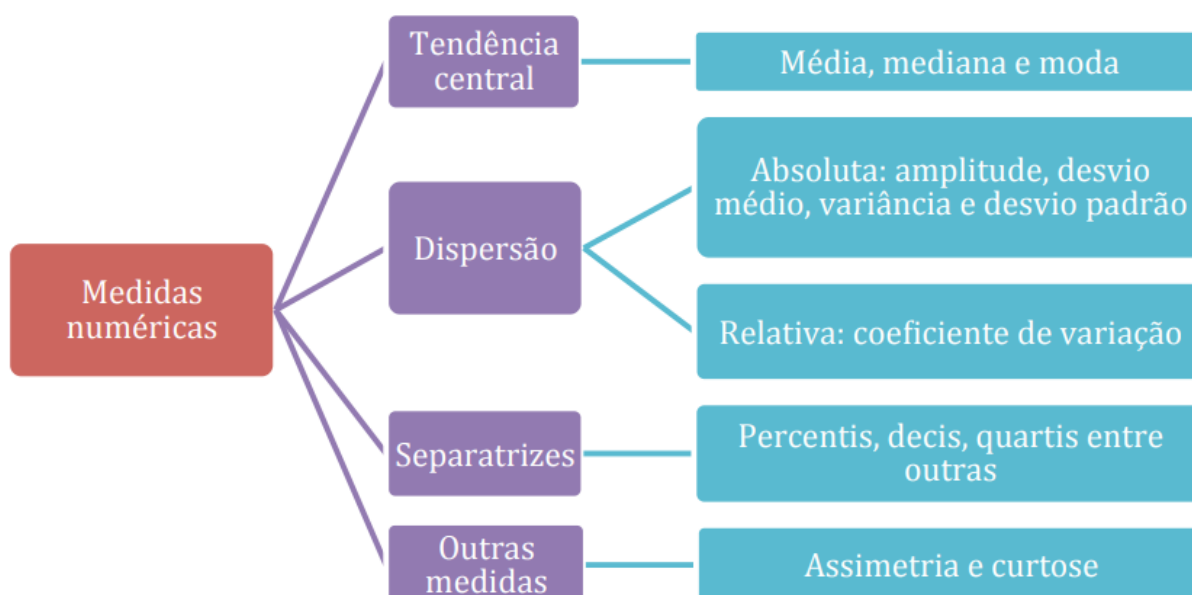
Na próxima seção, apresentaremos o estudo do objeto matemático, analisando os principais elementos envolvidos no ensino e na aprendizagem da Estatística por meio da resolução de problemas após o desenvolvimento de projetos de pesquisa.

4 ESTUDO DO OBJETO

Nesta seção, apresentaremos alguns conceitos básicos da Estatística, contemplados no Currículo do Estado de São Paulo para o ensino médio, vigente em 2018 (SÃO PAULO, 2012), ano da nossa coleta de dados.

Iniciaremos abordando as medidas estatísticas de tendência central e de dispersão, conforme destacado por Vladimir de Andrade (2013) em sua tese.

Figura 8 – Medidas estatística numéricas



Fonte: ANDRADE, 2013, p. 82.

Com base nessas medidas, fizemos um recorte e falaremos sobre os dois primeiros blocos, por serem os que estão presentes nos livros didáticos adotados pelos alunos sujeitos da pesquisa e, sobretudo, no material didático oficial, o Caderno do Aluno do Estado de São Paulo (SÃO PAULO, 2014): medidas de tendência central — moda, mediana e média aritmética — e de dispersão absoluta — variância, desvio padrão, desvio médio e amplitude total.

A seguir, tratamos dos conceitos de variabilidade e sistemas de representação, sendo eles gráficos e tabelas estatísticas, investigados por meio de nossos dois instrumentos de coleta: o questionário individual e a atividade em grupos com resolução de problemas estatísticos.

4.1 Determinação de medidas de tendência central e de dispersão

O Currículo do Estado de São Paulo prevê que o professor de Matemática aborde em suas aulas do 4º bimestre no 3º ano:

- gráficos estatísticos: cálculo e interpretação de índices estatísticos;
- medidas de tendência central: média, mediana e moda;
- medidas de dispersão: desvio médio e desvio padrão.

(SÃO PAULO, 2012, p. 70)

Novaes e Coutinho (2009) salientam que a função da Estatística na organização dos dados é oferecer uma visão panorâmica do problema viabilizando o estudo, por meio do qual poderão ser observadas as relações entre variáveis, padrões e regularidades que permitam estimativas, previsões.

Não é o objetivo da Educação Estatística priorizar cálculos da Estatística Descritiva, mas sim estudar problemas do ensino e da aprendizagem da Estatística. Entretanto, os alunos aprendem a obter tais medidas e, sobretudo, empregá-las na Estatística Inferencial. Campos, Wodewotzki e Jacobini (2013) destacam a importância pedagógica da investigação e reflexão em detrimento de uma perspectiva que enfatize aspectos técnicos e operatórios.

Cláudia da Silva (2007), em sua tese de Doutorado acerca do pensamento estatístico e raciocínio sobre variação com professores de Matemática, afirma que os conceitos de variação e variabilidade não são completamente compreendidos pelos docentes, o que, obviamente, compromete a aprendizagem de seus alunos. Falta entendimento conceitual para analisar os dados aplicando as técnicas abordadas nas aulas. A autora admite que:

[...] é preciso ensinar mais do que simplesmente o cálculo de medidas e elaboração de gráficos. É preciso discutir sobre o significado e aplicabilidade das medidas e representações e, principalmente, relacionar esses conceitos. (Ibid., p. 21)

Nessa perspectiva, pretendemos analisar os conceitos e concepções dos alunos a respeito da Estatística, concebidos em um contexto de resolução de problemas após o desenvolvimento de projetos, conforme a Teoria das Concepções.

4.1.1 Medidas de tendência central

De acordo com Novaes e Coutinho (2009, p. 51), são as medidas-resumo que permitem “a construção de um 'retrato' dos dados tratados, complementando as representações tabulares e gráficas” (grifo das autoras). Medidas essas que, para Batanero e Godino (2001), são representadas pelo conjunto de valores ao redor dos

quais se agrupam os dados, dentro deles, a média, a mediana e a moda. Silvia Galicia (2009, p. 13, tradução nossa) define, de modo simples, as três principais medidas de tendência central:

A medida da tendência central mais intuitiva é o valor obtido adicionando os valores observados e dividindo esta soma pelo número de dados que estão no grupo. A média resume as características de uma variável, considerando todos os casos, mas apenas pode ser usada com variáveis quantitativas. A mediana e moda são outras medidas de tendência central também aplicáveis a outras variáveis, a primeira aos dados ordinais e a segunda aos dados qualitativos. A mediana é intuitivamente definida como o valor que está no centro da distribuição, quando os dados são ordenados de modo crescente, e a moda, como o valor de maior frequência.

Batanero e Díaz (2011) observam que, até mesmo no ensino superior, os alunos encontram dificuldades em saber quando e como utilizar cada uma das medidas-resumo e, em alguns casos, simplesmente não sabem calculá-las, em outros, o algoritmo é aplicado mecanicamente, sem compreensão de seu significado.

Dentre os erros mais comuns no cálculo das medidas-resumo de tendência central, as autoras destacam: supor que a moda sempre existe e é única; não ordenar os dados ao calcular a mediana; trocar a moda pela mediana; e não atribuir pesos distintos ao calcular a média ponderada. Ainda, Batanero e Díaz (2011) identificaram as seguintes variáveis didáticas que afetam a dificuldade das tarefas:

- formato (tabela, série de números e gráfico);
- valores das variáveis muito maiores do que os das frequências;
- uma das frequências muito mais elevada do que as demais.

Com base nessas considerações sobre as medidas-resumo, nas próximas subseções, vamos analisá-las, brevemente, uma a uma.

4.1.1.1 Moda

A moda é uma medida-resumo útil quando a variável é qualitativa, tornando impossível a determinação da média, sobretudo quando é qualitativa nominal, o que também impossibilita a determinação da mediana, como observam Batanero e Godino (2001). Tal medida-resumo é definida por Pedro Morettin e Wilton Bussab (2013, p. 35) como “a realização mais frequente do conjunto de valores observados”.

Régnier (2007) traz duas definições de moda: quando a variável é quantitativa discreta ou qualitativa, a moda é o efetivo ou frequência máximos; e, quando a variável é quantitativa contínua, a moda corresponde à densidade da frequência máxima.

A isso, Batanero e Godino (2001) acrescentam que, em uma distribuição de frequência, pode haver mais de uma moda. Assim, se houver apenas uma moda, a chamamos de unimodal; duas, bimodal; e, caso haja mais de duas, multimodal (alguns autores preferem o termo 'plurimodal'). Além disso, a moda pode, até mesmo, não existir quando todas as realizações não se repetem (amostra amodal), como apontam Novaes e Coutinho (2009).

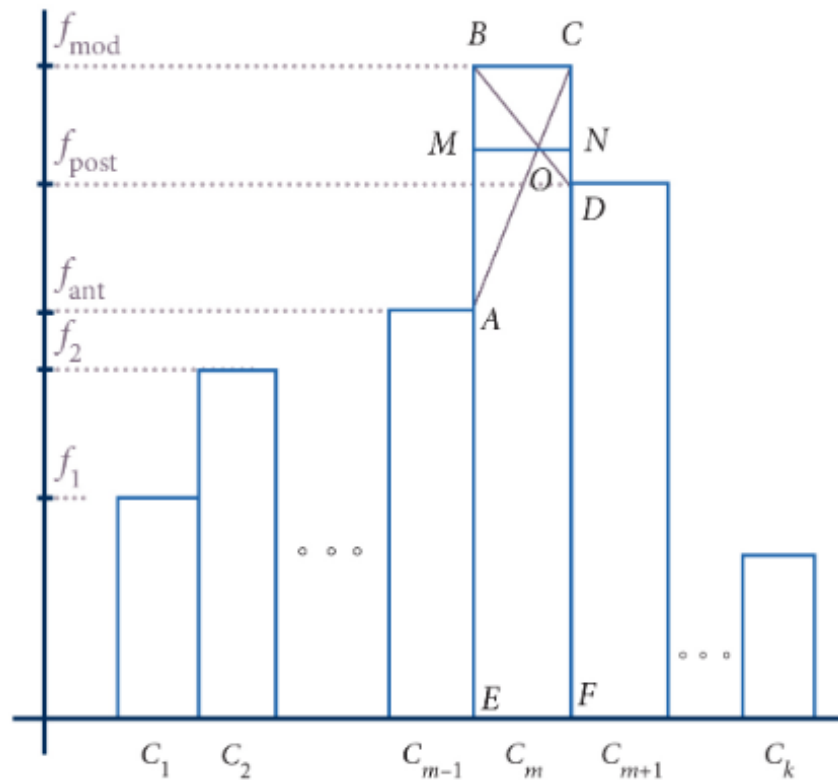
Quando a variável está agrupada em intervalos de classe, segundo Batanero e Godino (2001), a moda recai sobre a classe mais frequente, podendo ser determinada pela expressão $M_o = E_i [d_i / (d_i + d_{i+1})] a_i$, na qual M_o representa a moda; E_i , o limite inferior real da classe modal; d_i , a diferença entre a frequência absoluta da classe modal e a classe anterior; d_{i+1} , a diferença entre a frequência absoluta da classe modal e o seguinte; e a_i , a amplitude do intervalo de classe modal. De maneira aproximada, podemos tomar como forma o centro do intervalo modal (intervalo de frequência mais elevada).

Um método para a determinação da moda, baseado na Geometria, mais especificamente na semelhança de triângulos, é a moda de Czuber — não apresentada no Caderno do Aluno (SÃO PAULO, 2014), infelizmente —, a qual se trata de um processo gráfico, como descrito por Márcio Monteiro (2014).

Temos, no histograma a seguir (Figura 9): a variável x ; $C_1, C_2, \dots, C_k$, que são os intervalos de valores assumidos pela variável x , denominados por nós como classes; $f_1, f_2, \dots, f_k$, os quais são, respectivamente, as frequências dessas classes; e C_m , a classe modal, aquela com a maior frequência de ocorrências. Consideramos os limites l_i e l_s os extremos inferior e superior do intervalo C_m . Pelo método de Czuber, determinamos, nessa situação gráfica, um valor que chamaremos de moda de Czuber.

Traçamos os segmentos AC e BD , que se intersectam em O . A moda de Czuber é a soma $l_i + MO$. O segmento MO construído faz a moda de Czuber se aproximar da classe com maior frequência adjacente à classe modal.

Figura 9 – Construção gráfica da moda de Czuber



Fonte: MONTEIRO, 2014, p. 2.

Na figura anterior, temos que:

- f_{mod} é a frequência da classe modal;
- f_{ant} é a frequência da classe imediatamente anterior à classe modal;
- f_{post} é a frequência da classe imediatamente posterior à classe modal;
- h_{mod} é a amplitude da classe modal (medida do segmento EF).

Os triângulos AOB e COD são semelhantes, logo, $\frac{MO}{AB} = \frac{NO}{CD}$.

Mas $AB = f_{mod} - f_{ant}$ e $CD = f_{mod} - f_{pos}$, logo, $\frac{MO}{f_{mod} - f_{ant}} = \frac{NO}{f_{mod} - f_{pos}} =$

$$\frac{MO+NO}{2f_{mod} - f_{ant} - f_{pos}} = \frac{h_{mod}}{2f_{mod} - f_{ant} - f_{pos}}.$$

Assim, a moda de Czuber, M_o , será:

$$M_o = l_i + MO = l_i + h_{mod} \left(\frac{f_{mod} - f_{ant}}{2f_{mod} - f_{ant} - f_{pos}} \right).$$

Dentre as limitações da moda, Batanero e Godino (2001, p. 40-41) destacam:

- a) Se frequências se concentrem fortemente em alguns valores da variável, a moda não é uma medida eficaz de tendência central.
- b) Uma mesma distribuição, com os valores agrupados em classes distintas, pode resultar em uma moda distinta, no cálculo aproximado.

Quando a moda não é representativa, podemos recorrer a outras medidas de tendência central, como a mediana, tratada a seguir.

4.1.1.2 Mediana

A Mediana permite, segundo Novaes e Coutinho (2009, p. 52), uma “rápida visualização de como os dados estão distribuídos, onde existe a maior concentração ou dispersão”.

Batanero e Godino (2001) destacam que esses valores numéricos indicam sua posição no conjunto de dados ordenados, uma vez que uma determinada fração dos dados tenha um valor variável menor ou igual à medida estatística. O mais importante é a mediana, que também é uma medida de posição central. Desse modo, os autores consideram que:

Se assumirmos todos os valores de uma variável ordenada, do menor para o maior, um número é chamado mediana se existirem tantos valores da variável maiores ou iguais quanto menores ou iguais a ela. (Ibid., p. 42, tradução nossa)

Tukey (1970) atribuía grande importância ao valor da mediana, considerando-a uma função da distribuição empírica mais robusta em relação às demais medidas-resumo.

O cálculo da mediana é considerado por Belén Cobo (2003) um processo complexo, porque seu algoritmo de cálculo é diferente quando há um número par ou ímpar de dados ou quando os dados estão apresentados em tabelas de valores agrupados com intervalos de classe ou não agrupados. A autora também ressalta que, desde a definição de mediana como "valor da variável estatística que divide em duas partes iguais aos indivíduos da suposta população, organizados em ordem crescente" até seu cálculo, baseado na representação gráfica de frequência acumulada, existe muitos pontos a serem esclarecidos.

Para Estepa (1993), calcular a mediana com base nas representações gráficas das frequências acumuladas oferece alguns desafios:

- para as variáveis quantitativas discretas, tal representação corresponde a um efeito ‘de escada’, com descontinuidades, saltos, que os alunos não estudam com frequência;
- o valor da mediana, geralmente, é indeterminado, o que gera confusão entre os alunos (há mais de um valor que poderia corresponder à definição de mediana);
- os alunos não têm domínio suficiente sobre o tratamento das desigualdades que emerge da definição de mediana e de seu cálculo;
- no caso de variáveis contínuas, dependendo da amplitude do intervalo, o valor da mediana varia, o que também acontece com as outras medidas estatísticas. Contudo, o aluno não está acostumado com isso nas aulas de Matemática, pois, na escola, a maioria dos problemas matemáticos apresentados tem solução única.

Estepa (ibid.) considera este último uma falha do contrato didático ao ensinar a Estatística nas aulas da disciplina.

A fim de calcular a mediana, Eric Langford (2006) apresenta 15 métodos diferentes, alguns deles com o auxílio de *softwares* — tais como *Excel*, *Minitab* e *SAS* —, chegando a valores diferentes, embora todos válidos, o que proporciona elementos para uma boa discussão sobre as vantagens e desvantagens de cada escolha. No entanto, no Caderno do Aluno (SÃO PAULO, 2014), essa discussão não é sugerida, tampouco essa diversidade de métodos é apresentada. Na verdade, nesse material, é apresentado um único método para calcular a mediana.

Uma maneira possível de expor a mediana ao aluno seria por meio de um gráfico, como o *boxplot*, desenvolvido por Tukey, que tem o objetivo de possibilitar a verificação de uma distribuição de dados. Assim, ao analisar um *boxplot*, tiramos as seguintes conclusões: centro dos dados (a média ou mediana); a amplitude dos dados (máximo – mínimo); a simetria ou assimetria do conjunto de dados; e a presença de *outliers* (dados que se diferenciam drasticamente de todos os outros, pontos “fora da curva”). Além de informações importantes sobre localização e dispersão, esse gráfico fornece dados a respeito da cauda da distribuição.

Vejamos um exemplo apresentado por Batanero e Godino, em que representam pesos de *varones* (sexo masculino) e *hembras* (sexo feminino) em um gráfico *boxplot*.

Figura 10 – Pesos de varones y hembras

Peso en Kg.											
Varones						Hembras					
55	64	70	74	75	70	60	45	46	50	47	55
64	93	60	62	70	80	49	52	50	46	50	52
61	60	62	68	65	65	52	48	52	63	53	54
66	68	70	72	72	71	54	54	53	55	57	44
						56	56	56	53	60	65
						67	61	68	55	64	60

Fonte: BATANETO; GODINO, 2001, p. 56.

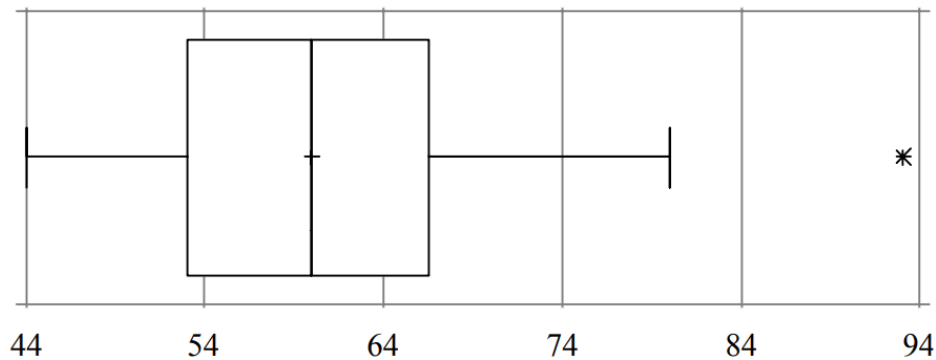
As extremidades do eixo são os valores mínimo e máximo, 44 e 93 kg, respectivamente. Como vemos, essa "caixa" indica, graficamente, o intervalo de 50% de variação dos valores centrais em uma distribuição que, para o peso dos alunos, varia de 53 a 66,5 kg.

A caixa é dividida em 2 partes, traçando uma linha na altura da mediana (60 kg em nosso caso). Portanto, $\frac{1}{4}$ dos alunos tem um peso entre 44 e 53 kg, sendo incluso os outros quartos nas seguintes faixas de peso: 53 a 60 kg, 60 a 66,5 kg e 66,5 a 93 kg. Como aqui o peso mínimo é de 44 kg e o intervalo interquartil é de $66,5 - 53 = 13,5$, subtraindo o primeiro quartil — $Q_1 = 53$ — uma vez e a metade do intervalo interquartil (II), obtemos: $Q_1 - 1,5 (13,5) = 53 - 20,25 = 32,75$. O máximo entre 44 e 32,75 é 44, então o segmento inferior a ser desenhado no gráfico da caixa deve ir até 44.

No peso dos homens, o máximo é de 93 kg. Assim, temos: $Q_3 + 1,5 (13,5) = 66,5 + 20,25 = 86,75$. Desse modo, a extremidade superior do segmento agora deve se estender apenas para 86,75 (vale que não é exatamente esse valor que encontramos no gráfico *box-plot* abaixo, na figura 11, mas é assim que se encontra no original e nós mantivemos os valores presentes no original, pois nesse caso, o que mais nos interessa é a discussão dos autores em torno das medidas de posição). Se algum dos dados estiver fora do intervalo coberto pela caixa e por esses segmentos, como no exemplo com o aluno pesando 93 kg, é indicado por um asterisco no gráfico, ou seja, ele é um *outlier*, o qual pode indicar um erro ou um caso excepcional.

O gráfico que, dentre outras coisas, evidencia a mediana correspondente ao segundo quartil Q_2 fica desta forma:

Figura 11 – Pesos de varones y hembras



Fonte: BATANERO; GODINO, 2001, p. 56.

Dentre as características da mediana, Batanero e Godino (2001, p. 47-49, tradução nossa) apontam:

- a) Ao calcular a mediana, não usamos todos os valores observados da variável, o que a limita como uma medida de tendência central.
- b) Não pode ser aplicada às distribuições de variáveis qualitativas.
- c) Como medida de tendência central, apresenta certas vantagens sobre a média, em algumas distribuições, uma vez que não é afetada por valores extremos das observações. A mediana não varia se um valor inferior a ela for diminuído ou se um valor superior a ela for aumentado, pois somente os valores centrais da variável são levados em consideração. É adequada para distribuições assimétricas ou quando existem valores extremos.
- d) Preserva mudanças de origem e escala. Se somarmos, subtrairmos, multiplicarmos ou dividirmos cada elemento do conjunto de dados pelo mesmo número, esta operação é transportada para a mediana. Isso significa que ela é expressa na mesma unidade de medida que os dados.
- e) A mediana é uma medida estatística resistente: com pequenas flutuações na amostra, seu valor não muda. Um ou mais dados podem ser alterados sem alterar o valor de mediana, basta não modificar as duas partes do mesmo tamanho em que divide a distribuição.
- f) Se os dados são ordinais, a mediana existe, enquanto a média não faz sentido, já que seu cálculo é baseado nos valores (numéricos, necessariamente) dos dados.
- g) Para dados agrupados em intervalos, com um deles aberto, também é preferível a mediana que a média. Nesses casos, o intervalo aberto é omitido ou não é possível calcular a média, pois uma das marcas da classe, a correspondente a esse intervalo, estaria ausente.

A mediana apresenta vantagens e desvantagens em relação à moda ou a média. Segundo Lorí Vialí (2016), por exemplo, ela é utilizada junto ao desvio médio absoluto (dma) para calcular o coeficiente de dispersão (CD), como veremos mais adiante. Contudo, o material didático distribuído aos alunos sujeitos de nossa pesquisa privilegia amplamente o uso da média aritmética, em detrimento de outras medidas de dispersão que permitem a percepção da variabilidade.

Vamos, então, tratar dessa medida de tendência central.

4.1.1.3 Média aritmética simples

A compreensão da média aritmética simples, doravante tratada apenas como média, está profundamente arraigada no cotidiano das pessoas, em nossa sociedade, como observa Irene Cazorla (2002, p. 29):

É comum ler nos jornais ou ouvir nas reportagens frases do tipo: a renda per capita do Nordeste é inferior à do Sudeste, a expectativa de vida da mulher é maior que a do homem, ou informações referentes à chuva média mensal, à escolaridade média, ao número médio de filhos por casal e assim por diante. As pessoas estão acostumadas a estimar o tempo médio gasto no percurso de casa para o trabalho, o tempo médio que demoram para fazer as compras do mês no supermercado, o tempo despendido na fila de banco dentre outras estimativas.

Segundo Elon Lima et al. (2006), a média representa uma medida de posição de extrema importância na Matemática, isto é: “A média de uma lista de números é um valor que pode substituir todos os elementos dessa lista sem alterar uma certa característica da lista.” (ibid., p. 138). Os autores ainda ressaltam que existem várias médias: aritmética (simples ou ponderada), geométrica, quadrática e harmônica.

Trataremos, neste momento, da média aritmética. Mais à frente, quando abordarmos o conceito de desvio padrão, a média quadrática será utilizada.

Assim, Cobo (2003) apresenta duas definições similares a essas para a média, sendo soma ponderada dos valores possíveis para as frequências e média aritmética de um conjunto de dados. Novaes e Coutinho (2009, p. 80) a definem como “um ponto de equilíbrio dos desvios dos valores da distribuição”. E, para Morettin e Bussab (2013, p. 35), ela resulta da “soma das observações, dividida pelo número delas”.

Karla Priscila Schreiber, Gabriela Braz, Cristina Cavalli Bertolucci, João Alberto da Silva e Mauren Porciúncula (2019) lembram que essa medida-resumo embasa o cálculo de outras medidas estatísticas, como a variância, o desvio padrão, o coeficiente de variação, a assimetria, a curtose e a correlação. Os autores ressaltam a importância, para o letramento estatístico, de que o ensino da média aritmética não esteja restrito ao mecanismo de cálculo, mas que envolva situações significativas, nas quais os alunos possam valer-se de outros recursos, como as propriedades da média, identificadas por Sidney Strauss e Efraim Bichler (1988). Além disso, para Schreiber et al. (2019), a compreensão dessa medida pode se dar, inclusive, com bases nos dados concretos presentes em situações em que esses estudantes consigam visualizar a distribuição e abstrair o conceito de média.

Mabel Marques, Gilda Guimarães e Verônica Gitirana (2011, p. 727) discutem sobre as propriedades da média — com base nas ideias de Strauss e Bichler (1988) —, enunciando o reconhecimento dos alunos sobre:

1. a média está localizada entre os valores extremos;
2. a soma dos desvios a partir da média é zero;
3. a média é influenciada por cada um e por todos os valores;
4. a média não necessariamente coincide com um dos valores que a compõem;
5. a média pode ser um número que não tem um correspondente na realidade física, ou seja, não corresponde a um valor possível da variável considerada;
6. o cálculo da média leva em consideração todos os valores inclusive os nulos e os negativos;
7. a média é um valor representativo dos dados a partir dos quais ela foi calculada.

As autoras acrescentam, ainda, os quatro significados atribuídos à média:

- a. estimativa de uma quantidade desconhecida, em presença de erros de medida;
 - b. obtenção de uma quantidade equitativa a repartir para conseguir uma distribuição uniforme;
 - c. aplicação de que a média serve de elemento representativo de um conjunto de valores dados, cuja distribuição é aproximadamente simétrica;
 - d. necessidade de conhecer o valor que se irá obter com maior probabilidade ao contar com um dado faltando em uma distribuição.
- (MARQUES; GUIMARÃES; GITIRANA, 2011, p. 727-728)

Em sua pesquisa acerca da compreensão do conceito de média aritmética, diagnosticaram a necessidade de um trabalho sistematizado sobre sua presença nos anos iniciais de escolarização. Tanto professores quanto alunos precisam compreender melhor o conceito de média.

Embora o trabalho de Marques, Guimarães e Gitirana (2011) esteja focado nos anos iniciais do ensino fundamental, o reconhecimento de que esse assunto não tem sido tratado de forma adequada nesse segmento implica em reflexos drásticos nas etapas posteriores do ensino e da aprendizagem da Estatística. Tais significados também não são imediatos para a maioria dos alunos do ensino médio e estão presentes em diversos obstáculos epistemológicos, os quais estudaremos nesta pesquisa.

A esse respeito, Galicia (2009) afirma que os novos currículos, sobretudo na Espanha e nos Estados Unidos da América, recomendam abordagens na perspectiva da AED com alunos da secundária (ensino médio), a qual dá mais peso para o uso da mediana, destacando valores atípicos (*outliers*) e representações gráficas baseadas

na mediana — como o gráfico *boxplot* — e para outras medidas de posição — como os quartis.

Vale ressaltar que o conhecimento das medidas de tendência central, em especial a média aritmética, são de extrema importância para a vida do indivíduo fora da escola, como enfatiza a autora, pois apresentam importantes aplicações no cotidiano dos alunos — como PIB *per capita*, taxa de natalidade, custo ou índice de preços e taxas de desemprego —, além de aplicações em contextos escolares nas demais disciplinas — como para calcular o centro de gravidade em Física; o baricentro em Geometria; e uma partilha equitativa em situações de definição de renda *per capita* ou a densidade demográfica em Geografia. Reconhecer as dificuldades dos estudantes na compreensão dessas medidas é o primeiro passo para superá-las.

4.1.1.4 Articulando média, mediana e moda

As três medidas de posição até aqui abordadas são determinadas de forma distinta, de acordo com a distribuição — com dados não agrupados e agrupados sem intervalos de classe ou com intervalos de classe. É relevante discutir com os alunos quando convém ou não organizar os dados dessa ou daquela maneira, mas o importante é saber quando e por que utilizar cada uma delas e, mais ainda, como inseri-las de forma significativa e útil na análise estatística.

Isso dependerá, segundo Novaes e Coutinho (2009), dos objetivos da pesquisa bem como da forma que os dados se apresentam. As autoras explicam que a média não pode ser obtida sobre variáveis qualitativas, que a mediana pode ser mais adequada quando existe uma grande variabilidade entre os dados e que a moda pode ser mais útil quando queremos colocar um valor em destaque diante dos demais.

Cobo (2003) ressalta que média, mediana e moda são representantes de um coletivo e, como tal, proporcionam informações dos conjuntos de dados, não de elementos concretos distintos. Morettin e Bussab (2013) reforçam essa ideia, ao afirmarem que a natureza da variável condiciona a medida de posição a ser utilizada, por exemplo: para variáveis qualitativas nominais, só podemos utilizar a moda, enquanto para qualitativas ordinais, é possível utilizar moda e mediana.

A opção por uma ou outra medida-resumo, assim como pela forma de calculá-la, envolve, diretamente, um importante elemento do pensamento científico, apresentado por Wild e Pfannkuch (1999): a transnumeração, que se refere à mudança da representação dos dados para chegar a uma compreensão mais

profunda do fenômeno investigado. De acordo com os autores, a transnumeração se faz presente em todas as análises estatísticas de dados e a cada mudança de olhar sobre os dados estatísticos.

Para Coutinho, Silva e Almouloud (2011, p. 500):

O trabalho com a transnumeração é uma ideia fundamental de um enfoque estatístico para a aprendizagem e consiste em formar e mudar representações de dados de aspectos de um sistema para chegar à melhor compreensão desse sistema.

A transnumeração está presente na reclassificação e reorganização dos dados, na escolha da forma de calcular uma medida-resumo, assim como no cálculo das médias dos dados agrupados e não agrupados, ou na escolha da forma de representar dados brutos em rol ou transformados em tabelas e gráficos. Mesmo que a escolha seja de fazer a representação por meio de um gráfico estatístico, a seleção do tipo de gráfico ou a mudança de um gráfico para outro implica em transnumeração.

Para tanto, nos interessa, nesta pesquisa, identificar e caracterizar as escolhas transnumerativas realizadas espontaneamente pelos alunos, pois elas nos permitem analisar os conceitos e concepções mobilizadas por eles.

4.1.2 Medidas de dispersão

Enquanto as medidas de tendência central indicam os valores em torno dos quais os dados estão distribuídos, as medidas de dispersão — as quais são valores maiores ou iguais a zero, sendo que o zero denota a total ausência de dispersão — indicam como os dados estão distribuídos ao redor das medidas de tendência central.

Ademais, medidas de dispersão são medidas-resumo que permitem avaliar a variabilidade de um conjunto de dados, ou seja, os desvios das observações em relação à média das observações, segundo Morettin e Bussab (2013), tais como desvio médio, variância, desvio padrão, coeficiente de variação e intervalo interquartílico (SILVA, 2007).

Para Batanero (2001), o estudo de uma distribuição de frequências não pode se limitar às medidas de tendência central, até porque diferentes medidas de tendência central podem implicar em diferentes graus de variabilidade. Um erro frequente é ignorar a dispersão de dados ao comparar duas ou mais amostras ou populações. A média sem uma medida de dispersão associada não tem o poder de

caracterizar uma distribuição de frequência, e uma amostra bastante heterogênea pode apresentar a mesma média que uma amostra homogênea.

Por exemplo, imaginemos dois alunos A e B, sendo que o primeiro obtém notas bimestrais 5, 4, 5 e 6, enquanto o segundo, 2, 4, 6 e 8. Ao analisar o desempenho escolar de ambos, nos chama a atenção a regularidade de A e o crescimento das notas de B. Embora ambos tenham obtido a mesma média anual, certamente serão observados, pelos pais e professores, de modos muito distintos. Logo, compreender essa variabilidade é essencial para compreender a Estatística.

Sobre variabilidade, Maxine Pfannkuch e Dani Ben-Zvi (2011, p. 326, tradução nossa) ressaltam que:

A variabilidade é onipresente em todo o ciclo investigativo estatístico e é fundamental para o pensamento estatístico (Moore, 1990). Na verdade, "a solução de problemas estatísticos e tomada de decisão dependem de compreender, explicar e quantificar a variabilidade nos dados" (Franklin et al., 2005, p. 6). Os professores precisam aprender sobre como lidar com muitas fontes de variação inerentes aos dados e como usar ideias de variação para projetar a investigação.

O reconhecimento da importância da variabilidade é fundamental no ensino e aprendizagem da Estatística.

A seguir, apresentaremos algumas medidas de dispersão, contempladas no Currículo do Estado de São Paulo (SÃO PAULO, 2012).

4.1.2.1 Variância

Nos livros didáticos, bem como no Caderno do Aluno (SÃO PAULO, 2014), a variância é apresentada apenas como um passo intermediário para chegar ao desvio padrão.

Segundo Novaes e Coutinho (2009), a fim de evitar que valores negativos sejam abatidos dos positivos, anulando a soma dos desvios, em vez de considerá-los em módulo, elevamos cada um dos desvios em relação à média ao quadrado e calculamos a média aritmética desses quadrados.

Silva (2007, p. 44) define, assim, a variância:

A variância de uma variável aleatória X é uma medida de dispersão dos valores de uma variável em torno da sua média. Ela é definida como a média aritmética dos quadrados dos desvios dos valores de X ($X_1, X_2, \dots, X_N$) em relação à média dos próprios valores de X .

Chris Reading e Michael Shaughnessy (2004) destacam que a complexidade nos cálculos necessários para a obtenção dessa medida podem ser um obstáculo no estudo da Estatística. Contudo, da mesma forma que podem desmotivar os alunos, esses cálculos podem representar uma boa oportunidade para mobilizar conceitos matemáticos elementares, como potenciação, radiciação, fração, além de enriquecimento da linguagem matemática com a introdução do símbolo de somatório.

Batanero e Godino (2001) observam que a variância não muda quando realizamos uma inversão, ou seja, quando adicionamos ou subtraímos a mesma quantidade de todos os dados. Segundo os autores, o uso da variância como medida de dispersão traz um inconveniente: ela não é expressa na mesma unidade de medida que a variável estudada. Por esse motivo, é mais comum a opção pelo uso do desvio padrão.

4.1.2.2 Desvio padrão

O desvio padrão é utilizado para medir a dispersão de um conjunto de valores em torno de sua média. Quanto menor o desvio padrão, mais homogênea é a população (DIAZ, 2016). Geralmente apresentado como a raiz quadrada da variância, mede a propagação e a variabilidade ao redor da média. A autora ressalta, também, que essa medida é preferida em detrimento à variância, porque esta não é expressa na unidade de variável, já que o valor que assume é ao quadrado e dificulta a interpretação da sua ordem de magnitude.

Silva (2007) afirma que essa é a medida de variação mais conhecida bem como a mais utilizada na Estatística. Segundo a autora, a maioria dos livros didáticos apresenta o desvio padrão e outras medidas de dispersão por meio de uma tabela de distribuição de frequências. O desvio padrão mede a intensidade com que os dados se desviam da média.

Além disso, Batanero e Godino (2001) observaram que muitos livros didáticos enfatizam mais a heterogeneidade do que seu desvio da posição central, por meio de palavras como variação, dispersão, diversidade e flutuação. Isso pode acarretar dificuldades para a compreensão do aluno a respeito de variabilidade, como discutiremos a partir de algumas questões resolvidas, no capítulo 5 desta tese.

4.1.2.3 Desvio médio absoluto

A média dos valores absolutos dos desvios em relação à média é o chamado desvio médio absoluto (dma). De acordo com Morettin e Bussab (2013), subtraímos a média aritmética de cada valor do conjunto de dados e, para evitar que valores negativos sejam abatidos dos positivos, anulando sua soma, consideramos apenas seus valores em módulo.

Silva (2007) ressalta que, mesmo utilizando a média aritmética como referência, é a mediana que permite uma compreensão melhor sobre variabilidade. Segundo a autora, o desvio médio absoluto é pouco explorado nos livros de Matemática e Estatística. Para Maurice Kendall e George Yule (1948), o fato de as medidas de dispersão mais comuns serem representadas em unidades da própria variável inviabiliza operações com unidades distintas. De forma a permitir comparações entre valores que não dependam das respectivas unidades, é possível operar com o coeficiente de variação que resulta das divisões entre desvio médio e média aritmética, desvio médio e moda ou desvio padrão e média aritmética.

De acordo com Viali (2016, p. 10), o dma corresponde à “média das distâncias que os valores do conjunto se encontram da média”. Esse desvio pode ser utilizado para determinar o coeficiente de dispersão (CD), razão entre o desvio médio absoluto da mediana e a mediana dos dados.

Por exemplo, em $CD = \frac{1}{n} \frac{\sum |m-x|}{m}$, n é o tamanho da amostra, m é a mediana da amostra, $\sum$ corresponde ao somatório e x é a variável.

4.1.2.4 Amplitude total

Viali (2016) expõe que a amplitude total (h) é a diferença entre o valor máximo ($x_{\max}$) e o valor mínimo ($x_{\min}$) de uma distribuição de frequência, logo, $h = x_{\max} - x_{\min}$. Para Silva (2007), é uma medida de variação simples e eficaz, pois permite observar a variação total de uma distribuição de frequência. Além disso, dentre suas aplicações, podemos destacar que ela permite o planejamento e organização de intervalos de classes para as representações gráficas e tabulares.

Vale ainda destacar que a amplitude total não depende de uma medida central como ponto de referência, tal qual a média aritmética ou mediana. No entanto, Viali (2016) ressalta que, a despeito da facilidade em calculá-la, não considera todos os valores do conjunto, apenas os extremos. Assim, geralmente optamos por trabalhar com medidas que utilizem toda a informação disponível.

4.2 Registros de representação: tabelas de frequência e gráficos estatísticos

Dada a sua grande presença na mídia, o aprendizado de tabelas e gráficos estatísticos tem conquistado espaço na educação, como podemos ver claramente na BNCC (BRASIL, 2018). Um exemplo diretamente relacionado às nossas investigações foi o fenômeno das eleições majoritárias no Brasil em 2018, ano em que realizamos a coleta de dados da nossa pesquisa. Durante o período, era possível ver gráficos, tabelas e quadros sobre intenção de voto e tendências do eleitorado brasileiro quase todos os dias na televisão, nos jornais impressos e na internet.

Nos estudos sobre registros de representação, uma das figuras mais eminentes é Raymond Duval¹¹, cuja Teoria dos Registros de Representação Semiótica (TRRS) tem sido amplamente empregada nas mais variadas pesquisas da Educação Matemática. Segundo o filósofo, um dos papéis da ensino da Educação Matemática é contribuir para o desenvolvimento das capacidades de raciocínio, análise e visualização dos alunos (DUVAL, 2003).

De forma sucinta, Bernard Coutanson (2010, p. 160, tradução nossa) apresenta o conceito de registro de representação semiótica do teórico:

[...] um registro semiótico atende às três características a seguir:

- de comunicação (ideia inequívoca e óbvia, sem necessidade de esclarecimentos-chaves de leitura, para apresentar as informações);
- de tratamento (possível com os dados contidos graças à própria ferramenta exposição utilizada);
- de conversão (possibilidade de traduzir os dados expostos em outro registro, com a possibilidade de levá-los um por um e convertê-los da mesma maneira em outro registro).

O autor destaca que, para Duval (apud COUTANSON, 2010), se faz necessário recorrer, no mínimo, “ao confronto de representações fornecidas por dois registros diferentes”. Como exemplos, Coutanson (2010) destaca listas, linguagem natural, tabelas, diagramas, gráficos, mapas e ilustrações.

Além disso, ele nos propõe uma reflexão sobre as questões: “Como aprendemos a passar de um registro para outro? Como aprendemos a diferenciar um objeto da representação que oferecemos?” (ibid., p. 160). Na Estatística, em particular, essas perguntas sempre vêm à tona no momento de organização e apresentação dos dados, ao optar por uma tabela ou um gráfico ou ao escolher um

¹¹ Raymond Duval é um filósofo e psicólogo francês que desenvolveu importantes estudos no campo da Psicologia Cognitiva.

determinado tipo de gráfico — aquele que melhor se adequa aos dados possuídos bem como aos objetivos do pesquisador. O autor ainda acrescenta:

Uma atividade matemática é um ato de fusão, de acordo com dois tipos de transformação:

- o primeiro diz respeito a qualquer alteração no registro de representação;
- o outro consiste em as possibilidades de transformação específicas para cada registro.

(COUTANSON, 2010, p. 160, tradução nossa)

Regina Damm (2012) destaca que “toda a comunicação se estabelece com base em representações” e, em Matemática, em particular, seus objetos de estudo são “conceitos, propriedades, estruturas, relações que podem expressar diferentes situações”. A autora também afirma que: “Não existe conhecimento matemático que possa ser mobilizado por uma pessoa sem o auxílio de uma representação.” (ibid.).

Na Matemática — e por extensão na Estatística —, os objetos de estudo nem sempre são acessíveis à percepção imediata, exigindo a manipulação de uma linguagem rica em símbolos, signos, diagramas, códigos, tabelas, gráficos, algoritmos, desenhos etc. Então, para garantir o acesso ao conhecimento matemático, o professor busca, muitas vezes, abordá-lo de diferentes maneiras e sob diferentes pontos de vista, exigindo, para tanto, representações distintas. Assim, quando o aluno realiza com destreza as transformações necessárias entre os diferentes registros, demonstra domínio sobre o objeto estudado.

Os registros requerem um nível de funcionamento consciente — comunicação, objetivação e tratamento. Os códigos, por sua vez, mobilizam funções cognitivas não conscientes. Consequentemente, se revelam necessárias as mudanças de registro bem como a coordenação entre eles (ALMOULOU, 2007).

Essas mudanças podem ocorrer dentro de uma mesma representação (em língua natural, numérica, figural/geométrica, algébrica, gráfica, algorítmica e/ou tabular) — denominadas tratamentos — ou entre representações diferentes — chamadas de conversões —, como observado por Márcia Vieira (2008).

Damm (2012), nos lembra que mais importante do que o registro de representação é a forma como ele é utilizado. Contudo, a existência de muitos registros requer uma leitura mais atenta, ou seja, um refinamento na percepção e cognição sobre o objeto, pois cada registro oferece vantagens e desvantagens, e nem todas as suas formas são adequadas aos objetos de estudo.

Na Educação Estatística, essa compreensão é fundamental. A escolha entre um registro numérico, tabular ou gráfico implica em possibilidades e limitações distintas. Mesmo dentro de uma mesma espécie de registro, existe, aparentemente, um grande número de opções — por exemplo, em gráfico, há gráfico de barras verticais ou colunas, barras horizontais, histograma, setores, linha ou curva, *boxplot*, *dot plot*, ramo e folha, pictograma, cartograma, polar etc.) — e, na verdade, dada a natureza das variáveis a serem representadas e dos objetivos do estudo, as escolhas não são tantas. Assim, em um registro tabular, podemos optar por representar a frequência simples ou a acumulada, a absoluta ou a relativa ou mesmo em um rol.

Para Coutinho, Silva e Almouloud (2011, p. 500):

A compreensão da Estatística implica em transitar entre os diversos registros de representação semiótica envolvidos no processo de construção de conceitos estatísticos, seguindo, assim, os princípios da Análise Exploratória de Dados.

A riqueza de possibilidades a serem exploradas através das mudanças de registro na Estatística tem, no reverso da moeda, as dificuldades na escolha e ajustes a cada registro. Vieira (2008, p. 23) afirma que

Pensando em Estatística, poderíamos considerar como tratamento a passagem da enumeração em série de rol para uma distribuição de frequências ou, no caso de variáveis qualitativas ou quantitativas discretas, a mudança de escala em um gráfico.

Gal (2002) nos lembra que, para construir, ler e interpretar tabelas e gráficos estatísticos, algumas competências são requeridas, tais como a capacidade de interpretar e de avaliar criticamente dados ou argumentos embasados em fenômenos estatísticos, presentes em diversos lugares — incluindo a mídia, mas não limitado a ela —, discutindo e/ou comunicando resultados e opiniões a respeito delas.

A linguagem gráfica, segundo Pedro Arteaga et al. (2009), é fundamental para a organização, descrição e análise de dados como instrumento de transnumeração, uma forma de raciocínio estatístico básico que consiste em alcançar novas informações com base nas transformações de um sistema de representação em outro. Por exemplo, baseado em uma lista de dados não ordenados a um histograma, visualiza-se a moda e percebe-se a simetria ou assimetria da distribuição.

Concordando com Duval, Pfannkuch e Ben-Zvi (2011) destacam a importância do reconhecimento de padrões nos dados buscando os significados escondidos neles e por trás deles. De acordo com os autores, o raciocinar com base nos gráficos requer

a capacidade de perceber, decodificar, avaliar e julgar as mensagens inerentes ao contexto.

Sobre a importância de tabelas e gráficos para nossa sociedade, Arteaga et al. (2011, p. 57) acrescentam:

A importância de tabelas e gráficos, se deve também ao fato de que a ciência as utiliza como representações semióticas externas para construir e comunicar conceitos abstratos. Portanto, a aprendizagem de conceitos científicos está ligada tais representações e a seus processos de construção e de transformação. Essas representações são também utilizadas na ciência como uma ponte entre os dados experimentais e formalização científica e ajudam a determinar as relações entre as variáveis que intervêm nos fenômenos, para modelá-los.

Em conformidade com os autores, Batanero et al. (2013) asseveram que tabelas e gráficos são um instrumento essencial da transnumeração, graças a seu papel na descrição, organização e análise de dados.

Aliás, tabelas e gráficos são ferramentas poderosas para apresentar a distribuição de frequência das variáveis envolvidas na pesquisa. Segundo Pfannkuch e Ben-Zvi (2011, p. 326, tradução nossa): “[...] distribuições são um elemento fundamental do raciocínio estatístico uma vez que os estatísticos olham a variação através de uma lente que é distribuição”.

Trataremos, primeiramente, dos gráficos estatísticos.

4.2.1 Gráficos estatísticos

Boa parte dos professores, conforme observado por Batanero (2001), julgando a elaboração de tabelas e gráficos muito simples, dedicam pouco tempo a seu ensino. Para os alunos, contudo, isso não é tão simples, a começar pelos diferentes tipos de frequência neles representados: absoluta e relativa, simples e acumulada.

Com o fácil acesso a *softwares* e aplicativos gratuitos ou de baixo custo, escolher e trocar um determinado tipo de gráfico ficou muito simples. Entretanto, é necessário compreender bem as propriedades de cada tipo, para que a escolha acertada seja feita. Régnier (1998, p. 1, tradução nossa) faz uma profunda reflexão sobre um tipo muito popular de gráfico estatístico, o histograma:

Ao se tornar um objeto de ensino, podemos nos perguntar sobre as deformações sofridas pelo conceito de domínio da estatística matemática neste processo de transposição didática. O termo histograma é amplamente utilizado em vários livros didáticos, em alguns *softwares* como planilhas ou *softwares* estatísticos. As principais perguntas que nos guiam são: Para atender a quais objetivos planejamos ensinar o conceito de histograma? Qual

é o significado da noção de histograma para o professor quando é realmente ensinado? Que(ais) sentido(s) essa noção leva no aluno?

Analisando suas definições, desde sua origem histórica, passando pelos dicionários até os livros didáticos atuais franceses, o autor alerta, ainda, para a necessidade de haver clareza acerca dos objetivos do estudo bem como das propriedades do gráfico ao tomar a decisão de construir um histograma, pois essa construção “não se limita a obtenção de um gráfico figurativo para fins estéticos e decorativos” (RÉGNIER, 1998, p. 8).

Também, segundo ele, a falta de compreensão do professor sobre um tipo de gráfico e a discussão das propriedades gráficas com seus alunos “obstrui o entendimento do histograma quando ele tiver que ser usado de outra maneira que não a de uma ilustração que funcione como gráfico de barras” (ibid., p. 20). O autor conclui ser necessário, além de atrair a atenção de docentes de Matemática, envolver nessa discussão docentes de outras disciplinas que também fazem uso de gráficos estatísticos.

Mas a noção de histograma também é assunto de muitas outras disciplinas acadêmicas e seria útil conectar nossa reflexão a outras similares, conduzidas em Geografia, Física ou Biologia, por exemplo. Nós permanecemos certo de que situações didáticas relevantes podem ser organizadas pelos professores. É a esse preço que nossas conjecturas serão validadas ou refutadas e que nosso objetivo não será dogmatizado. (RÉGNIER, 1998, p. 24, tradução nossa)

Poderíamos ampliar essa discussão envolvendo outros tipos de gráficos estatísticos, mas pretendemos, aqui, apenas endossar o ponto de vista de Batanero (2001) de que os professores devem dedicar uma maior atenção ao ensino dos gráficos, de modo a permitir uma leitura mais profunda deles.

Isso também é proposto por Frances R. Curcio (1989), que descreveu três níveis diferentes de entendimento dos gráficos, os quais podem ser aplicados a tabelas e gráficos estatísticos, sendo:

- Nível 1 - Leitura dos dados – requer uma apreensão literal das informações contidas no gráfico, não sua interpretação. A leitura do gráfico requerida neste nível tem um custo cognitivo muito baixo;
- Nível 2 - Leitura entre os dados – inclui a interpretação e integração dos dados dos gráficos e requer a capacidade de comparar as quantidades e o uso de outros conceitos e habilidades matemáticas, ou seja,

habilidades que permitam ao leitor combinar e integrar dados e identificar as relações matemáticas expressas no gráfico, exigindo, minimamente, um certo grau de inferência;

- Nível 3 - Leitura além dos dados – requer do leitor previsões e inferências com base nos dados sobre informações que não são apresentadas diretamente no gráfico, extraíndo novos esquemas e padrões.

Susan Friel, Curcio e George Bright (2001) ampliaram essa classificação, incluindo um quarto nível:

- Nível 4 - Leitura por trás dos dados – requer a avaliação sobre a confiabilidade e integridade dos dados, o que implica em avaliar criticamente o método de coleta de dados, sua validade e confiabilidade bem como as possibilidades de estender as conclusões.

Arteaga et al. (2009, p. 101, tradução nossa) alertam que:

[...] a simplicidade da linguagem gráfica é aparente, pois mesmo o mais simples dos gráficos pode ser considerado um modelo matemático. Ao reduzir os dados, com base nos valores individuais de uma variável, se introduz a distribuição de frequência, conceito complexo que se refere ao agregado (população ou amostra) e não a dados em particular. Por outro lado, o mesmo tipo de gráfico (por exemplo, um gráfico de barras simples) pode ser usado para representar diferentes objetos matemáticos, tais como frequências absolutas, percentuais relativos e frequências acumuladas, médias ou outras medidas-resumo estatísticos. Gráficos estatísticos estão presentes no dia a dia, na mídia e na Internet, em livros didáticos de diferentes materiais e no trabalho profissional.

Jim Ridgway, James Nicholson e Sean McCusker (2008, p. 65, tradução nossa) afirmam ser necessário ensinar aos alunos as seguintes heurísticas, diretamente relacionadas à compreensão dos gráficos e tabelas:

- Ser crítico com a fonte de dados, exigindo qualidade dos dados.
- Identificar as variáveis do estudo, seu tipo (qualitativa ou quantitativa) e seu papel no estudo (dependente ou independente).
- Descrever e explorar os dados a fundo antes de tentar obter conclusões.
- Buscar relações não lineares entre as diferentes variáveis e mudanças ao longo do tempo.
- Em caso de levar a cabo um estudo de inferência, avaliar o efeito das variáveis explicativas (tamanho da diferença das médias entre os grupos analisados), e não apenas o significado estatístico dos dados. Comprovar, também, mediante cálculo de intervalos de confiança, que as diferenças da variável dependente nos grupos são substancialmente maiores do que aquelas do erro aleatório.

Essa criticidade tão almejada, essa compreensão da realidade e esse nível de raciocínio estatístico não surgem facilmente nas situações didáticas tradicionais. Para isso, se faz necessário o desenvolvimento de uma abordagem contextualizada que coloque o aluno como figura central do processo de investigação, como ator na pesquisa, e não como receptor passivo.

Sobre a importância da contextualização dessa ciência, Batanero et al. (2013) ressaltam que a Estatística é inseparável de suas aplicações e que sua justificativa final em diversos meios, inclusive no escolar, pode ser atribuída à sua utilidade na resolução de problemas externos à própria Estatística.

A transnumeração — pensamento estatístico já mencionado nesta tese — depende da diversidade de recursos disponíveis no repertório do aluno. Dessa forma, para ampliar esse repertório e desenvolver o pensamento estatístico, Coutinho (2017) sugere a inserção de duas novas representações gráficas nas aulas de Estatística: *dot plot* e *boxplot*. Além disso, propõe a utilização do *software* GeoGebra para a exploração de técnicas transnumerativas e o aprofundamento da análise gráfica.

O Currículo do Estado de São Paulo (SÃO PAULO, 2012) não menciona explicitamente os tipos de gráficos que devem ser estudados no ensino médio, como vemos na BNCC (BRASIL, 2018), apenas trata deles de modo vago, por meio da habilidade de “Saber construir e interpretar tabelas e gráficos de frequências com base nos dados obtidos em pesquisas por amostras estatísticas.” (SÃO PAULO, 2012, p. 70).

O Caderno do Aluno (idem, 2014) traz somente gráficos de barras/colunas/histogramas e gráficos de linha/curva. Além disso, as avaliações de aprendizagem em processo (AAP) — avaliações institucionais elaboradas pela Secretaria da Educação do Estado de São Paulo e aplicadas bimestralmente nas escolas dessa rede de ensino — incluíam ainda gráficos de setores.

Nesse aspecto, concordamos com Coutinho (2017): isso não basta, se faz necessário explorar melhor a riqueza de recursos oferecida pelas representações gráficas, sempre que possível, empregando recursos tecnológicos digitais, como o *software* GeoGebra. Ademais, a autora propõe a introdução do estudo do gráfico de pontos e do *boxplot*, ideia que começaria a ser introduzida nos currículos brasileiros a partir da implementação da BNCC em 2018.

A seguir, abordamos as tabelas estatísticas.

4.2.2 Tabelas de frequência

Junto aos gráficos, as tabelas constituem o principal meio de representação de informação (PALLAUTA; GEA; VENEGAS, 2019). Segundo esses autores, que analisaram coleções de livros didáticos chilenos, as tabelas estatísticas não recebem a atenção que merecem no currículo e no ensino, pois observaram que as tabelas estatísticas de dupla entrada têm escassa presença nos livros didáticos, chegando à ausência em algumas coleções. Isso vale tanto para o Chile quanto para o Brasil, de acordo com a nossa revisão da literatura.

Danilo Díaz-Levicoy et al. (2016) afirmam que as tabelas, como elementos da cultura estatística, devem ser contempladas nos currículos formais, garantindo presença nas salas de aula desde a educação infantil, uma vez que, além de comporem problemas de Estatística e Probabilidade, são encontradas em livros didáticos de outras disciplinas escolares bem como na apresentação de dados das mais diversas mídias.

Para Batanero (2001), há três tipos de problemas simples de associação estatística: tabelas de contingência, de correlação numérica e de comparação de uma variável numérica em diferentes amostras — independentes ou relacionadas. As tabelas servem para resumir a distribuição de frequências de uma população ou amostra classificadas em relação a duas variáveis estatísticas.

Segundo a autora, apesar da aparente simplicidade, cruzar dados estatísticos pode constituir um desafio complexo para o aluno, pois, inicialmente, ele deve distinguir os casos de associação direta, de associação inversa ou de independência. Devemos levar em conta a capacidade do estudante de, conforme o contexto, considerar a existência ou não de relação com a associação empírica na tabela, além de suas próprias crenças sobre a possível associação que é esperada no contexto dado — crenças essas que podem conduzi-lo ao que Batanero (2001) chama de correlação ilusória.

Howard Wainer (1995) propôs três níveis de compreensão na leitura e interpretação de tabelas: elementar, associado à capacidade do indivíduo de extrair dados pontuais das tabelas sem necessitar de comparação e análise; intermediário, relacionado à interpolação de dados em tabelas bem como à compreensão das possíveis relações entre eles; e avançado, referente à compreensão mais ampla da

estrutura dos dados apresentados em sua totalidade, avaliando tendências e relações implícitas nelas.

Adriana Pagan et al. (2008) detectaram as mesmas lacunas no processo de letramento estatístico em toda a educação básica. No Estado de São Paulo, podemos supor que isso seja um efeito da descontinuidade da presença da Estatística no currículo oficial.

4.3 Variabilidade

A variabilidade é um conceito central a toda e qualquer compreensão da Estatística bem como de suas aplicações. Mais do que isso, podemos dizer que se trata de um conceito fundador do quadro teórico da Análise Exploratória de Dados (TUKEY, 1970).

Moore (1997) reconhece a onipresença da variabilidade na Estatística. O autor destaca três elementos fundamentais para o ensino e a aprendizagem da Estatística: enfatizar os elementos centrais do pensamento estatístico — reconhecimento da necessidade de dados e importância da produção de dados (onipresença da variabilidade e a medida e modelagem da variabilidade); valorizar uma abordagem mais conceitual do que procedimental; e implementar estratégias de aprendizagem ativas.

Concordando com ele, Wild e Pfannkuch (1999, p. 235) chamam a nossa atenção para três características essenciais à sua compreensão: “a variabilidade é onipresente; a variabilidade pode ter sérias consequências práticas; e as estatísticas nos dão um meio de entender um mundo variado”.

Então, se a variabilidade é onipresente, está ao nosso redor, sendo facilmente observável, embora nem sempre bem compreendida. Graças a ela, vivemos em um mundo muito rico em diferenças, seja quanto à biodiversidade ou às variações de peças, mesmo quando feitas pelas mesmas máquinas e sob as mesmas condições de trabalho. Ainda que em condições de aparente simetria, ela está presente. No corpo humano, por exemplo, os braços, pernas ou olhos de uma pessoa parecem simétricos, mas raramente são.

Garfield e Gal (1999) consideram fundamental na investigação estatística, para se fazer previsões, compreender a relação entre as medidas de posição e a variabilidade de um conjunto de dados, identificando quais são as medidas mais apropriadas em cada caso e como elas representam tal conjunto.

Gary Kader e Mike Perry (2006) destacam que a resolução de problemas e a tomada de decisões baseadas em dados estatísticos dependem, diretamente, da análise e compreensão da variabilidade dos dados. Segundo os autores, é o foco na variabilidade dos dados que enfatiza a diferença entre a Estatística e a Matemática.

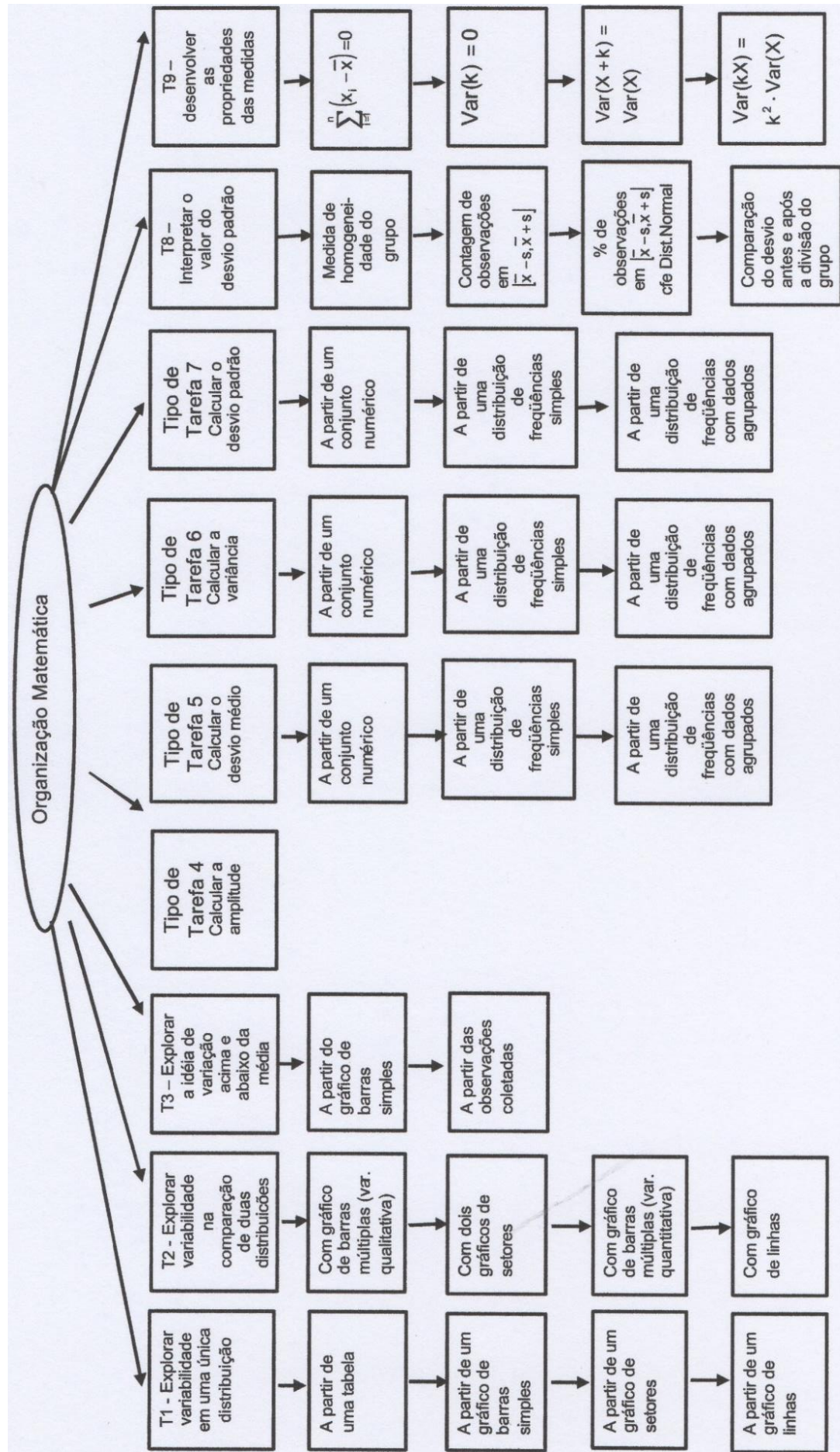
Embora muitos autores tratem os conceitos de variação e variabilidade como sinônimos, Silva (2007, p. 20) fez questão de diferenciá-los. Para ela, o termo "variabilidade" é "uma característica da entidade que é observável", ao passo que o termo "variação" consiste na "descrição ou medida desta característica" (ibid.). Em outras palavras, variabilidade é uma "forma substantiva do adjetivo variável" (ibid.).

Silva (ibid.) considera fundamental para o ensino e a aprendizagem da Estatística preparar os professores de Matemática para abordar a variabilidade desde os anos iniciais do ensino fundamental, por se tratar de conceito central dessa ciência, tanto para o letramento quanto para o raciocínio e pensamento estatísticos. Segundo a autora, a variabilidade está associada a todas as medidas-resumo da Estatística, sendo particularmente evidenciada nas medidas de dispersão, como variância e desvio padrão.

Uma de suas conclusões das investigações, que particularmente nos interessa aqui, é a de que poucos professores de Matemática promoviam a realização de pesquisas com seus alunos e, quando faziam, era de "maneira restrita, utilizando apenas distribuição de frequência e sua respectiva representação gráfica" (SILVA, 2007). Isso difere bastante tanto da proposta da abordagem da Estatística por meio de projetos quanto das orientações da BNCC (BRASIL, 2018).

A autora ainda elaborou um esquema representativo da presença da variabilidade nos livros didáticos:

Figura 12 – Organização matemática envolvendo variabilidade nos livros didáticos



Fonte: SILVA, 2007, p. 88.

Nos interessa, nesta pesquisa, o reconhecimento e a caracterização das concepções mobilizadas por alunos na resolução de tarefas similares a essas.

Ainda sobre atividades matemáticas envolvendo variabilidade, Campos (2007, p. 62) observa que:

Atividades que desafiam os estudantes a explicar o que faz um desvio padrão ser maior ou menor podem, por exemplo, ajudar no desenvolvimento do raciocínio sobre a variabilidade (DELMAS, 2001). Atividades que permitem uma simulação visual de amostras de uma população, variando o tamanho da amostra ou os parâmetros da população, ajudam os estudantes a desenvolver o raciocínio sobre distribuição amostral (GARFIELD, 2002). Se os professores estiverem atentos aos tipos de raciocínio que precisam reforçar em seus estudantes, podem promover atividades para ajudar a desenvolvê-los.

Silva, Verônica Kataoka e Cazorla (2011), ao discutirem o conceito de variabilidade, ressaltam que a variação é um dos componentes do pensamento estatístico e sem ela a Estatística não existiria, pois alguns teóricos a definem como a ciência da variação. O conceito de variação já fazia parte do currículo de Matemática no ensino médio, com base na publicação dos PCN de 1997, 1998 e 2000, e agora mais ainda, devido à implantação da BNCC em 2018. Segundo as autoras, os resultados de pesquisas recentes têm apontado na direção do desenvolvimento desse conceito, de modo intuitivo e precoce, em crianças.

Novaes (2011) investigou as concepções de professores da Educação Básica sobre variabilidade estatística, à luz da Teoria das Concepções, sendo esta a pesquisa brasileira que mais se aproxima da nossa.

Segundo Novaes e Coutinho (2011, p. 4), a mobilização de diferentes saberes estatísticos possibilita a transformação do conceito de variabilidade. Um nível de apreensão da variação é reconhecido na "sua medida em função da natureza da variação".

Batanero et al. (2013) destacam que, embora em diversos campos da Matemática as variáveis sejam utilizadas — sendo esperado que os dados se encaixem perfeitamente em um modelo prévio —, raramente se investiga a adequação a tal modelo. Na Estatística, o estudo da variabilidade assume um caráter peculiar, pois, conforme os autores, a Estatística permite um maior entendimento sobre variabilidade para fazer previsões. Assim, por meio de seu estudo, os alunos podem compreender melhor a natureza da variabilidade e gerenciar modelos que permitam controlá-la e prevê-la. Para explorar essa possibilidade, há algumas

estratégias de ensino que podem ser utilizadas, como a modelagem e a própria abordagem por meio de projetos.

A partir dessas considerações, iniciamos nossa coleta de dados com alunos envolvidos na resolução de problemas, após o desenvolvimento de projetos, na abordagem da Estatística.

Nesta seção, realizamos um estudo conceitual considerando os conteúdos estatísticos que os alunos do 3º ano do ensino médio deveriam conhecer para atender as demandas da rede estadual de ensino paulista — a saber, tabela e gráficos estatísticos; medidas de tendência central (média, mediana e moda); e medidas de dispersão (desvio médio e desvio padrão). Segundo o Currículo do Estado de São Paulo: Matemática e suas tecnologias, vigente em 2018, quando realizamos nossa coleta de dados, os estudantes deveriam, no 4º bimestre desse ano, desenvolver estas habilidades:

- Saber construir e interpretar tabelas e gráficos de frequências a partir de dados obtidos em pesquisas por amostras estatísticas.
- Saber calcular e interpretar medidas de tendência central de uma distribuição de dados: média, mediana e moda.
- Saber calcular e interpretar medidas de dispersão de uma distribuição de dados: desvio padrão.
- Saber analisar e interpretar índices estatísticos de diferentes tipos.

(SÃO PAULO, 2012, p. 70)

Assim, com base em nosso estudo do objeto, consideramos que nossos apontamentos sobre medidas de tendência central e de dispersão, registros de representação — tabela e gráficos estatísticos — e variabilidade — conceito que permeia toda a Estatística — foram essenciais para fundamentar nossas análises, apresentadas na seção a seguir.

5 O QUE OS ALUNOS SABEM SOBRE ESTATÍSTICA

Neste capítulo, apresentaremos e analisaremos alguns resultados obtidos nos questionários individuais dos alunos, os quais abordaram seus conhecimentos sobre Estatística, anteriores ao desenvolvimento de projetos, na perspectiva da ASI. Uma versão integral do questionário pode ser encontrada no Apêndice 1 desta tese.

Para atingir nossos objetivos, optamos pela análise estatística de dados multidimensionais, que nos permite organizá-los e sintetizá-los, evidenciando as relações entre os agentes envolvidos nos fenômenos estudados (ORÚS; ZAMORA; GREGORI, 2009). Utilizando o *software* CHIC¹², analisamos as respostas fornecidas por 86 alunos, com idades entre 16 e 19 anos, provenientes de 4 turmas do 3º e último ano do ensino médio de uma escola pública brasileira.

Esse *software* gerou grafos de similaridade, de implicação e de coesão das relações entre as variáveis identificadas pelo instrumento de coleta de dados, com base no cruzamento delas. A multidimensionalidade permitiu, assim, que várias características (variáveis) relativas a um mesmo grupo de sujeitos fossem simultaneamente consideradas na análise dessas pessoas quanto a um tema específico — em nosso caso, conhecimentos sobre Estatística.

O CHIC foi criado pelo pesquisador Régis Gras e seus colaboradores na década de 1990, com o objetivo de sintetizar e estruturar as respostas, caracterizando comportamentos por meio de um tratamento multidimensional de dados estatísticos. Tal *software* tem por funções essenciais extrair de um conjunto de dados, cruzando sujeitos e atributos, regras de associação entre eles, gerando um índice de qualidade de associação e uma estruturação das variáveis obtidas por meio dessas regras bem definidas — atendendo bastante nossas necessidades.

A opção pelo questionário também se mostrou adequada. Jean-Claude Oriol (2007, p. 259, tradução nossa) destaca que:

[...] a administração de questionários é uma tarefa fundamental para a apropriação de técnicas e do campo da estatística; ela não se limita ao simples tratamento dos dados e, nesse sentido, permite enriquecer toda as situações que servem como referência.

¹² Versão 7.0, copywrite 2014, licenciado para uso pelo grupo de pesquisa PEA-MAT na PUC-SP. O coordenador do grupo, o Prof. Dr. Saddo Ag Almouloud, foi o responsável pela primeira versão do *software*, durante sua pesquisa de Doutorado sob orientação do Prof. Dr. Régis Gras, criador do tipo de análise e do *software*.

O questionário foi aplicado em sala de aula pelo autor desta tese, durante 50 minutos, sendo organizado em 3 partes distintas: a primeira delas, composta por 1 questão aberta e 10 questões objetivas — Q1 a Q11, ver Apêndice 1 —, buscou identificar as concepções e conhecimentos dos alunos a respeito da Estatística, os quais contemplam o modelo de letramento estatístico proposto por Gal (2002), conforme apresentado no capítulo 3 desta tese.

Destacamos, ainda, que foi inclusa a variável suplementar ‘sexo’, a qual, embora não fizesse parte de nenhuma das 29 questões de nosso instrumento de coleta de dados, pôde ser apurada por identificação verbal do aluno durante as aulas.

A segunda parte do questionário, composta por 13 questões objetivas — Q12 a Q24, ver Apêndice 1 —, procurou identificar os conhecimentos estatísticos e indícios de letramento dos alunos a respeito de Probabilidade. No entanto, optamos, no desenvolvimento da pesquisa, por focar nas questões que abordassem apenas a Estatística, deixando essa análise sobre letramento probabilístico para futuros artigos.

Por fim, a terceira parte trazia 5 questões objetivas — Q25 a Q29, ver Apêndice 1 — extraídas das avaliações do Exame Nacional do Ensino Médio – ENEM — o qual possibilita o ingresso, bolsa e/ou financiamento de estudos no ensino superior brasileiro — dos últimos anos, o que permitiu investigarmos os conhecimentos estatísticos dos alunos referentes aos conteúdos específicos da Estatística, previstos no Currículo do Estado de São Paulo (SÃO PAULO, 2012). Vale ressaltar que os estudantes não haviam estudado Estatística no ensino médio nessa escola porque, segundo o documento, ela deve ser ensinada durante um único bimestre do 3º ano do ensino médio.

Uma vez que faremos uso, neste capítulo, das siglas que utilizamos na ASI, elaboramos o quadro a seguir para melhor compreendê-las:

Quadro 3 – Siglas utilizadas neste capítulo

Q1	Questão aberta: O que é Estatística? Ela é uma ciência exata ou uma ciência social? É uma área da Matemática?
Q2NL	Não se lembra se estudou Estatística no ensino fundamental.
Q2NE	Não estudou Estatística no ensino fundamental.
Q2EF	Estudou Estatística no ensino fundamental.
Q3NL	Não se lembra se estudou Estatística no ensino médio.
Q3NE	Não estudou Estatística no ensino médio.
Q3EM	Estudou Estatística no ensino médio.
Q4S	Estudou Estatística em outro curso, fora do ensino regular.
Q4N	Não estudou Estatística em outro curso fora do ensino regular.
Q51	Nenhum conhecimento em Estatística.
Q52	Baixo nível de conhecimento em Estatística.
Q53	Bom nível de conhecimento em Estatística.
Q54	Domínio perfeito dos conhecimentos em Estatística.
Q61	Discorda parcialmente de que a Estatística é importante para compreender as notícias do rádio, da TV, da internet e dos jornais.
Q62	Discorda totalmente de que a Estatística é importante para compreender as notícias do rádio, da TV, da internet e dos jornais.
Q63	Concorda parcialmente que a Estatística é importante para compreender as notícias do rádio, da TV, da internet e dos jornais.
Q64	Concorda totalmente que a Estatística é importante para compreender as notícias do rádio, da TV, da internet e dos jornais.
Q71	Discorda parcialmente de que a Estatística é importante para tomar decisões em seu dia a dia.
Q72	Discorda totalmente de que a Estatística é importante para tomar decisões em seu dia a dia.
Q73	Concorda parcialmente que a Estatística é importante para tomar decisões em seu dia a dia.
Q74	Concorda totalmente que a Estatística é importante para tomar decisões em seu dia a dia.
Q81	Não acredita totalmente que irá utilizar a Estatística no(s) curso(s) superior(es) que pretende fazer.
Q82	Não acredita parcialmente que irá utilizar a Estatística no(s) curso(s) superior(es) que pretende fazer.
Q83	Acredita parcialmente que irá utilizar a Estatística no(s) curso(s) superior(es) que pretende fazer.
Q84	Acredita totalmente que irá utilizar a Estatística no(s) curso(s) superior(es) que pretende fazer.
Q91	Não acredita totalmente que a Estatística é uma área da Matemática, portanto, uma ciência exata.
Q92	Não acredita parcialmente que a Estatística é uma área da Matemática, portanto, uma ciência exata.
Q93	Acredita parcialmente que a Estatística é uma área da Matemática, portanto, uma ciência exata.
Q94	Acredita totalmente que a Estatística é uma área da Matemática, portanto, uma ciência exata.
Q101	Não confia totalmente em pesquisas estatísticas.
Q102	Não confia parcialmente em pesquisas estatísticas.
Q103	Confia parcialmente em pesquisas estatísticas.
Q104	Confia totalmente em pesquisas estatísticas.

Q111	Não sabe como são realizadas as pesquisas estatísticas.
Q112	Sabe pouco sobre como são realizadas as pesquisas estatísticas.
Q113	Sabe bem como são realizadas as pesquisas estatísticas.
Q114	Sabe muito bem como são realizadas as pesquisas estatísticas.
Q12 a Q24	Essas questões versam sobre Probabilidade. No decorrer da pesquisa, optamos por desconsiderá-las e manter o foco na Estatística.
Q25a	Assinalou a alternativa a da questão 25 (alternativa errada).
Q25b	Assinalou a alternativa b da questão 25 (alternativa errada).
Q25c	Assinalou a alternativa c da questão 25 (alternativa errada).
Q25d	Assinalou a alternativa d da questão 25 (alternativa correta).
Q25e	Assinalou a alternativa e da questão 25 (alternativa errada).
Q26a	Assinalou a alternativa a da questão 26 (alternativa errada).
Q26b	Assinalou a alternativa b da questão 26 (alternativa errada).
Q26c	Assinalou a alternativa c da questão 26 (alternativa errada).
Q26d	Assinalou a alternativa d da questão 26 (alternativa errada).
Q26e	Assinalou a alternativa e da questão 26 (alternativa correta).
Q27a	Assinalou a alternativa a da questão 27 (alternativa errada).
Q27b	Assinalou a alternativa b da questão 27 (alternativa correta).
Q27c	Assinalou a alternativa c da questão 27 (alternativa errada).
Q27d	Assinalou a alternativa d da questão 27 (alternativa errada).
Q27e	Assinalou a alternativa e da questão 27 (alternativa errada).
Q28a	Assinalou a alternativa a da questão 28 (alternativa errada).
Q28b	Assinalou a alternativa b da questão 28 (alternativa correta).
Q28c	Assinalou a alternativa c da questão 28 (alternativa errada).
Q28d	Assinalou a alternativa d da questão 28 (alternativa errada).
Q28e	Assinalou a alternativa e da questão 28 (alternativa errada).
Q29a	Assinalou a alternativa a da questão 29 (alternativa correta).
Q29b	Assinalou a alternativa b da questão 29 (alternativa errada).
Q29c	Assinalou a alternativa c da questão 29 (alternativa errada).
Q29d	Assinalou a alternativa d da questão 29 (alternativa errada).
Q29e	Assinalou a alternativa e da questão 29 (alternativa errada).

Fonte: o autor, 2020.

Tendo como base os modelos de letramento estatístico de Gal (2002), a questão Q1 teve o intuito de identificar conhecimentos e concepções dos alunos sobre a Estatística — seu objetivo, seu método e sua relação com as demais ciências —; avaliou habilidades de letramento referentes ao domínio da língua natural, conhecimentos estatísticos e questionamento crítico acerca dos elementos de conhecimento de Gal (ibid.); e, quanto aos elementos disposicionais, permitiu que examinássemos crenças, atitudes e postura crítica dos estudantes.

As questões Q2, Q3, Q4 e Q5 avaliaram o reconhecimento do aluno sobre sua própria aprendizagem em Estatística. As questões de Q6 a Q9 consideraram o conhecimento de contexto e questionamento crítico bem como suas crenças. E as questões Q10 e Q11 abordaram crenças e atitudes dos estudantes a respeito da produção do conhecimento científico em Estatística, das etapas do ciclo investigativo e da confiabilidade das investigações estatísticas.

Na segunda parte, a ênfase foi no letramento probabilístico. Por isso, as questões Q12 a Q15 consideraram o reconhecimento que o aluno possuía sobre sua própria aprendizagem em Probabilidade. As questões seguintes, de Q16 a Q24, investigaram os conhecimentos sobre ideias básicas em Probabilidade — tais quais variação, aleatoriedade, independência, previsibilidade e incerteza — bem como o questionamento crítico. Em termos de elementos disposicionais, destacamos crenças, atitudes e sentimentos quanto à incerteza e risco. Porém, como já apontamos, após a aplicação do questionário, escolhemos não as considerar em nossa análise e focar nos conhecimentos sobre a Estatística.

Por último, as perguntas da terceira — Q25 a Q29 — pretenderam indicar conhecimentos matemáticas e estatísticos, com ênfase na leitura e interpretação de tabelas e gráficos estatísticos. No entanto, o nível das habilidades de letramento exigidas foi o básico. Vamos dedicar uma atenção maior a estas cinco questões, pois, por meio delas, podemos inferir possíveis concepções mobilizadas pelos alunos em nossa análise prévia.

A questão Q25 requer a leitura e interpretação de uma tabela.

Figura 13 – Questão 25 do questionário aplicado

25 - (ENEM/2012) A tabela a seguir mostra a evolução da receita bruta anual nos três últimos anos de cinco microempresas (ME) que se encontram à venda.

<i>ME</i>	<i>2009 (em milhares de reais)</i>	<i>2010 (em milhares de reais)</i>	<i>2011 (em milhares de reais)</i>
<i>Alfinetes V</i>	<i>200</i>	<i>220</i>	<i>240</i>
<i>Balas W</i>	<i>200</i>	<i>230</i>	<i>200</i>
<i>Chocolates X</i>	<i>250</i>	<i>210</i>	<i>215</i>
<i>Pizzaria Y</i>	<i>230</i>	<i>230</i>	<i>230</i>
<i>Tecelagem Z</i>	<i>160</i>	<i>210</i>	<i>245</i>

Fonte: Dados fictícios

Um investidor deseja comprar duas das empresas listadas na tabela. Para tal, ele calcula a média da receita bruta anual dos últimos três anos (de 2009 até 2011) e escolhe as duas empresas de maior média anual. As empresas que este investidor escolhe comprar são:

- a) Balas W e Pizzaria Y.
- b) Chocolates X e Tecelagem Z.
- c) Pizzaria Y e Alfinetes V.
- d) Pizzaria Y e Chocolates X.**
- e) Tecelagem Z e Alfinetes V.

Fonte: o autor, 2020.

Pallauta, Gea e Venegas (2019) destacam que as tabelas não recebem a devida atenção nos currículos e no ensino e acrescentam que essas tabelas estão praticamente ausentes nos livros didáticos de Matemática dos dois países. Entretanto, estão em livros de outras disciplinas, como Geografia, o que pode justificar o fato de 57% dos alunos terem acertado a questão ao assinalar a alternativa "d".

Por outro lado, a opção errada com o maior número de escolhas — 19% dos alunos — foi a alternativa "e", a qual aponta para as duas empresas cujas receitas apresentaram a maior amplitude total. Esse argumento pode ser reforçado pela redação do enunciado da questão, a qual menciona a “evolução da receita bruta anual”, o que poderia levar o estudante a pensar na maior variação entre os limites inferior e superior encontrados, respectivamente, nas empresas Z e V.

Pode existir, por detrás dessa resposta, uma concepção $\mathbb{C}_a$, caracterizada pelos elementos:

- P – campo de problemas: determinação da média aritmética simples com base nos valores apresentados em uma tabela de dupla entrada.

- R – operadores: localização dos valores das receitas brutas anuais associadas às respectivas empresas de origem em uma tabela e verificação da diferença entre eles.
- L – conjunto de representações: representação numérica e representação tabular.
- Σ – estrutura de controle: média aritmética simples como valor diretamente proporcional à variação entre os limites inferior e superior da amostra (amplitude total).

Assim, podemos inferir a existência de uma possível concepção Φ_a , deslocada de seu domínio de validade, segundo a qual o valor da média será maior, quanto maior for a diferença entre os dois extremos na distribuição de frequência: o limite inferior e o limite superior da amostra.

Vejamos, agora, a questão 26, em que o aluno, novamente, precisa determinar a média aritmética simples (X , segundo o problema) com base nos dados presentes em uma tabela. No entanto, ele também precisa determinar os valores da mediana (Y) e da moda (Z). Em relação à questão anterior, o percentual de acertos caiu de 57% para 27%.

Figura 14 – Questão 26 do questionário aplicado

26 - (ENEM/2010) O quadro seguinte mostra o desempenho de um time de futebol no último campeonato. A coluna da esquerda mostra o número de gols marcados e a coluna da direita informa em quantos jogos o time marcou aquele número de gols.

Gols marcados	Quantidade de partidas
0	5
1	3
2	4
3	3
4	2
5	2
7	1

Fonte: Dados fictícios

Se X , Y e Z são, respectivamente, a média, a mediana e a moda desta distribuição, então:

- $X = Y < Z$.
- $Z < X = Y$.
- $Y < Z < X$.
- $Z < X < Y$.
- $Z < Y < X$.

Fonte: o autor, 2020.

Ao contrário da questão anterior, que concentrou 76% das respostas em duas alternativas — a correta e a do erro analisado —, o resultado desta foi bastante heterogêneo: o erro mais frequente (alternativa "a") foi escolhido por 30% dos alunos, ao passo que o acerto (alternativa "e") foi assinalado por 27% dos respondentes, e as alternativas "b", "c" e "d" foram apontadas, respectivamente, por 12%, 14% e 17% deles.

Segundo Batanero (2001), a moda é, entre as medidas de tendência central, aquela com o custo cognitivo muito baixo. Identificando a moda — no caso deste problema, 0 gols marcados —, o respondente ficaria entre as alternativas "b", "d", "e".

Podemos inferir duas hipóteses sobre a confusão entre média (X) e mediana (Y): os alunos podem ter confundido o valor da mediana — no caso, 2 — com o valor da frequência da classe intermediária, a 4ª classe — no caso, 3; ou os estudantes trocaram os algoritmos da média aritmética simples e da mediana.

Essa confusão entre média e mediana ainda aparecerá nas respostas dos alunos — tanto na escrita quanto no discurso —, coletadas na segunda fase da pesquisa, apresentada mais à frente. Cobo (2003) considera que a determinação da mediana para o estudante é mais complexa do que parece ser, pois seu cálculo é diferente quando há um número par ou ímpar de dados ou quando os dados estão apresentados em tabelas de valores agrupados com intervalos de classe ou não agrupados.

Uma possível concepção mobilizada por quem errou esta questão relacionada à compreensão da mediana em uma tabela — no caso, 73% dos alunos — pode ser representada pelos elementos a seguir:

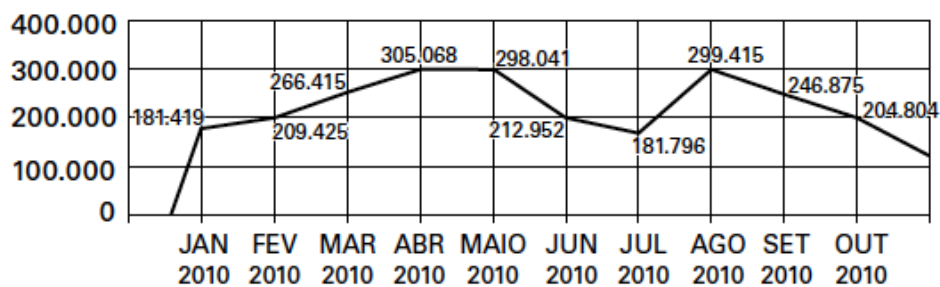
- P – campo de problemas: determinação da mediana e comparação com os valores de média aritmética simples e moda.
- R – operadores: localização da classe de maior frequência na tabela e comparação, em ordem crescente, com os valores atribuídos à moda e à média aritmética simples.
- L – conjunto de representações: representação numérica e representação tabular.
- Σ – estrutura de controle: verificação da frequência da classe intermediária, em um número ímpar de classes, compreendida pelo estudante que mobilizou essa concepção como mediana.

Sendo assim, podemos inferir a existência de uma possível concepção Φ_b , segundo a qual o valor da mediana em uma tabela corresponde à frequência da classe intermediária, em vez de corresponder ao valor intermediário — considerando um número ímpar de elementos — ou à média dos dois valores intermediários — considerando um número par de elementos. Neste caso, como temos sete classes, a mediana seria a frequência correspondente à quarta classe, sendo, portanto, maior do que a média e a moda.

Na questão seguinte, Q27, o diferencial em relação às duas anteriores é o fato de que, agora, os valores estão representados em um gráfico de linha. Outro detalhe que nos chama a atenção é a distribuição das respostas: 24%, 21%, 21%, 14% e 20% para as alternativas "a", "b", "c", "d" e "e", respectivamente, isto é, todas muito próximas dos 20% esperados para uma escolha feita ao acaso. A alternativa correta, "b", teve 21% das escolhas.

Figura 15 – Questão 27 do questionário aplicado

27 - (ENEM/2012) O gráfico apresenta o comportamento de emprego formal surgido, segundo o CAGED, no período de janeiro de 2010 a outubro de 2010.



Disponível em: www.mte.gov.br. Acesso em: 28 fev. 2012 (adaptado).

Pelo gráfico, o valor da parte inteira da mediana dos empregos formais surgidos no período é:

- a) 212.952.
- b) 229.913.**
- c) 240.621.
- d) 255.496.
- e) 298.041.

Fonte: o autor, 2020.

O entendimento exigido nesta questão, tomando por base a classificação de Curcio (1989), ampliada por Friel, Curcio e Bright (2001), é o de nível 3 da leitura dos dados, pois, além de exigir a comparação de quantidades bem como o uso de outros conceitos e habilidades matemáticas, combinando e integrando dados presentes no

gráfico, requer informações que não são apresentadas diretamente no gráfico, extraindo novos esquemas e padrões para, nesse caso, a determinação da mediana. Exercícios como esse, não são comuns nos livros didáticos brasileiros, como observou Giordano (2016).

Os livros didáticos recomendados nos últimos anos pelo Programa Nacional do Livro Didático – PNLD possuem exercícios pouco favoráveis ao desenvolvimento do letramento estatístico dos alunos. Coutinho (2013) e Amari Goulart (2015) asseveram que os livros didáticos destinados ao ensino médio têm apresentado os tópicos tradicionais da Estatística, como gráficos, tabelas e medidas de posição, porém não contam com propostas de atividades que estimulem uma análise crítica dos resultados, em contextos que despertem o interesse e a motivação dos estudantes.

Dessa forma, predominam em tais livros os exercícios que valorizam a aplicação pura e simples de um algoritmo, sem explicitar os conceitos envolvidos. Considerando que os alunos tiveram pouco ou nenhum contato com a Estatística nas aulas de Matemática, conforme mostrou o questionário da primeira etapa de nossa coleta de dados, o resultado da questão Q27 é compreensível.

Uma possível concepção mobilizada, dentro de seu domínio de validade, por quem acertou essa questão relacionada à compreensão da mediana em gráfico de linha — neste caso, apenas 21% dos alunos — pode ser representada pelos seguintes elementos:

- P – campo de problemas: determinação da mediana a partir de uma dezena de valores relativos às vagas de emprego criadas no período referido, apresentadas em um gráfico de linha.
- R – operadores: localização de valores relativos às vagas de emprego criadas no período referido, identificação dos dois valores intermediários e cálculo da média aritmética simples entre os dois, no caso, o quinto e sexto maiores valores.
- L – conjunto de representações: representação numérica e representação gráfica.
- Σ – estrutura de controle: mediana como média aritmética simples entre os dois valores intermediários da distribuição de frequência após a organização dos dados em ordem crescente ou decrescente.

Portanto, podemos inferir a existência de uma possível concepção Φ_c , segundo a qual a mediana, estabelecida com base em um gráfico de linha que apresenta um número par de elementos (10), corresponde à média aritmética simples entre os dois intermediários (quinto e sexto maiores valores), quando são organizados em rol, em ordem crescente ou decrescente.

Passemos, agora, à questão Q28, em que, tal qual verificamos nas questões Q25 E Q26, os dados necessários para a resolução estão dispostos em uma tabela. No entanto, como é menos comum nas provas do ENEM, é exigido o cálculo de uma medida de dispersão, não uma de tendência central.

Figura 16 – Questão 28 do questionário aplicado

28 - (ENEM/2010) Marco e Paulo foram classificados em um concurso. Para classificação no concurso, o candidato deveria obter média aritmética na pontuação igual ou superior a 14. Em caso de empate na média, o desempate seria em favor da pontuação mais regular. No quadro a seguir são apresentados os pontos obtidos nas provas de Matemática, Português e Conhecimentos Gerais, a média, a mediana e o desvio padrão dos dois candidatos.

Dados dos candidatos no concurso

	Matemática	Português	Conhecimentos Gerais	Média	Mediana	Desvio Padrão
Marco	14	15	16	15	15	0,32
Paulo	8	19	18	15	18	4,97

Fonte: Dados fictícios

O candidato com pontuação mais regular, portanto mais bem classificado no concurso, é

a) Marco, pois a média e a mediana são iguais.

b) Marco, pois obteve menor desvio padrão.

c) Paulo, pois obteve a maior pontuação da tabela, 19 em Português.

d) Paulo, pois obteve maior mediana.

e) Paulo, pois obteve maior desvio padrão.

Fonte: o autor, 2020.

Em grande parte da literatura científica sobre Educação Estatística, da qual podemos destacar Batanero (2001), Batanero e Godino (2001), Cobo (2003), Galicia (2009) e Andrade (2013), recomenda-se a análise de medidas de tendência central associadas a medidas de dispersão, como média e desvio padrão ou mediana e desvio médio.

Além disso, apesar de o enunciado ser relativamente simples, apresenta um contexto bastante familiar ao aluno, envolvendo uma situação de avaliação e notas em diferentes disciplinas. Embora requeira um nível de compreensão conceitual mais

profundo do que a primeira ou segunda questões aqui analisadas, esta obteve um nível relativamente bom de acertos: 44% dos respondentes assinalaram a alternativa "b". Podemos inferir, então, que isso também ocorreu porque os valores de média, mediana e desvio padrão foram dados no enunciado, não exigindo cálculo. A alternativa incorreta mais assinalada foi a "d", erro frequente, denotando dificuldade do aluno em associar homogeneidade/heterogeneidade/regularidade ao conceito de dispersão.

A questão é mais conceitual do que procedimental e requer compreensão de um conceito onipresente na Estatística: o de variabilidade. Segundo Wild e Pfannkuch (1999), esse é um conceito central a toda e qualquer compreensão da Estatística bem como de suas aplicações. Garfield e Gal (1999) consideram a variabilidade fundamental, em toda e qualquer investigação estatística, para compreender a relação entre as medidas de posição e variabilidade de um conjunto de dados.

Uma possível concepção mobilizada por quem acertou essa questão relacionada à compreensão do desvio padrão em dados organizados em tabela — neste caso, apenas 21% dos alunos — pode ser representada pelos seguintes elementos:

- P – campo de problemas: determinação da maior regularidade de um conjunto de dados, considerando sua homogeneidade/heterogeneidade a partir de duas medidas de tendência central — média aritmética simples e mediana — e uma medida de dispersão — desvio padrão.
- R – operadores: identificação das medidas de posição em uma tabela.
- L – conjunto de representações: representação numérica e tabular.
- Σ – estrutura de controle: reconhecimento do desvio padrão como indicador do nível de regularidade/variabilidade de um conjunto de dados.

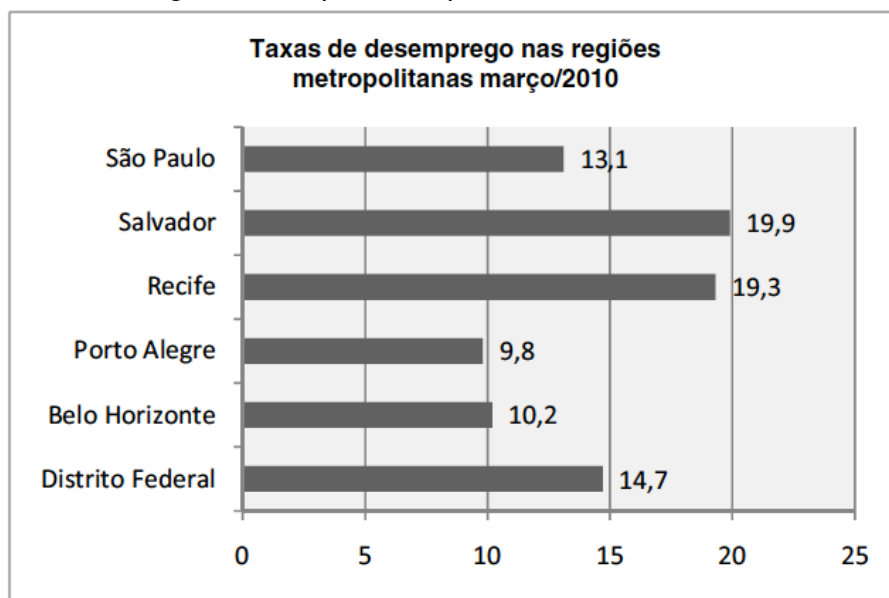
Assim, podemos inferir a existência de uma possível concepção Φ_d , segundo a qual o desvio padrão é um indicador inversamente proporcional à regularidade de um conjunto de dados: quanto menor o desvio padrão, maior a regularidade, e vice-versa.

Por fim, observemos a questão Q29, a qual acusou o segundo maior número de respostas corretas nessa atividade — 50% dos respondentes. Embora o enunciado não informe que a referida taxa de desemprego foi apresentada em termos

percentuais, os alunos, aparentemente, concluíram isso, possivelmente pela facilidade de encontrar, nas mais diferentes mídias, valores estatísticos representados por porcentagens.

Figura 17 – Questão 29 do questionário aplicado

29 - (ENEM/2010) Os dados do gráfico seguinte foram gerados a partir de dados colhidos no conjunto de seis regiões metropolitanas pelo Dieese.



Disponível em: <http://g1.globo.com>. Acesso em: 28 abr. 2010 (adaptado).

Supondo que o total de pessoas pesquisadas na região metropolitana de Porto Alegre equivale a 250 000, o número de desempregados em março de 2010, nessa região, foi:

- a) 24 500.
- b) 25 000.
- c) 220 500.
- d) 223 000.
- e) 227 500.

Fonte: o autor, 2020.

O enunciado da questão Q29 destaca a cidade de Porto Alegre, direcionando a resposta do aluno. Então, basta que ele localize a taxa de desempregados e, supondo que ela esteja representada em termos percentuais, efetue o cálculo $9,8 / 100 \times 250.000$. A questão requer quase que apenas uma apreensão literal do gráfico, mas com alguma comparação e integração muito básica. Ela corresponde ao nível de leitura 2 — apresentado por Curcio (1989) —, com um custo baixo cognitivo.

Uma possível concepção Φ_e , mobilizada dentro do domínio de validade por quem acertou essa questão — ou seja, metade dos alunos —, pode ser representada pelos seguintes elementos:

- P – obtenção de um dado de frequência absoluta, localizado em um gráfico de barras horizontais (o número total de trabalhadores desempregados daquela cidade) com base em um dado relativo (a taxa de desemprego), supostamente representada em termos percentuais considerando o número total de pessoas pesquisadas daquela cidade.
- R – operadores: localização do valor de uma dada frequência absoluta e, com base nela, determinação do valor correspondente à frequência relativa considerando o total de elementos da amostra.
- L – conjunto de representações: representação numérica e representação gráfica.
- Σ – estrutura de controle: comparação parte-todo.

Dessa maneira, podemos inferir a existência de uma possível concepção Φ_e , segundo a qual basta empregar uma técnica simples, como a regra de três, para obter um valor absoluto correspondente a um valor relativo, expresso em termos de porcentagem.

Ressaltamos que voltaremos a tratar de concepções no capítulo 6 deste estudo, já considerando respostas individuais de alunos.

Ainda em relação a nossa análise, nossos dados foram coletados em junho de 2018, poucas semanas antes do início do terceiro bimestre, momento em que os alunos têm seu primeiro contato com a Estatística no ensino médio nas aulas de Matemática. Alguns deles, eventualmente, travaram contato com a Estatística em outras disciplinas durante o ensino fundamental — como Ciências, Educação Física e Geografia — ou mesmo no ensino médio — como Biologia, Física e Sociologia.

Não podemos desconsiderar seu contato com a Estatística por meio das mais diversas mídias. No ano de realização da nossa coleta, houve eleições majoritárias no Brasil e as primeiras pesquisas eleitorais já eram, regularmente, apresentadas em jornais, rádio, TV e *internet*.

A primeira pergunta, e única questão aberta do nosso questionário, não foi respondida por cerca de 16,3% dos alunos. As demais foram respondidas pelos 86 estudantes. Para nossa variável complementar, sexo, obtivemos como resposta 51,2% de homens e 48,8% de mulheres. Suas idades variam de 16 a 19 anos, prevalecendo

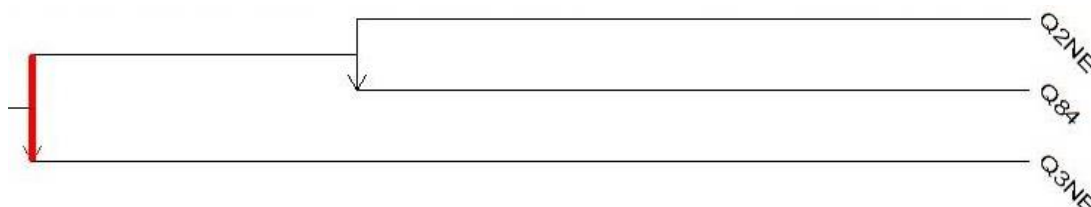
17 anos (82,6%), idade esperada para um aluno concluinte do ensino médio que nunca tenha sido reprovado em seu percurso escolar.

5.1 Análise Coesitiva

Feitas essas observações, trataremos, especificamente, dos dados obtidos por meio do *software* CHIC, começando pela análise coesitiva das várias metas-regras que associam as variáveis em jogo, apresentadas na figura 18 a seguir.

Iniciamos a análise coesitiva pelas classes assinaladas na figura 18, aquela que tem o maior número de nós significativos — assinalados em vermelho, nos termos do *software* —, e pelo conjunto de variáveis que ela envolve. Destacamos tais classes para explicitar as metas-regras identificadas, lembrando que, nessa análise, tais metas-regras têm a forma “se A ocorre, então provavelmente B ocorre”.

Figura 19 – Trecho da árvore coesitiva I



Fonte: o autor, 2020.

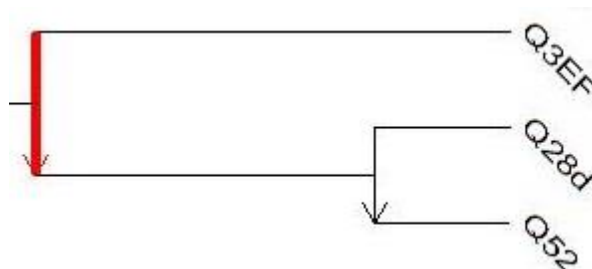
Apontamos as seguintes legendas para as siglas da figura 19:

- Q2NE – Não estudou Estatística no ensino fundamental.
- Q84 – Acredita totalmente que irá utilizar a Estatística no(s) curso(s) superior(es) que pretende fazer.
- Q3NE – Não estudou Estatística no ensino médio.

Reafirmamos que nossos apontamentos nesta seção, na perspectiva da ASI, assumem um caráter determinista: se o aluno afirma A, então provavelmente acontece B.

Essas relações evidenciam o cenário do ensino de Estatística na rede estadual paulista. Os alunos que acreditavam que iriam utilizar os conhecimentos estatísticos no ensino superior alegaram não ter estudado essa ciência no ensino fundamental, tampouco no ensino médio, até o momento de nossa coleta de dados, como já havíamos observado.

Figura 20 – Trecho da árvore coesitiva II



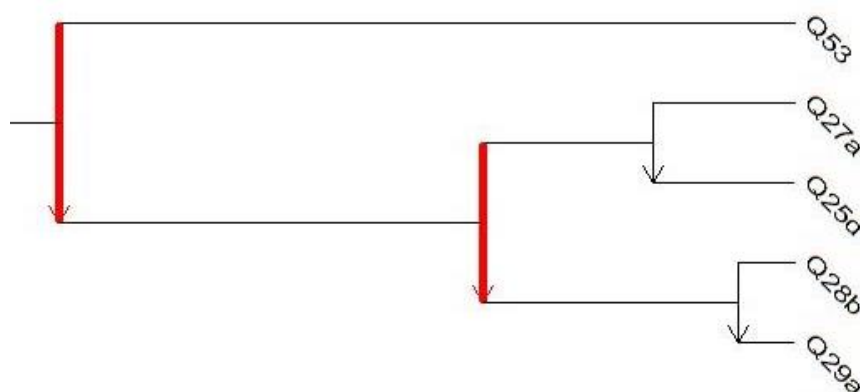
Fonte: o autor, 2020.

Apontamos as seguintes legendas para as siglas da figura 20:

- Q3EF – Estudou Estatística no ensino fundamental.
- Q28D – Assinalou a alternativa d da questão 28 (alternativa errada).
- Q52 – Baixo nível de conhecimento em Estatística.

Os alunos que alegaram ter estudado Estatística no ensino fundamental e admitiram possuir baixo nível de conhecimento na área erraram a questão 28, que envolvia o cálculo de desvio padrão, cujo estudo não era previsto no Currículo paulista de Matemática (SÃO PAULO, 2012).

Figura 21 – Trecho da árvore coesitiva III



Fonte: o autor, 2020.

Apontamos as seguintes legendas para as siglas da figura 21:

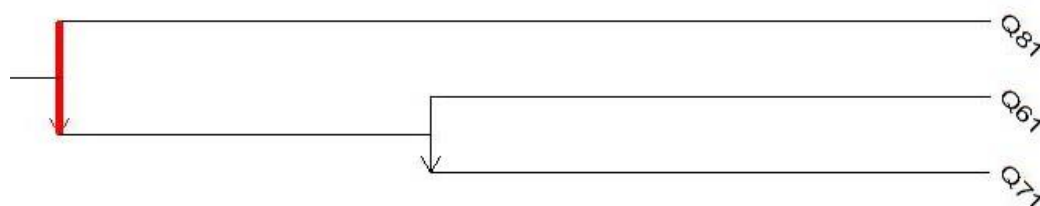
- Q53 – Bom nível de conhecimento em Estatística.
- Q27a – Assinalou a alternativa a da questão 27 (alternativa errada).
- Q25d – Assinalou a alternativa d da questão 25 (alternativa correta).
- Q28b – Assinalou a alternativa b da questão 28 (alternativa correta).
- Q29a – Assinalou a alternativa a da questão 29 (alternativa correta).

O bom nível de conhecimento em Estatística, declarado por alguns alunos, está associado por metas-regras às questões Q25, Q27, Q28 e Q29. As variáveis em jogo nessa classe são: Q27a – opção "a" para a questão 27, tratando de um problema que envolve mediana — a resposta certa seria "b"; Q25d – opção "d", correta, para a questão 25, que envolve o cálculo de uma média; Q28b – resposta "b", correta, para a questão 28, que envolve a análise do desvio padrão; e Q29a – resposta "a", correta,

para a questão 29, que trata da taxa de desemprego de regiões brasileiras dadas em um gráfico de barras.

Observamos aqui que apenas uma questão não teve a opção correta assinalada, e a leitura da árvore coesitiva permitiu afirmar que, se o aluno errou a questão 27 ao responder o item "a", provavelmente, acertou a questão 25; se acertou a questão 28, provavelmente, acertou a questão 29; e, por fim, se teve o primeiro comportamento assinalado, provavelmente, apresentou o segundo comportamento assinalado. Assim, aventamos a hipótese de as relações entre os conhecimentos — presença ou ausência — sobre média e desvio padrão serem articuladas, bem como os conhecimentos práticos sobre o cálculo de uma taxa.

Figura 22 – Trecho da árvore coesitiva IV



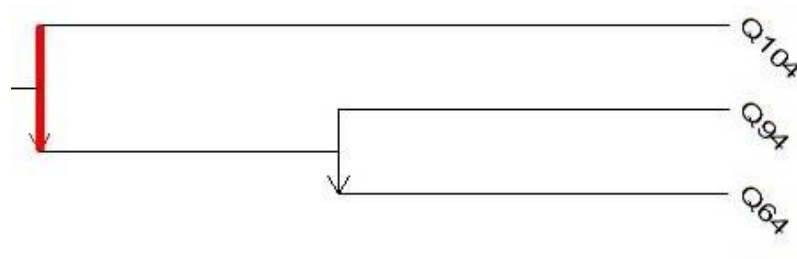
Fonte: o autor, 2020.

Apontamos as seguintes legendas para as siglas da figura 22:

- Q81 – Não acredita totalmente que irá utilizar a Estatística no(s) curso(s) superior(es) que pretende fazer.
- Q61 – Discorda parcialmente de que a Estatística é importante para compreender as notícias do rádio, da TV, da internet e dos jornais.
- Q71 – Discorda parcialmente de que a Estatística é importante para tomar decisões em seu dia a dia.

Os alunos que não estavam certos se utilizariam a Estatística nos cursos de sua preferência na universidade provavelmente manifestaram incerteza quanto à sua utilidade para a compreensão das notícias veiculadas em diferentes fontes e a seu peso na tomada de decisões em seu dia a dia.

Figura 23 – Trecho da árvore coesitiva V



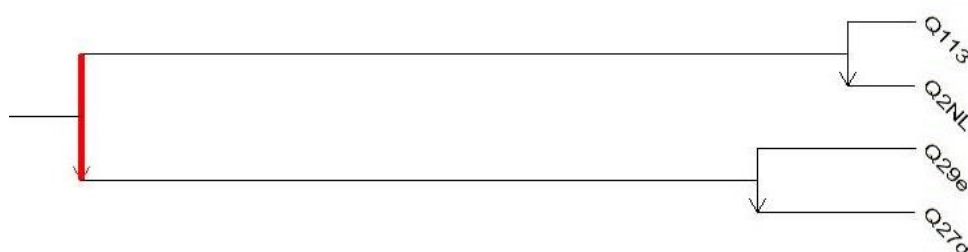
Fonte: o autor, 2020.

Apontamos as seguintes legendas para as siglas da figura 23:

- Q104 – Confia totalmente em pesquisas estatísticas.
- Q94 – Acredita totalmente que a Estatística é uma área da Matemática, portanto, uma ciência exata.
- Q64 – Concorda totalmente que a Estatística é importante para compreender as notícias do rádio, da TV, da internet e dos jornais.

Os alunos que afirmaram confiar em pesquisas estatísticas provavelmente manifestaram plena confiança em sua utilidade para a compreensão das notícias veiculadas em diferentes fontes, julgando-a uma ciência exata, um campo da Matemática. É possível que tenham associado à Estatística o rigor e a precisão geralmente atribuídos à Matemática, considerando-a, por extensão, tão confiável quanto a última.

Figura 24 – Trecho da árvore coesitiva VI



Fonte: o autor, 2020.

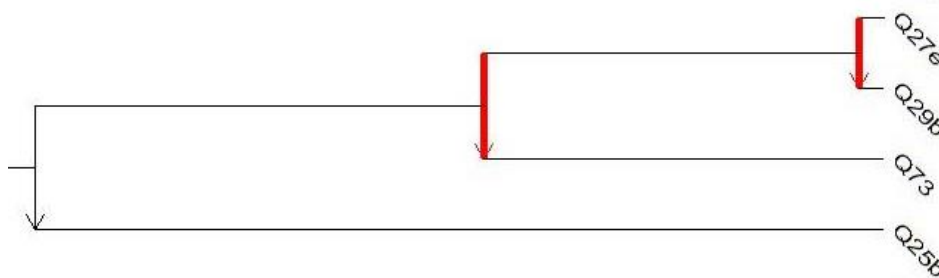
Apontamos as seguintes legendas para as siglas da figura 24:

- Q113 – Sabe bem como são realizadas as pesquisas estatísticas.
- Q2NL – Não se lembra se estudou Estatística no ensino fundamental.
- Q29e – Assinalou a alternativa e da questão 29 (alternativa errada).

- Q27d – Assinalou a alternativa d da questão 27 (alternativa errada).

Os alunos alegaram saber bem como são realizadas as pesquisas estatísticas, no entanto, não recordaram ter estudado Estatística no ensino fundamental e, pelos dados que obtivemos em nossa pesquisa, é pouco provável que a tenham estudado. Assim, erraram a questão 27 (que obteve o menor índice de acerto, de apenas 24% dos entrevistados) e, provavelmente, erraram a questão 29 (que teve o maior índice de acerto, 50%) assinalando o item menos escolhido pelos entrevistados (por apenas 9% deles).

Figura 25 – Trecho da árvore coesitiva VII



Fonte: o autor, 2020.

Apontamos as seguintes legendas para as siglas da figura 25:

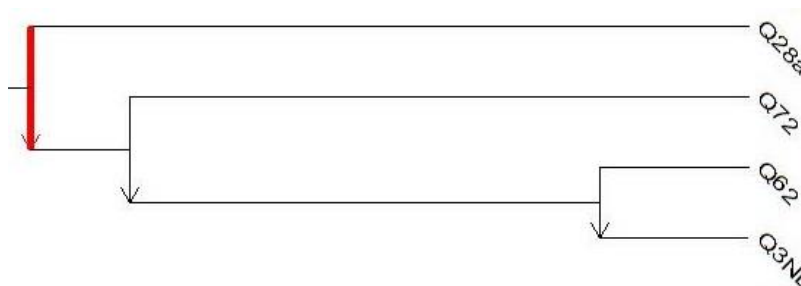
- Q27e – Assinalou a alternativa e da questão 27 (alternativa errada).
- Q29b – Assinalou a alternativa b da questão 29 (alternativa errada).
- Q73 – Concorda parcialmente que a Estatística é importante para tomar decisões em seu dia a dia.
- Q25b – Assinalou a alternativa b da questão 25 (alternativa errada).

Temos, aqui, as seguintes características: Q25b – opção "b", envolvendo média, como resposta ao problema proposto sobre compra de produtos, com base na análise de uma tabela, cuja alternativa correta seria a "d"; Q27e – opção "e", envolvendo mediana, como resposta ao problema proposto sobre comportamento do emprego formal no Brasil, cuja alternativa correta seria a "b"; Q29b – opção "b", envolvendo desvio padrão, alternativa correta, como resposta ao problema proposto sobre pontuação de candidatos em um concurso; e Q73, em concordância com o afirmado no item 7 de que “a Estatística é importante para tomar decisões no dia a dia”.

Das quatro variáveis envolvidas nessa classe, três exigiam conhecimentos estatísticos sobre medidas-resumo — a saber, média, mediana e desvio padrão —, ao passo que a outra exigia do aluno um posicionamento frente à assertiva sobre a aplicabilidade da Estatística. Na figura 25, vemos que: se o respondente errou a questão 27, provavelmente, errou a 29; se esse comportamento foi observado, provavelmente concordou que a Estatística é importante nas decisões de seu dia a dia; e, finalmente, se errou a questão 25, provavelmente apresentou o comportamento descrito.

Constatamos pouco conhecimento estatístico relativo ao estudo da variação de dados, tão necessárias para o pleno exercício da cidadania na tomada de decisões cotidianas.

Figura 26 – Trecho da árvore coesitiva VIII



Fonte: o autor, 2020.

Apontamos as seguintes legendas para as siglas da figura 26:

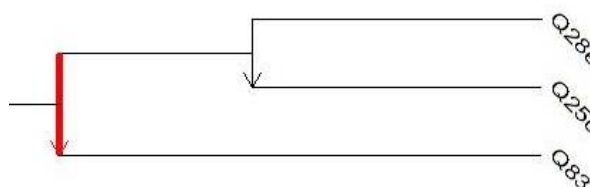
- Q28a – Assinalou a alternativa a da questão 28 (alternativa errada).
- Q72 – Discorda totalmente de que a Estatística é importante para tomar decisões em seu dia a dia.
- Q62 – Discorda totalmente de que a Estatística é importante para compreender as notícias do rádio, da TV, da internet e dos jornais.
- Q3NL – Não se lembra se estudou Estatística no ensino médio.

Se os alunos erraram a questão Q28 — que obteve o segundo maior índice de acertos (44% dos entrevistados) —, a qual tratou das medidas-resumo média, mediana e desvio padrão, provavelmente não se lembravam de ter estudado Estatística no ensino médio — a maioria não havia estudado na disciplina de Matemática até então, por motivos já explicados — e, possivelmente, não

consideravam essa ciência importante para tomar decisões em seu dia a dia, tampouco para compreender as notícias do rádio, da TV, da internet e dos jornais.

Por fim, as últimas relações que observaremos do grafo coesitivo estão representadas na figura 27.

Figura 27 – Trecho da árvore coesitiva IX



Fonte: o autor, 2020.

Apontamos as seguintes legendas para as siglas da figura 27:

- Q28e – Assinalou a alternativa e da questão 28 (alternativa errada).
- Q25c – Assinalou a alternativa c da questão 25 (alternativa errada).
- Q83 – Acredita parcialmente que irá utilizar a Estatística no(s) curso(s) superior(es) que pretende fazer.

Os alunos que erraram as questões Q28 e Q25, — que envolviam resumo média aritmética —, a despeito de suas dificuldades, identificavam a possível importância da Estatística para seu futuro profissional.

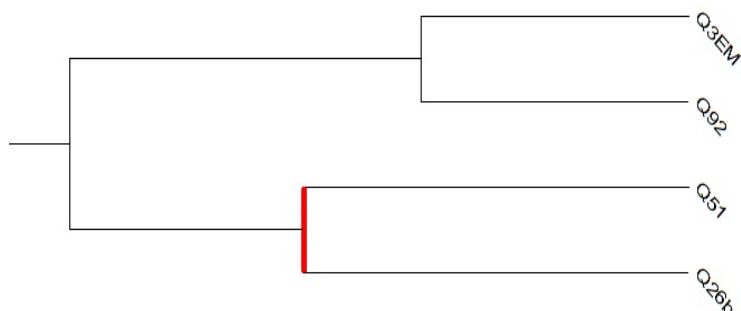
Reconhecemos que, até agora, fizemos uma série de suposições baseadas em evidências levantadas pelo grafo coesitivo. Para dar mais consistência a elas, analisaremos, a seguir, outros dois grafos gerados pelo CHIC: a árvore de similaridades e o grafo implicativo.

5.2 Análise de Similaridades

Com base na figura apresentada a seguir, buscaremos compreender as relações identificadas na árvore de similaridades.

Da figura anterior, extraímos o seguinte trecho:

Figura 29 – Árvore de similaridades: relações significativas I



Fonte: o autor, 2020.

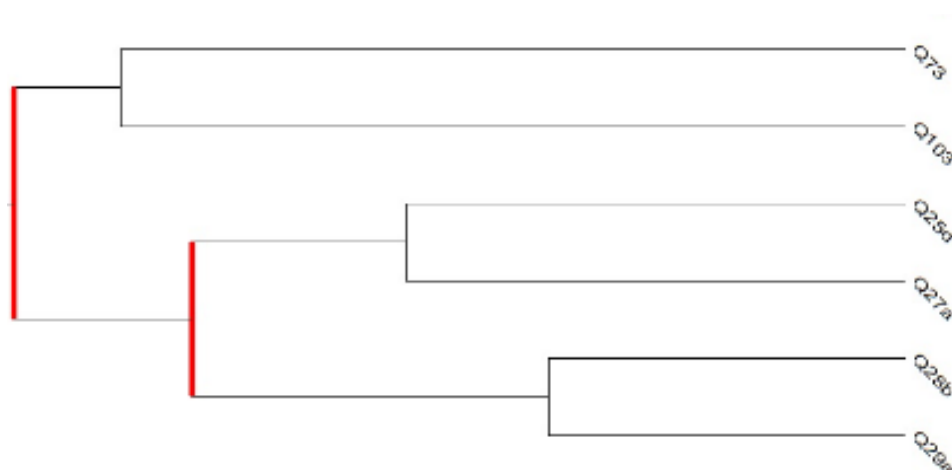
Apontamos as seguintes legendas para as siglas da figura 29:

- Q3EM – Estudou Estatística no ensino médio.
- Q92 – Não acredita parcialmente que a Estatística é uma área da Matemática, portanto, uma ciência exata.
- Q51 – Nenhum conhecimento em Estatística.
- Q26b – Assinalou a alternativa b da questão 26 (alternativa errada).

Os respondentes que afirmaram ter estudado Estatística no ensino médio (Q3EM) tiveram comportamento similar ao dos que discordaram da afirmação 9, ou seja, de que a Estatística é uma área da Matemática, sendo uma ciência exata. Podemos inferir que os alunos que possuíam algum conhecimento mínimo sobre o assunto já reconheciam a Estatística como uma ciência autônoma, e não como uma mera extensão da Matemática.

Quem afirmou não ter conhecimentos anteriores de Estatística (Q51) provavelmente tem comportamento similar ao de quem assinalou o item "b" na questão 26 — em que a opção correta era a "e". Esse reconhecimento é compreensível, pelos motivos já expostos nesta tese. Uma vez que, por questões curriculares, os alunos não estudaram esse tema, eles não possuíam domínio conceitual que os permitisse comparar a moda, a mediana e a média, ainda que o problema tratasse de um contexto familiar ao brasileiro: o futebol.

Figura 30 – Árvore de similaridades: relações significativas II



Fonte: o autor, 2020.

Apontamos as seguintes legendas para as siglas da figura 30:

- Q73 – Concorda parcialmente que a Estatística é importante para tomar decisões em seu dia a dia.
- Q103 – Confia parcialmente em pesquisas estatísticas.
- Q25d – **Assinalou a alternativa d da questão 25 (alternativa correta).**
- Q27a – Assinalou a alternativa a da questão 27 (alternativa errada).
- Q28b – **Assinalou a alternativa b da questão 28 (alternativa correta).**
- Q29a – **Assinalou a alternativa a da questão 29 (alternativa correta).**

Aqueles que concordaram que a Estatística é importante para a tomada de decisões em problemas do dia a dia (Q73) têm comportamento similar ao dos que confiaram em pesquisas estatísticas (Q103), posição intrinsecamente coerente, envolvendo os elementos disposicionais de letramento de Gal (2002): crenças e atitudes. Esse grupo manifestou comportamento similar ao daqueles que apresentaram bom aproveitamento nas questões do ENEM, com 75% de acerto no conjunto que abarca as questões Q25, Q27, Q28 e Q29.

Os alunos que erraram a questão 27 (Q27a) — que envolvia mediana em um gráfico de linha com valores relativamente elevados — acertaram outras 3 do ENEM, que envolveram média, moda, mediana e desvio padrão (Q25d, Q28b e Q29a). Como vimos na seção 5.1 desta tese, apenas 21% dos respondentes a acertaram. Entre as respostas erradas, nenhuma das opções se destacou, e a alternativa mais escolhida

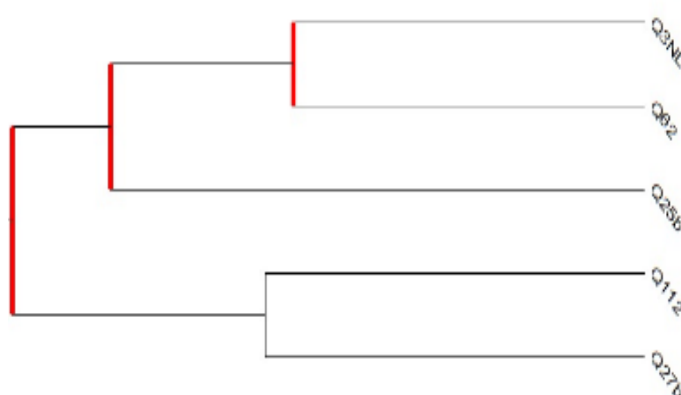
foi assinalada por apenas 24% dos alunos. Assim, podemos inferir que os estudantes fizeram uma escolha casuística.

O gráfico foi reproduzido da forma que estava no ENEM: pequeno; poluído visualmente, com números muito próximos uns dos outros; com valores elevados para os padrões do ENEM; e com uma grade desnecessária, que dificultava ainda mais a apreensão visual. Comparativamente, o gráfico de barras horizontais da questão Q29 era muito mais simples, menos poluído visualmente e com valores menores. Então, metade dos respondentes acertou essa questão.

As questões Q25, Q28 e Q29 envolveram valores menores. Tais valores são mais próximos àqueles que foram encontrados nos problemas de Matemática em livros didáticos, como observou Giordano (2016). Esses problemas podem ser melhor compreendidos pelo conhecimento prévio de seus respectivos contextos — empresas como pizzaria ou tecelagem; desemprego; e situação de avaliação escolar —, como defendido na Análise Exploratória de Dados – AED (BATANERO; ESTEPA; GODINO, 1991).

Embora esses alunos tenham afirmado possuir bons conhecimentos sobre Estatística (Q53), seu rendimento é justificável mais pelo seu conhecimento de mundo do que pelo nível de seus conhecimentos estatísticos no âmbito escolar, uma vez que careciam dessa experiência (BATANERO e DÍAZ, 2004).

Figura 31 – Árvore de similaridades: relações significativas III



Fonte: o autor, 2020.

Apontamos as seguintes legendas para as siglas da figura 31:

- Q3NL – Não se lembra se estudou Estatística no ensino médio.

- Q62 – Discorda totalmente de que a Estatística é importante para compreender as notícias do rádio, da TV, da internet e dos jornais.
- Q25b – Assinalou a alternativa b da questão 25 (alternativa errada).
- Q112 – Sabe pouco sobre como são realizadas as pesquisas estatísticas.
- Q27b – **Assinalou a alternativa b da questão 27 (alternativa correta).**

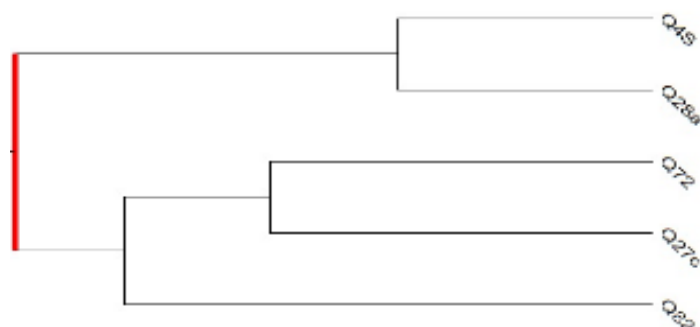
Aqueles alunos que não recordaram se já haviam estudado a Estatística no ensino médio (Q3NL) manifestaram incerteza quanto à importância dessa ciência para a compreensão das notícias veiculadas pelas mais diversas mídias (Q62), apresentando comportamento similar ao daqueles que erraram a questão Q25, embora seu enunciado seja passível de muitas críticas (ver Apêndice 1).

Tal questão demandou do aluno uma avaliação sobre a evolução de duas empresas, contudo, propôs como critério para a escolha das duas melhores apenas a média de sua receita bruta anual nos três últimos anos. A empresa Z — presente no item b — apresentou a maior evolução, embora as empresas X e Y — da alternativa correta — apresentassem médias maiores. A qualidade das questões de Estatística presentes no ENEM não favorece o letramento estatístico e probabilístico, segundo Goulart (2015). Neste caso, o conhecimento dos elementos do contexto, como exposto por Gal (2002), parece ter influenciado a análise dos estudantes, assunto já discutido na seção 5.1.

Os respondentes que alegaram desconhecer como são realizadas as pesquisas estatísticas (Q112) apresentaram comportamento similar ao dos alunos que acertaram a questão 27 (Q27b), a qual apresentou uma situação típica daquelas estudadas em outras disciplinas, como Geografia ou Sociologia, ou até mesmo por políticos em campanhas eleitorais, pois abordou o desemprego no Brasil — lembramos que o país vivia o ápice das discussões políticas sobre eleições majoritárias e que boa parte desses estudantes já eram eleitores¹³. Então, não ter conhecimento sobre os procedimentos típicos da realização de uma pesquisa estatística aparentemente não comprometeu análise do gráfico de linha.

¹³ No Brasil, o voto é facultativo a partir dos 16 anos de idade e obrigatório a partir dos 18.

Figura 32 – Árvore de similaridades: relações significativas IV



Fonte: o autor, 2020.

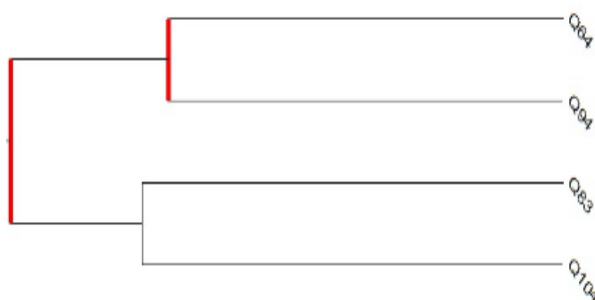
Apontamos as seguintes legendas para as siglas da figura 32:

- Q4S – Estudou Estatística em outro curso, fora do ensino regular.
- Q28a – Assinalou a alternativa a da questão 28 (alternativa errada).
- Q72 – Discorda totalmente de que a Estatística é importante para tomar decisões em seu dia a dia.
- Q27c – Assinalou a alternativa c da questão 27 (alternativa errada).
- Q82 – Não acredita parcialmente que irá utilizar a Estatística no(s) curso(s) superior(es) que pretende fazer.

Os alunos que estudaram Estatística em outro curso, fora do ensino regular (Q4S), apresentaram comportamento similar ao daqueles que não obtiveram êxito na questão (Q28a) que abordou o desvio padrão. Consideramos essa reação compreensível, pois, ainda que a proposta curricular vigente nessa escola previsse a Estatística no ensino fundamental, o estudo do desvio padrão estava reservado ao 3º ano do ensino médio, ou seja, em um momento posterior à nossa coleta de dados.

Os alunos que discordaram parcialmente de que a Estatística é importante para o processo de tomada de decisões em problemas do dia a dia (Q72) apresentaram comportamento similar ao dos que não foram bem-sucedidos na questão Q27 — que envolvia leitura e interpretação de um gráfico de linha. Esses grupos, por sua vez, manifestaram comportamento similar ao dos estudantes que não estavam muito certos quanto à importância da Estatística na tomada de decisões no enfrentamento dos problemas do dia a dia, revelando dificuldades encontradas com frequência no ensino e aprendizagem da Estatística (BATANERO et al., 1994).

Figura 33 – Árvore de similaridades: relações significativas V



Fonte: o autor, 2020.

Apontamos as seguintes legendas para as siglas da figura 33:

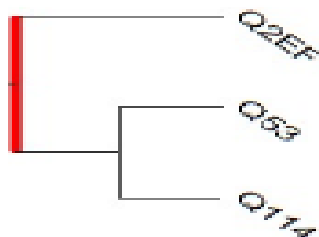
- Q64 – Concorda totalmente que a Estatística é importante para compreender as notícias do rádio, da TV, da internet e dos jornais.
- Q94 – Acredita totalmente que a Estatística é uma área da Matemática, portanto, uma ciência exata.
- Q83 – Acredita parcialmente que irá utilizar a Estatística no(s) curso(s) superior(es) que pretende fazer.
- Q104 – Confia totalmente em pesquisas estatísticas.

Os respondentes que avaliaram a Estatística como extremamente importante para a compreensão das mídias sociais (Q64) apresentaram comportamento similar àqueles que acertaram a classificação da Estatística como uma ciência exata (Q94), de modo semelhante àqueles que consideraram a possibilidade de utilizar os conhecimentos estatísticos no ensino superior (Q83) e tinham total confiança nas pesquisas estatísticas (Q104).

Sua compreensão da Estatística como ciência exata pode lhe conferir credibilidade, graças ao estereótipo dessa área. Além disso, o entendimento de que a Estatística é necessária para compreender notícias da mídia é reforçada por sua presença no material didático — livro didático e Caderno do Aluno — de outras áreas, como de Língua Portuguesa, no qual esteve presente, inclusive, em avaliações institucionais da disciplina, por exemplo na AAP¹⁴. Essa interpenetração e impregnação mútua das áreas de Matemática e Língua Portuguesa foi discutida por nós (GIORDANO, 2016).

¹⁴ Avaliação de Aprendizagem em Processo.

Figura 35 – Árvore de similaridades: relações significativas VII



Fonte: o autor, 2020.

Apontamos as seguintes legendas para as siglas da figura 35:

- Q2EF – Estudou Estatística no ensino fundamental.
- Q53 – Bom nível de conhecimento em Estatística.
- Q114 – Sabe muito bem como são realizadas as pesquisas estatísticas.

Os alunos que alegaram ter estudado a Estatística no ensino fundamental (Q2F) têm comportamento similar ao daqueles que julgam possuir bons conhecimentos sobre essa ciência (Q53) e declararam saber muito bem como são realizadas as pesquisas estatísticas (Q114), apresentando uma concepção simplista sobre ela. Porém, ao longo dos trabalhos em sala de aula, ficou claro que muitos deles confundiam pesquisas estatísticas com meras enquetes, subestimando a complexidade dessa ciência.

Figura 36 – Árvore de similaridades: relações significativas VIII



Fonte: o autor, 2020.

Apontamos as seguintes legendas para as siglas da figura 36:

- Q27d – Assinalou a alternativa d da questão 27 (alternativa errada).
- Q29e – Assinalou a alternativa e da questão 29 (alternativa errada).
- Q28d – Assinalou a alternativa d da questão 28 (alternativa errada).

Os alunos que erraram as questões Q27 e Q29 manifestaram comportamento similar àqueles que erraram a questão Q28. Podemos inferir, então, que a dificuldade em ler e interpretar gráficos estatísticos, tabelas e quadros pode ser, em parte, provocada pela precária exploração de tabelas e gráficos estatísticos da Matemática e da Estatística — como aponta a pesquisa de Lucas Cobello (2018).

Isso pode estar associado à natureza das atividades presentes no material didático dos alunos da rede estadual paulista de ensino. Tais atividades enfatizam aspectos algébricos procedimentais, em prejuízo aos aspectos crítico-reflexivos, valorizando a atividade individual em detrimento à produção coletiva e a realização de projetos de pesquisa. Essa tendência contraria os pressupostos da AED (BATANERO; ESTEPA; GODINO, 1991).

Tanto o livro didático quanto o material didático institucional — o Caderno do Aluno (SÃO PAULO, 2014) — apresentam uma abordagem que deixa muito a desejar, com exercícios que exigem quase que exclusivamente níveis de leitura básicos na escala de Curcio (1989) e Friel, Curcio e Bright (2001), como é discutido por Coutinho (2013), Goulart (2015) e por nós (GIORDANO, 2016).

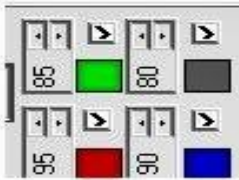
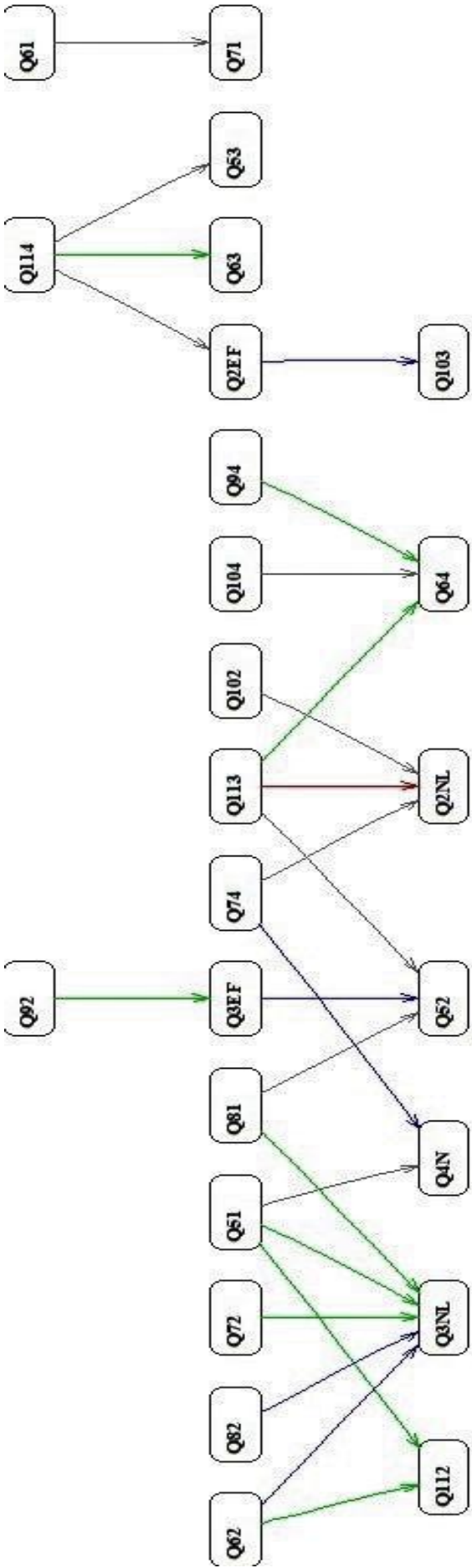
5.3 Análise Implicativa

Construímos nossa análise baseados nas relações emergentes da análise implicativa por meio da teoria clássica e da distribuição binomial. A primeira associa as variáveis A e B segundo uma meta-regra em que “se A ocorre, então provavelmente B ocorre”. Nela, as setas coloridas indicam os níveis de confiança, segundo as quais podemos ler tais associações.

Em nosso grafo, as setas cinza indicam valores menores ou iguais a 0,80; as verdes, valores entre 0,81 e 0,85; as azuis, variação entre 0,86 e 0,90; e, finalmente, as vermelhas, de 0,91 até 0,95. Tal escala é representada a seguir, na Figura 30.

Os resultados da análise implicativa, juntos aos da análise de similaridades e da análise coesitiva, nos permitiram elaborar um cenário para os conhecimentos estatísticos dos alunos que, na sequência da fase experimental de nossa pesquisa, construíram seus conhecimentos e desenvolveram seu letramento estatístico com base na resolução de problemas após o desenvolvimento por meio de projetos.

Figura 37 – Grafo implicativo



Fonte: o autor, 2020

Apontamos as seguintes legendas para as siglas da figura 37:

- Q113 – Sabe bem como são realizadas as pesquisas estatísticas.
- Q2NL – Não se lembra se estudou Estatística no ensino fundamental.
- Q74 – Concorda totalmente que a Estatística é importante para tomar decisões em seu dia a dia.
- Q102 – Não confia parcialmente em pesquisas estatísticas.
- Q3NL – Não se lembra se estudou Estatística no ensino médio.

Começamos nossa análise pelo caminho que apresenta uma seta vermelha, ou seja, o de maior nível de confiança: Q113 → Q2NL, o qual significa que se afirmou concordância com a afirmação “sabe como são realizadas as pesquisas estatísticas”, então, provavelmente, não se lembra se “estudou Estatística no ensino fundamental”.

A variável típica desse comportamento é o sexo masculino, com risco 0,195. Pela presença maciça da Estatística nas mídias, sobretudo considerando o período eleitoral em que a pesquisa foi realizada, podemos inferir que esse grupo compreendeu apenas de forma intuitiva as informações dadas nessas pesquisas. Temos, aqui, um primeiro indicativo sobre a necessidade de iniciar a abordagem dos conteúdos estatísticos sem pressupor conhecimentos estáveis.

Mantendo-nos nos caminhos que terminam também em Q2NL, passamos para Q74 → Q2NL — acreditam que a Estatística serve para a tomada de decisões no dia a dia — e Q102 → Q2NL — discordam que confiam em pesquisas estatísticas. Em ambos, observamos as mesmas variáveis típicas: sexo feminino com risco 0,246. Este grupo, apesar de partir de premissas distintas, não se lembra se cursou Estatística no ensino fundamental (Q2NL). Uma outra variável que também está relacionada aos vários caminhos implicativos é a Q3NL, a qual significa que não se lembra se estudou Estatística no ensino médio, assinalada por 41 dos 86 estudantes. Recordamos que os respondentes estão no 3º ano do ensino médio e ainda não cumpriram os conteúdos de Estatística do currículo do Estado de São Paulo, segundo o qual os alunos estudam Probabilidade ao longo do 2º ano do ensino médio, levando-nos a inferir que não associam a Probabilidade com a Estatística. Vale ressaltar que, tanto no Caderno do Aluno (SÃO PAULO, 2014) quanto nos livros didáticos (GIORDANO, 2016), a Estatística é apresentada de maneira isolada.

O quadro a seguir apresenta os caminhos encontrados.

Quadro 4 – Caminhos implicativos determinados na análise das variáveis em jogo

Caminho	Significado do ponto de partida
Q62 → Q3NL	Discordou de que a Estatística é importante para compreender as notícias do rádio, da TV, da internet e dos jornais.
Q82 → Q3NL	Não acreditou que iria utilizar a Estatística no(s) curso(s) superior(es) que pretende fazer.
Q72 → Q3NL	Discordou de que a Estatística é importante para tomar decisões em seu dia a dia.
Q51 → Q3NL	Classificou seu nível de conhecimentos em Estatística (1 – ausência de conhecimento).
Q81 → Q3NL	Desacreditou totalmente de que iria utilizar a Estatística no(s) curso(s) superior(es) que pretende fazer.

Fonte: o autor, 2020.

Os caminhos demarcados permitiram inferir a pouca importância que os alunos atribuem aos conceitos estatísticos em suas vidas até o momento: não acreditam em sua utilidade no dia a dia nem em seu futuro profissional e alegam possuir poucos conhecimentos estatísticos sobre esses conceitos, implicando, provavelmente, na não lembrança de já ter cursado a disciplina no ensino médio. Tais inferências são compreendidas em um nível de confiança entre 0,85 e 0,90. Sobre variáveis típicas, nada afirmaremos, uma vez que o risco assinalado nesses caminhos está em torno de 0,49, independentemente da variável assinalada.

Isso tudo nos leva a inferir a necessidade de uma cuidadosa abordagem dos conteúdos estatísticos, indicando a importância do desenvolvimento completo do letramento estatístico dos alunos, o que guia, certamente, as ações da fase experimental com a resolução de problemas após o desenvolvimento dos projetos por eles.

Em síntese, podemos dizer que esses alunos tiveram pouco contato com a Estatística em ambiente escolar. Deles, 65% não se recordam de ter estudado Estatística no ensino fundamental, enquanto 23% afirmam que realmente não a estudaram. Quanto a seu ensino no ensino médio, 48% não se lembram ter visto Estatística nas aulas, enquanto 26% alegam não a ter estudado até então e 87% asseveram não ter estudado Estatística fora do ensino regular.

Esses dados nos ajudam a compreender seu desempenho nas 5 questões extraídas de versões do ENEM: a maioria (57%) demonstrou possuir algum conhecimento sobre cálculo de média aritmética simples, acertando a questão Q25;

quando a média foi apresentada com a mediana e a moda na questão Q26, o percentual de acertos caiu para 27%; ao passo que a questão que solicitava somente o cálculo da mediana (Q27) teve apenas 21% de acertos.

Cabe ressaltar que, provavelmente, esses alunos calcularam média aritmética em outras disciplinas, como Geografia ou Física, e em contexto nos quais estavam familiarizados. É o que observamos na questão Q28 — a qual abordava o contexto escolar —, em que 44% dos estudantes conseguiram calcular a nota média, pois o conhecimento de contexto é um importante elemento no letramento estatístico, segundo Gal (2002). Além disso, a familiaridade com a presença da Estatística na Geografia, ou seja, no estudo de dados socioeconômicos, pode justificar o percentual de acertos na questão 29.

Assim, concluímos que, mesmo com pouco ou nenhum estudo específico de Estatística dentro da disciplina de Matemática, os alunos possuíam conhecimentos básicos sobre essa ciência, no que se refere a medidas de posição, sobretudo de tendência central, gráficos estatísticos e tabelas de distribuição de frequência.

A seguir, prosseguiremos com nossa análise, a partir da produção dos grupos de alunos, aferida por meio de gravações e registros escritos, à luz do modelo cK ϕ .

6 ANÁLISE DAS CONCEPÇÕES DOS ALUNOS

Neste capítulo, apresentaremos e analisaremos alguns resultados obtidos, com base nas gravações e registros por escrito das produções coletivas de quatro grupos de alunos ao resolverem problemas de Estatística, no contexto do trabalho por meio de projetos. Os procedimentos para coleta já foram explanados no capítulo 2 desta tese, ao passo que os registros sobre a produção coletiva dos estudantes são apresentados no Anexo 2 e no Apêndice 2 deste trabalho.

Nossa segunda fonte de dados é constituída pelas produções de quatro grupos de alunos (dois trios e duas duplas), que responderam a três questões de Estatística, em novembro de 2018, após o desenvolvimento de suas próprias pesquisas, em uma abordagem de ensino e aprendizagem por meio de projetos.

Nos moldes da Teoria das Concepções, é inviável analisar o desenvolvimento de um projeto do início ao fim, pois tal processo pode durar semanas e parte do trabalho é realizada fora do ambiente escolar, como a casa do aluno ou alguma biblioteca pública, e, no caso da coleta de dados, nos lugares mais variados, como praças, shoppings e hospitais.

Propusemos aos estudantes a filmagem de suas apresentações, mas só daríamos prosseguimento a essa ideia se todos concordassem. Uma vez que não houve unanimidade no consentimento, nos limitamos a gravar a atividade apresentada neste capítulo e filmar entrevistas com alunos e professores voluntários. Como já explicamos, excluímos as entrevistas da pesquisa e focalizamos a atividade de resolução de problemas em grupos. Então, apresentamos alguns problemas similares àqueles presentes nos projetos desenvolvidos pelos estudantes, para serem resolvidos de forma colaborativa, em pequenos grupos.

Diaz (2016) destaca que, segundo os alunos, o trabalho colaborativo em pequenos grupos reduz a ansiedade e contribui para uma maior compreensão de noções estatísticas e aquisição de experiência. De acordo com a autora, as atividades colaborativas são mais tranquilizadoras, motivadoras e estimulantes:

A observação de colegas, trabalhando ativamente, favorece a concentração sobre a tarefa e incentiva a emergência de uma diversidade de propostas. O incentivo à ajuda mútua permite a assimilação de conceitos, o progresso nas atividades, a redução da percepção de dificuldade da tarefa e diminuir a ansiedade. Os cuidados coletivos reduzem o fardo das dificuldades; trocar e compartilhar melhora a autoconfiança e compromisso coletivo. (ibid. p. 277, tradução nossa)

Concordamos com as observações de Diaz, as quais são muito próximas daquelas realizadas por Garfield (1993, 2013), em suas investigações sobre o trabalho cooperativo no ensino e aprendizagem de Estatística. Nessa perspectiva, propusemos as tarefas em pequenos grupos, após a realização dos projetos.

Inicialmente, falaremos um pouco sobre tais projetos, desenvolvidos durante o período de 1 bimestre letivo (nos meses de agosto e setembro) de 2018, com a frequência regular de 5 aulas de Matemática, de 50 minutos cada, por semana, e apresentados no anfiteatro da escola em outubro do mesmo ano.

No quadro a seguir, estão registrados os projetos desenvolvidos pelos estudantes, sendo que somente consideramos os concluídos e apresentados. Se algum grupo não cumpriu todas as etapas (alguns alunos não quiseram apresentar) ou se houve plágio (tivemos um caso), o trabalho foi desconsiderado.

Quadro 5 – Resumo dos projetos

Tema escolhido	Participantes	Objetivos	Sujeitos da pesquisa	Disciplina auxiliar
Regime Militar: 1964 – 1985	5 alunos de 17 a 19 anos de idade.	Avaliar o nível de conhecimentos dos respondentes sobre o regime militar (de 1964 a 1985) e saber se, na opinião deles, o país esteve melhor nesse período, em termos sociopolítico-econômicos.	22 pessoas de 20 a 90 anos de idade, de ambos os sexos e com diferentes níveis de escolaridade.	História
Influência das línguas estrangeiras no mercado de trabalho	2 alunos e 4 alunas de 17 ou 18 anos de idade.	Investigar a percepção dos entrevistados sobre a importância do conhecimento de línguas estrangeiras para o ingresso e permanência no mercado de trabalho brasileiro na atualidade.	42 pessoas com idades superiores a 16 anos, de ambos os sexos.	Inglês
Questões de gênero: homens e mulheres no esporte	6 alunas de 16 a 19 anos de idade.	Avaliar a opinião de professores e futuros professores de Educação Física sobre as questões de gênero no esporte, desde o incentivo à prática e formação de base até o investimento no esporte profissional.	50 professores e alunos de Educação Física de 23 a 52 anos de idade, de ambos os sexos.	Educação Física
Eleições majoritárias no Brasil, em 2018	5 alunos e 1 aluna de 17 a 19 anos de idade.	Investigar possíveis influências e relações existentes entre jovens de 16 a 18 anos e as eleições para presidente no Brasil em 2018.	50 alunos do ensino médio, de 16 a 18 anos de idade e de ambos os sexos.	Sociologia
Inflação no Brasil e no mundo	3 alunos e 1 aluna de 17 a 19 anos de idade.	Avaliar o nível de conhecimentos dos respondentes sobre a inflação e a economia mundial nos últimos 20 anos.	42 pessoas, (homens, mulheres e não binários) com idades superiores a 16 anos.	Geografia
Legalização do porte de armas	5 alunas e 1 aluno de 16 e 17 anos de idade.	Saber se o respondente defendia a legalização do porte de armas de fogo, se ele acreditava que isso contribuiria para a redução da violência no país e se ele se considerava apto a usar uma arma.	60 pessoas com idades superiores a 15 anos, de ambos os sexos.	Filosofia
Gravidez na adolescência	2 alunos e 3 alunas de 17 anos de idade.	Avaliar o nível de conhecimento e as opiniões dos respondentes sobre a relação entre gravidez precoce e saúde da mulher.	50 mulheres de 14 a 40 anos de idade.	Biologia

Métodos contraceptivos	2 alunas e 1 aluno de 16 a 19 anos de idade.	Avaliar o nível de conhecimento dos respondentes sobre os métodos contraceptivos mais populares, suas experiências com eles, qual o método preferido e o porquê.	35 pessoas de 16 a 67 anos de idade, de ambos os sexos.	Biologia
Legalização da maconha	5 alunos e 1 aluna de 16 a 19 anos de idade.	Saber se o respondente defendia a legalização do consumo da maconha, avaliar seu nível de conhecimento sobre ela e verificar se o respondente associava seu consumo à violência urbana.	36 pessoas de 15 a 44 anos de idade, de ambos os sexos.	Sociologia
Alcoolismo	2 alunas de 17 anos de idade.	Avaliar o nível de conhecimento e as opiniões dos respondentes sobre os efeitos dessa dependência química sobre o corpo humano.	30 pessoas de 18 a 70 anos de idade, de ambos os sexos.	Química
Uso de redes sociais	6 alunos, de 16 e 17 anos de idade.	Investigar sobre o uso de redes sociais e sua importância para os jovens brasileiros em nossos dias.	30 pessoas com idades superiores a 10 anos, de ambos os sexos.	Artes
Educação brasileira	3 alunas de 17 anos de idade.	Avaliar o nível de conhecimento e as opiniões dos respondentes sobre a qualidade do ensino e os rumos da educação brasileira.	30 pessoas com idades variadas, de ambos os sexos.	Língua Portuguesa
Violência sofrida por jovens estudantes	3 alunas de 17 a 19 anos de idade.	Avaliar o nível de conhecimento e as opiniões das estudantes sobre as formas de violência que sofreram e os direitos da mulher.	30 alunas de 14 a 18 anos de idade.	Sociologia
Tabagismo e orçamento familiar	2 alunos e 2 alunas de 16 a 18 anos de idade.	Tabagismo: impactos do consumo de cigarros no orçamento do fumante.	30 pessoas de 15 a 64 anos de idade, de ambos os sexos.	Física

Fonte: o autor, 2020.

Em cada trabalho, os alunos apresentaram:

- justificativa da escolha do tema, destacando a relevância da sua pesquisa;
- questão de pesquisa;
- objetivos da pesquisa;
- problematização;
- instrumento de coleta de dados;
- procedimentos metodológicos;
- apresentação dos dados por meio de medidas de tendência central e de dispersão;
- apresentação dos dados por meio de gráficos e tabelas estatísticas;
- discussão dos resultados;
- considerações finais;
- anexos (todos os questionários aplicados);
- pôster para a divulgação dos resultados.

Todos os grupos optaram pelo questionário impresso como instrumento de coleta de dados. Em anos anteriores, outros grupos optaram por entrevistas abertas, filmagem, coleta de dados realizada por telefone e pela internet utilizando redes sociais etc. Segundo Oriol (2007, p. 258, tradução nossa):

A pesquisa por questionários mobiliza quase a totalidade das competências em ação no trabalho de Estatística: coleta de dados, organização dos dados, verificação da coerência, escolhas técnicas, escolhas tecnológicas, tratamentos, validação.

Foram oferecidas aos alunos três formas de apresentação dos resultados de pesquisa: criação de um site de divulgação dos trabalhos; realização de seminários em sala de aula, apenas para os colegas da mesma turma; ou painel para toda a comunidade escolar (alunos de outros anos, professores, membros da equipe de gestão escolar, demais funcionários da escola, pais e amigos). A escolha dos estudantes, por votação direta, foi pelo painel.

Após as apresentações das pesquisas realizadas pelos alunos no anfiteatro da escola, solicitamos voluntários para participar de uma atividade que envolvia a resolução de problemas de Estatística, semelhantes àqueles que eles encontraram no desenvolvimento de seus projetos. Dez estudantes apresentaram-se e foram

divididos em quatro grupos de dois ou três alunos: um grupo de cada uma das quatro classes do terceiro ano do ensino médio dessa escola.

Sua participação voluntária durante a resolução de problemas de Estatística foi audiogravada, e cada grupo levou de uma a três sessões de 100 minutos, em três dias distintos de uma mesma semana, para concluir as tarefas. No Apêndice 2, nos Protocolos 1, 2, 3 e 4, apresentamos a íntegra das respostas e das interações intragrupos.

Tal produção foi analisada à luz da Teoria das Concepções. Essa análise será apresentada a seguir.

6.1 Concepções mobilizadas pelos alunos

Nosso objetivo principal nessa atividade foi o de promover a reflexão e a discussão sobre a importância das ferramentas da Estatística para a análise de uma situação realista na resolução de problemas após o desenvolvimento de projetos (disponível no Apêndice 2 e Anexo 2). Esperávamos que os alunos pudessem identificar, medir, descrever e justificar a variação de dados.

Entre os elementos de Estatística explorados pelos alunos, destacamos as medidas-resumo de tendência central e de dispersão, além da leitura, da construção e da interpretação de tabelas e gráficos estatísticos.

Buscamos identificar as concepções mobilizadas com seus respectivos operadores e estruturas de controle, tal como foi explicado no capítulo 3 desta tese.

Em consonância com o Currículo de Matemática do Ensino Médio (SÃO PAULO, 2012), consideramos estes conteúdos matemáticos:

- medidas de tendência central;
- medidas de dispersão;
- gráficos estatísticos;
- tabelas de distribuição de frequências.

Balacheff e Margolinas (2005), ao exemplificar concepções, apresentaram estes quatro exemplos:

Concepção 1: «cantiga» IIIII & IIII

P: Problemas do tipo: Estas duas coleções de 5 objetos e 4 objetos fazem juntos quantos objetos? (As coleções estão presentes, distintas e manipuláveis, os números são pequenos).

R: Apontar e nomear duas coleções distintas, começando pela primeira e continuando pela segunda para obter o número total.

L: Gesto de apontar/nomear, cantiga oral.

Σ : Controle do número obtido, controle do procedimento.

Concepção 2: sobrecontagem de 16 & 4

P: Problemas de tipo: Quantos são 16 e 4? (Os números são nomeados, as coleções estão ausentes, um dos números não deve exceder cerca de 10).

R: Pronunciar os números enquanto desdobra os dedos sucessivamente.

L: Gesto da implantação dos dedos, pronunciar um determinado número.

Σ : Controle audiovisual da correspondência entre o número de dedos indicados e os números listados, o controle da pronúncia.

Concepção 3: adição «decimal» 16+23.

P: Todo problema de adição a dois ou mais números inteiros.

R: Algoritmo (usual) da adição.

L: Escritura decimal dos números.

Σ : Controle passo a passo do algoritmo.

Concepção 4: calculadora [1][6][+][2][3][=].

P: Todo problema de adição a dois ou mais números inteiros ou decimais relativos.

R: Continuação das teclas da calculadora.

L: Gestos sobre as teclas e exibições sucessivas na tela.

Σ : Controle dos gestos sobre as teclas.

(BALACHEFF; MARGOLINAS, 2005 p. 82, tradução nossa)

Desses quatro elementos, focalizaremos os operadores (R) e as estruturas de controle (Σ). Mas é importante dizer que, no campo de problemas (P), o aluno deveria ler e interpretar problemas; escolher uma ou mais medidas-resumo para explicar a variabilidade das idades, da renda familiar e da escolaridade dos pais dos elementos da amostra; determinar tais medidas-resumo; escolher uma ou mais representações gráficas para explicar a variabilidade das idades, da renda familiar e da escolaridade dos pais dos elementos da amostra; construir gráficos estatísticos bem como quadros e tabelas estatísticas; justificar suas escolhas; relacionar as medidas-resumo e representações gráficas/tabulares; e explicar seus procedimentos utilizados, por meio de produção escrita e/ou oral. Além disso, no conjunto das representações (L), temos representações orais, textuais, numéricas, algébricas, tabulares e gráficas (ver Apêndice 2 e Anexo 2).

Passaremos, então, à análise das respostas dos alunos, com ênfase nos operadores e nas estruturas de controle.

6.2 Análise das possíveis concepções mobilizadas pelos alunos na resolução de problemas após a realização de projetos

Um dos nossos objetivos, apresentados no capítulo 1, foi identificar as concepções estatísticas mobilizadas pelos alunos. Para tanto, realizamos a atividade apresentada no Apêndice 2 e no Anexo 2.

Acrescentamos que a atividade reflete, em parte, os projetos que os alunos desenvolveram. Assim, por exemplo, os eles determinam a mediana com base em dados agrupados, na referida atividade, porque foi dessa maneira que constatamos ter sido feito na maioria dos projetos por eles apresentados.

A atividade foi executada por 10 estudantes, distribuídos em 4 grupos de 2 ou 3 alunos cada e, ao todo, foram registradas pouco mais de 10 horas ao longo de 10 sessões, como podemos observar no quadro a seguir:

Quadro 6 – Número e duração das sessões de resolução de problemas dos quatro grupos de alunos

Grupo	N.º de sessões	Duração
Grupo I	3	02:49:27
Grupo II	1	01:27:56
Grupo III	2	03:32:58
Grupo IV	2	02:30:22
Total	8	10:20:43

Fonte: o autor, 2020.

As concepções constituem fenômenos muito específicos, são casos particulares e refletem o olhar do pesquisador, com seus conhecimentos, valores e idiossincrasias, sobre um sujeito único, a despeito de todo o rigor científico da pesquisa. Entretanto, essa também é uma característica de nosso tipo de investigação, o estudo de caso (PONTE, 2006; FIORENTINI e LORENZATO, 2007; CRESWELL, 2010; YIN, 2015). Assim, iniciaremos nossa análise buscando identificar as concepções mobilizadas por esses alunos.

Observamos, nos quatro grupos, a existência de ideias contraditórias, confrontadas explicitamente em seus diálogos, indícios de mobilização de diferentes concepções, até que uma delas prevalecesse. No Grupo I, por exemplo, a aluna denominada "S", de 19 anos de idade, interage com as colegas denominadas "D" e "Z", ambas de 17 anos, enquanto resolvem o primeiro exercício proposto (disponível no Apêndice 2):

Figura 38 – Questão 1 dos exercícios aplicados

1) Em uma escola pública de Santo André, um grupo de alunos do terceiro ano do Ensino Médio, realizando uma pesquisa escolar, resolveu investigar a opinião dos colegas sobre a política de cotas para o ingresso em universidades federais. Essa escola possuía 938 alunos, sendo 417 estudantes do Ensino Fundamental, estudando no período vespertino (6º ao 9º ano) e 521 no Ensino Médio, estudando no período matutino (204 no 1º ano, 179 no 2º ano e 138 no 3º ano). Desses 938 alunos, 513 são mulheres. Esse grupo decidiu entrevistar 40 alunos, sendo 20 (dez homens e dez mulheres) do período matutino e 20 (dez homens e dez mulheres) do período vespertino. Elaborou um questionário com doze perguntas, cujas respostas podem ser observadas no banco de dados que segue. Analise os resultados utilizando as representações que julgar adequadas e as medidas que julgar adequadas. Lembre-se de sempre explicar ao máximo seu pensamento.

a) Com relação à variável “idade”, os alunos forneceram as seguintes respostas:

Idade (anos)	1	1	1	1	1	1	1	1	1	2	2	Total
	1	2	3	4	5	6	7	8	9	0	1	
Freq. simples	3	4	4	6	5	8	4	3	2	0	1	40

Como você explica a variação das idades utilizando as ferramentas estatísticas que aprendeu: construa uma ou mais representações gráficas, explicando-as. Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de idades do grupo de 40 alunos pesquisados.

Fonte: o autor, 2020.

A transcrição de todos os diálogos a seguir é literal, isto é, não fizemos correção alguma, a fim de preservar sua naturalidade.

O Grupo I foi formado por três alunas, todas colegas de uma mesma turma: D, de 17 anos de idade; S, de 19 anos; e Z, de 17 anos.

O diálogo entre as duas começa com uma discussão sobre o que é média. Embora o enunciado da referida questão solicitasse, em primeiro lugar, representações gráficas, as alunas iniciaram as tarefas pela determinação das medidas-resumo, tal qual é apresentado no seu material didático.

3. Z: Vamos achar primeiro a média, né?
4. S: Mas tem que fazer esse aqui...
5. Z: Não, é 'a', tem que fazer primeiro a 'a'. (Referência ao item 'a' apresentado acima.)
6. S: Tá.
7. Z: Tá, primeiro a gente vai achar a média. Que que é a média?
8. S: Da idade. A média é o que tem mais?
9. Z: Não, essa aí é a moda. A média é que tem que pegar o número, fazer vezes e depois dividir, não é?
10. Z: Tá, primeiro a gente vai achar a moda. A moda é o que tem mais.
11. S: Moda... 16 anos, não é?
12. Z: É.

Na fala número 7, Z pergunta à amiga o que é a média, mas, logo em seguida, na fala 9, corrige a colega S, que confunde a moda com a média, demonstrando possuir conhecimentos estatísticos sobre a média ligados ao algoritmo.

A aluna Z corrige a colega, como podemos ver na fala 28, mas volta a manifestar sua dúvida, na fala 32 a seguir:

20. Z: Faz as contas aí...
21. S: Primeiro tem que fazer 11×3 ... que vai dar 33. Aí depois a gente tem que fazer...
22. Z: A moda é 16 anos.
23. S: ... 12×4 ... Eu acho que tem que pegar...
24. Z: Fala baixo!
25. S: Eu acho, ué? Eu acho que tem que pegar 33 somar com 48 e dividir por 2, né?
26. D: Dividir por 40!
27. S: Pegar todo o total e dividir por 40. É a mediana.
28. Z: A mediana não, esse é a média.
29. S: É, essa é a média, e a mediana?
30. Z: Vamo achar primeiro a moda, a moda é qual? 16, né? 16 anos.
31. D: 16 anos!
32. Z: E a média é o que?
33. S: Não sei, eu tava falando, aí você começou...
34. Z: Taqui ó...
35. S: Somar tudo e dividir por 4.
36. D: Por 40!

A aluna S soma os produtos das idades pelas suas respectivas frequências e chega ao resultado errado: 606. Z, então, desconfia e pede que o valor seja recalculado:

46. Z: Soma de novo! <RELÊ A TABELA>
47. S: Acho que eu erreí.
48. Z: Deu quanto?
49. S: 15,25.

50. Z: Deu quanto?

51. D: 601.

52. Z: O seu também, né? Erramos na primeira! Tá vendo? Então dá quanto?

53. S: Dividido por 40? 15,025.

O grupo chega ao valor correto, mas, conforme uma concepção já apresentada por outros autores, S não aceita o resultado, pois não encontra a idade na tabela, ignorando uma das propriedades da média, discutida por Strauss e Bichler (1988), de que a média não coincide, necessariamente, com um dos valores da amostra. O grupo assume, então, o próximo valor na tabela: 16 anos. Z, no entanto, rejeita e considera que a média está entre 15 e 16 anos, visto na fala 61 a seguir:

55. Z: Deixa eu ver.

56. S: 15,025.

57. Z: Então a média tá...

58. S: Tá no '16 anos'.

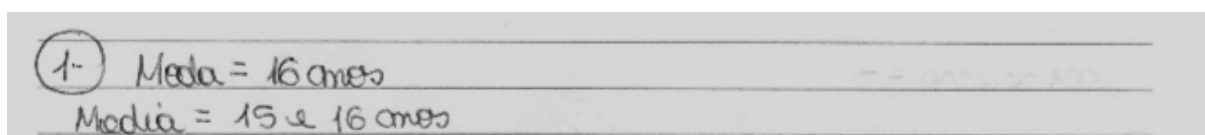
59. Z: Então a gente achou a média e a moda.

60. D: Coloca aí.

61. Z: Entre o '15' e o '16' está a média.

O registro escrito, no entanto, ficou por conta de S, que anotou a média como “15 e 16 anos”, mesmo após as respostas “15,025” e “entre o 15 e o 16”:

Figura 39 – Excerto I do Apêndice 2, Protocolo 1 – Grupo I



Fonte: o autor, 2020.

Apesar de manifestar suas dúvidas, Z apresentou estabilidade de sua concepção de média aritmética, mediando a discussão entre S e D. Mais adiante, Z volta a manifestar dúvidas com relação à média, à mediana e à moda (ao todo, temos 368 falas nas três sessões deste trio), mas, assumindo a liderança do grupo, reafirma suas concepções sobre essas medidas de tendência central, dando indícios de aprendizagem.

Pode existir, por detrás dessa resposta, uma concepção Φ_1 , caracterizada pelos elementos:

- P – campo de problemas: determinação da média aritmética simples com base nos valores apresentados em uma tabela.

- R – operadores: localização dos valores das idades associadas aos sujeitos da amostra em uma tabela e verificação de suas respectivas frequências. Estabelecimento da razão entre o somatório dos produtos desses dois valores e o número total de elementos da distribuição de frequências.
- L – conjunto de representações: representação numérica e representação tabular.
- Σ – estrutura de controle: controle passo a passo do algoritmo da média aritmética simples.

Assim, podemos inferir a existência de uma possível concepção Φ_1 , segundo a qual o valor da média será o quociente entre o somatório dos produtos das frequências f pelos respectivos valores x , assumidos pela variável em análise, e o número total n de elementos da amostra.

Diferentemente da média, a moda foi rapidamente determinada, confirmando as observações de Batanero (2001) e Batanero e Godino (2001), ao comparar o grau de dificuldade envolvido na determinação das medidas de tendência central: moda, mediana e média aritmética simples. Consideramos a possível existência da concepção Φ_2 , da aluna S, caracterizada pelos elementos:

- P – campo de problemas: determinação da moda com base nos valores apresentados em uma tabela.
- R – operadores: localização das frequências associadas à variável 'idade' em uma tabela.
- L – conjunto de representações: representação numérica e representação tabular.
- Σ – estrutura de controle: comparação dos valores das frequências associadas à variável analisada e identificação do maior valor.

Assim, podemos inferir a existência de uma possível concepção Φ_2 , segundo a qual, se a variável for quantitativa discreta ou qualitativa, a moda é o efetivo associado à frequência máxima.

Prosseguindo a atividade, S, Z e D tentaram determinar a mediana:

62. S: O que é agora? A mediana?

63. Z: A mediana...
64. S: Tem que passar à caneta.
65. Z: Depois a gente passa. Como é que era a mediana, mesmo? <PAUSA>
Ah! A mediana é metade. O total é 40. A metade de 40 é 20. A gente tem que achar quem é que tá no '20'. Quem é que tá dentro do '20'.
66. D: 19?
67. Z: Não.
68. D: 21 e 20?
69. Z: Não! 40. Quem tá dentro do '20'. É a metade. Tipo, qual idade tá dentro do '20', entendeu?
70. D: Então é 10?
71. Z: Não.
72. D: A metade de 20 é 10.
73. Z: Não. <RISOS> Não tem nem 10 aqui!
74. S: Não tem nem 10! <RISOS> Começa no '11', amiga.
75. Z: Tem que achar a metade! Tá vendo isso aqui tudo?
76. D: Aham.
77. Z: Deu 40, certo? A metade de 40? 20.
78. S: 20.
79. Z: Mais ou menos aqui. Deixa eu ver... Então a gente vai ter que fazer as contas pra ver quem tá no '20'. Deixa eu ver... <PAUSA> O total deu o quê?
80. S: Deu 40.
81. Z: 40 pessoas?
82. S: É.
83. Z: É, 40 alunos. O 20º aluno é o que eu quero.
84. S: Tá, mas onde tá o 20º aluno?
85. Z: A gente vai ter que fazer as contas. <RISOS> <PAUSA> Abre o livro aí, vê se você acha alguma coisa, um exemplo pra gente ver. <PAUSA> Abre na página 46. <SUSSURROS>
86. S: Nós tem que mudar aqui, você colocou 106, é 101. Empresta a borracha. <PAUSA LONGA>
87. Z: Deixa eu ver, mano, deixa eu ver aqui rapidinho. Que coisa que é essa? 12? <PAUSA> Certo! A mediana é um negócio complicado!
88. S: Eu sei, mas você anotou aqui!
89. Z: Eu não anotei nada... Anotei as coisa que ele tinha falado... 40, 40 pessoas, certo... <RELÊ A TABELA> Até onde já deu 20? Deixa eu ver... 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22... Aqui, mais ou menos, deu 20. <PAUSA> 22, certo? Deixa eu ver: 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40. Aqui, ó! É aqui. A mediana. Espero que seja isso. Entre 14 e 16 anos. Não, entre 14 e 15 anos. A mediana está entre 14 e 15 anos. Qual que era a próxima? O desvio. <PAUSA>

Consideramos a possível existência da concepção $\mathbb{C}_3$, caracterizada pelos elementos:

- P – campo de problemas: determinação da mediana com base nos valores apresentados em uma tabela.

- R – operadores: localização das frequências associadas à variável 'idade' em uma tabela.
- L – conjunto de representações: representação numérica e representação tabular.
- Σ – estrutura de controle: reconhecimento da posição da frequência intermediária.

Assim, podemos inferir a existência de uma possível concepção errônea Φ_3 , mobilizada pela aluna Z, deslocada de seu domínio de validade, segundo a qual a mediana não se encontra entre os dois efetivos intermediários extraídos da amostra, organizados em ordem crescente, mas, sim, entre as duas classes intermediárias (a classe que contém o maior efetivo entre os menores e a que contém o menor efetivo entre os maiores).

Após obter as três medidas de tendência central, seguindo a ordem do material didático, e não a do exercício, as alunas D, S e Z trabalharam para encontrar as medidas de dispersão. Assim, a seguir, apresentamos um longo diálogo envolvendo consulta às anotações de sala de aula das estudantes e ao material didático adotado pela escola. Ao encontrar a suposta mediana, Z pergunta para as colegas: “Qual que era a próxima? O desvio?”:

90. S: -4, né?

91. Z: 11 anos, -4; 12, -3; 13, -2; 14 -1; agora... pra cima. 1, 2, 3, 4, 5, 6. Pra cima agora, 21, 6.

Figura 40 – Excerto II do Apêndice 2, Protocolo 1 – Grupo I

Mediana = 14 e 15 anos

Desvio =

	11	12	13	14	15	16	17	18	19	20	21	total
	3	4	4	6	5	8	4	3	2	0	1	40

Fonte: o autor, 2020.

92. S: Vai parecer que é 11 - 4, 12 - 3...

93. Z: Põe como?

94. S: Põe assim, tipo, 'igual', não põe 'menos'.

95. Z: Ué, mas não é igual. 11 não é igual a -4. É o desvio.

96. S: Sim, mas é como se fosse... Mas é resposta, não é?

97. Z: É. Aqui no meu caderno...

98. S: Então, mas coloca igual. Pra não confundir.

99. Z: 11 = -4 é a mediana, 12 = -3 negativo. Agora pra cima, positivo. 21 = 6, seis casas acima, 20 = +5, 19 = +4, 18 = +3, 17 = +2, 16 = 1. Agora a variância. A variância é o que, mesmo?

Na fala de número 105, apresentada mais à frente, elevaram esses desvios ao quadrado.

Apesar de algumas confusões, com avanços e retrocessos, percebemos que as alunas compreenderam o conceito de desvio parcial bem como seu valor absoluto e relativo. Elas entendem que é necessário calcular todos os desvios simples antes de calcular a variância. Além disso, embora registrem os desvios ora com o sinal negativo, ora sem ele, e se esquecem de colocar os parênteses ao efetuar a potenciação, calculam os quadrados de forma correta após algumas correções, como poderemos observar nas falas 124, 125 e 126, logo adiante.

Vemos, também, que elas confundem diversas vezes os conceitos de variância e desvio padrão.

100. S: Pega todos e divide por...

101. Z: Olha aqui, ó. É pra somar.

102. D: Dividir, ele falou que era dividir.

103. Z: Soma todos... eleva ao quadrado.

104. S: Acho que é soma.

105. Z: Eu tinha anotado mais. Eu acho que isso, gente... A variância é a multiplicação da idade com a frequência. Ao quadrado. Mais 40.

Figura 41 – Excerto III do Apêndice 2, Protocolo 1 – Grupo I

$11 = -4^2 = 16$	$21 = 6^2 = 36$
$12 = -3^2 = 9$	$20 = 5^2 = 25$
$13 = -2^2 = 4$	$19 = 4^2 = 16$
$14 = -1^2 = 1$	$18 = 3^2 = 9$
	$17 = 2^2 = 4$
	$16 = 1^2 = 1$
	$15 = 0 = 0$

Fonte: o autor, 2020.

106. D: Pra mim ele falou dividir.

107. Z: Quer ir lá perguntar pra ele? Se vai dividir ou somar por 40 <D SAI DA SALA> <LONGA PAUSA> O desvio padrão é só elevar ao quadrado. Enquanto ela vê a variância, a gente vê o desvio padrão.

108. S: Mas tem que fazer os dois? O desvio e o desvio padrão?

109. Z: É esse aqui, o desvio padrão é ao quadrado. Quem sobra é quem aumenta da mediana. Quem sobra é quem aumenta, que é os positivos

e os negativos. A variância, que são os números ao quadrado mais ou dividido por 40. A gente responde. O desvio padrão... ele é... ele é ao quadrado, da variância. <SUSSURROS> <PAUSA> Fazer os gráficos, enquanto isso.

110. D: <RETORNA> É dividido mesmo, e ela anotou errado.
111. Z: Ah...
112. S: E tem um que tem que fazer a raiz quadrada, não é?
113. D: É essa.
114. Z: A variância?
115. D: Não é? Aqui. Ele falou uma coisa assim... que, tipo, é 5, quando elevar ao quadrado, vai ficar 25. Aí vai ser 25 dividido por 40.
116. Z: Ah! É mesmo! Aqui, ó! Só que não é 5, mano. É 4, porque como tá aqui no meio.
117. S: Aí é 4 dividido por 40?
118. Z: 4 elevado ao quadrado. Vamos elevar logo aqui pra ver o resultado que vai dar.
119. S: Vai devagar, gente...
120. Z: Então vamos voltar para o desvio.
121. D: E ele falou que se tiver empacadinho, não é para ir lá perguntar, pula pra próxima.
122. S: Tá bom. <RISOS>
123. Z: Ó, vou passar aqui.
124. D: 2. É 8, né?
125. Z: 4 é 2^2 . É 4^2 .
126. S: 16.
127. Z: É, $4^2 = 16$. Agora 3...
128. S: 9.
129. Z: 2×2 .
130. S e D: 4.
131. Z: 1×1 .
132. S: 1.
133. Z: 6×6 ?
134. D: 12.
135. S: 36. <RISOS>
136. Z: 5×5 ?
137. S: 25.
138. Z: 4×4 ?
139. S: 16.
140. Z: 3×3 , 9. 2×2 ...
141. D: 4.
142. Z: 4. 1×1 , 1. Esse aqui é o desvio, certo? Agora a variância.
143. S: É, mas não tem que pegar esse resultado e dividir por 40?
144. Z: Soma tudo?
145. S: É, mas não tem que pegar esse resultado e dividir por 40? Tipo, 36 dividido por 40?
146. Z: Nem dá pra dividir por 40.
147. D: Vai dar, tipo, 0,1.
148. S: É, nem dá.
149. Z: Agora a gente vai fazer a variância: o número multiplicado pelo outro ao quadrado, dividido por 40.

150. D: HUUUUUUUMMM!
151. Z: É, professor, no final ela explica tudo. <RISOS> Vai, 11×3 , faz aí.
152. S: $11 \times 3 = 33$, ao quadrado... 33×33
153. D: Dá quebrado, 1,0...
154. S: MIL...
155. Z: 1.089.
156. D: Deu número quebrado. É 1,089.
157. Z: Dá um murro nela...
158. D: <RISOS> Ah, é verdade, é 1.089.
159. Z: 14, oh, 12×4 .
160. D: 48.
161. Z: 48^2 ... 2.304. Vai, 13...
162. S: 13... dividido...
163. D: Vezes 4.
164. Z: 13×4 ...
165. S: 52.
166. Z: 52^2 ... 2.304. 14×6 ...
167. D: 84.
168. Z: 84^2 .
169. D: Não é esse aqui? Não é esse resultado aqui?
170. Z: 75^2 ... 5.625. $128^2 = 16.384$... 17... 68^2 .
171. S: 68?
172. Z: 4.624 ... 17×4 . Quanto é que é?
173. D: 68.
174. Z: 68 dividido por... oh, 68×68 ? Certo? 4.624. 18×3 ... 54... 54^2 ... 2.916... 19×2 , dá 38, 38^2 dá 1.444, 20×0 é 0, 0^2 é 0, 21×1 é 21, 21^2 é 21×21 ...
175. D: 441.
176. Z: 441. Pronto, essa é a variância. Agora vai dividir por... Somar tudo e dividir por... 40? Acho que a gente vai dividir cada uma por 40.
177. D: O resultado, né?
178. Z: Cada um, pra dar o resultado... 1.089 sobre 40.
179. S: O meu deu esse resultado aqui, eu vou pôr sem ponto, gente.
180. D: Ai, deu esse, agora.
181. Z: Não fala "deu esse"! Qual que é? 27.225. 2.304 sobre 40... 57,6. 2.704 sobre 40... 67,6. 7.056 sobre 40?
182. D: 176,4.
183. Z: 5.625 sobre 40.
184. D: 140,62.
185. Z: 16.374 sobre 40... 408,6.
186. Z: 84...
187. D: Aqui deu 409,6.
188. Z: 16.384.
189. S: Eu tinha invertido a ordem.
190. Z: Dividido por 40... 409,6... 4.624 sobre 40... 115,6... 2.916 sobre 40... 72,9... 1.444 sobre 40... 36,1... 441 sobre 40... 11,025. Pronto, a gente achou a variância. Agora é desvio padrão.
191. S: É nesse aí que a gente tira a raiz quadrada, né?
192. D: Tem a fórmula aqui.
193. S: Falei.

194. Z: O desvio padrão é esse.
195. D: O desvio padrão é aquele que tira a raiz quadrada.
196. Z: De 40, né?
197. S: Do número! Do que deu. Do resultado.
198. Z: Que resultado?
199. D: Esse aí, ó!
200. S: É esse resultado aqui. Não dá nem pra tirar a raiz quadrada!
201. Z: Desvio padrão é o da variância. Será que a gente não tinha que somar tudo e dividir por 40?
202. S: Acho que a gente fez alguma coisa errada aí.
203. Z: Acho que a gente não tinha que somar tudo e dividir por 40.
204. D: Qual foi aquela que ele falou assim, que tem que dar uma idade que tá aqui?
205. Z: É a moda. Acho que a gente não tinha que somar tudo e dividir por 40. Não dividir cada um. Porque a variância é uma só, não é 300!
206. D: Olha aqui, uma variância aqui, ó!
207. Z: <LENDO UM EXEMPLO DO LIVRO> A gente fez de cada um. Agora tem que somar tudo e dividir por 40. Vamos pra dois, vai, qual que é? Tem que achar o desvio padrão.
208. S: A gente fez errado!
209. Z: <LENDO UM EXEMPLO DO LIVRO>
210. D: A gente erramos porque esquecemos de colocar o mais.
211. Z: Vamos fazer igual tá aqui.
212. A: A gente tem que fazer com esses números. Não com os números que tão aí.
213. Z: <LENDO UM EXEMPLO DO LIVRO> <REFAZ MULTIPLICANDO CADA IDADE PELA FREQUÊNCIA E ELEVANDO AO QUADRADO> Deu quanto o resultado agora?
214. D: 8.848 sobre 40... 211,2.
215. S: Coloca aproximadamente.
216. Z: Agora, desvio padrão. Raiz quadrada de 211,2... Está entre 14 e 15 anos. Agora vamos fazer o gráfico. Eu vou fazendo enquanto vocês leem a questão.

Essa longa discussão exemplifica bem a importância do trabalho em grupo, como proposto por Garfield (1993, 2013) e Diaz (2016).

Na sessão seguinte, as três alunas retomaram o mesmo exercício e refizeram os cálculos, no entanto, elas elevam ao quadrado as idades dos sujeitos, não os desvios, obtendo os seguintes resultados:

Figura 42 – Excerto IV do Apêndice 2, Protocolo 1 – Grupo I

$$\text{Variancia} =$$

$$V = 8 \cdot (11)^2 + 4 \cdot (12)^2 + 4 \cdot (13)^2 + 6 \cdot (14)^2 + 5 \cdot (15)^2 + 8 \cdot (16)^2 + 4 \cdot (17)^2 + 3 \cdot (18)^2 + 2 \cdot (19)^2 + 0 \cdot (20)^2$$

$$40$$

$$V = 9255 \div 40 = 231,3$$

Fonte: o autor, 2020.

Entre as falas numeradas e apresentadas anteriormente, do número 92 ao 98, o grupo demonstrou algum entendimento sobre desvios. Como haviam calculado a média aritmética das idades com o valor igual a 15 anos, determinaram: "11 = -4, [...] 12 = -3 negativo. Agora pra cima, positivo. 21 = 6, seis casas acima, 20 = +5, 19 = +4, 18 = +3, 17 = +2, 16 = 1." Na fala número 103 e no registro por escrito, observamos que elevaram os desvios ao quadrado, entretanto, sentem-se inseguras e consultam o material didático.

Nas falas 207, 209 e 2013, Z parece dispersar sua atenção e, então, encontramos uma discrepância entre a sua fala e o registro escrito. A interação acontece, mas Z assume a função de escriba e, alheia ao grupo, toma para si a responsabilidade de resolver. Destacamos a fala número 216, "Agora, desvio padrão. Raiz quadrada de 211,2... Está entre 14 e 15 anos. Agora vamos fazer o gráfico. Eu vou fazendo enquanto vocês leem a questão.", em que a frase "Eu vou fazendo enquanto vocês leem" contraria os princípios que norteiam o bom trabalho colaborativo. O registro, como já dissemos, é discrepante, mas como as colegas do grupo não estão acompanhando, não fazem a verificação.

Então, finalmente, o desvio padrão é apresentado:

Figura 43 – Excerto V do Apêndice 2, Protocolo 1 – Grupo I

$$\text{Desvio padrão} = \sqrt{231,3} \approx 15,2 \quad 15 \text{ e } 16 \text{ anos}$$

Fonte: o autor, 2020.

Identificamos, por parte de Z, a mobilização da concepção Φ_4 , caracterizada pelos elementos:

- P – campo de problemas: determinação do desvio padrão com base nos valores apresentados em uma tabela.
- R – operadores: localização das frequências associadas à variável 'idade' em uma tabela, desvio compreendido como a diferença entre um dado efetivo e a média aritmética simples da amostra e variância como desvio quadrático da média.
- L – conjunto de representações: representação numérica e representação tabular.
- Σ – estrutura de controle: controle passo a passo do algoritmo do desvio padrão, comparação dos procedimentos com registros dos cadernos de cada aluna e do material didático adotado pela escola.

Assim, podemos inferir a existência de uma possível concepção Φ_4 mobilizada pela aluna Z, segundo a qual o desvio padrão corresponde à raiz quadrada da variância, muito embora ela tenha, no decorrer dos cálculos, trocado os quadrados dos desvios pelos quadrados das idades dos estudantes fictícios apresentados no problema.

Mais adiante, na sequência do exercício, D, S e Z retomam suas ideias em torno das medidas de posição ao tentarem resolver o item 'c' do exercício proposto:

Figura 44 – Questão c dos exercícios aplicados

- c) Em relação às questões 5 e 6, escolaridade dos pais, os alunos elaboraram a tabela a seguir. A tabela construída simplifica o banco de dados. Como você evidenciaria essa comparação se quisesse utilizar gráficos? Explique seu pensamento. É possível determinar uma medida-resumo que contribua com a análise desses dados? Qual? Como determinar? Explique seu pensamento.

Escolaridade	do pai	da mãe
não alfabetizado	2	1
ens. fundamental incompleto	2	1
ens. fundamental completo	4	6
ens. médio incompleto	4	0
ens. médio completo	9	15
ens. superior incompleto	6	1
ens. superior completo	8	12
pós-graduação incompleta	3	0
pós-graduação completa	2	4
Total	40	40

Fonte: o autor, 2020.

Ao considerar o tipo de variável envolvida, Z não parece encontrar dificuldades para determinar as medidas de tendência central que podem ou não ser obtidas, como podemos ver na fala 289:

289. Z: A moda dá pra achar, a moda é o 'ensino médio completo'. A média não dá pra achar, porque é qualitativa. A mediana dá pra achar também.

Figura 45 – Excerto VI do Apêndice 2, Protocolo 1 – Grupo

3) moda = ~~5~~ Ensino médio completo
 mediana = Esta entre o Ensino médio completo e o Ensino Superior incompleto.

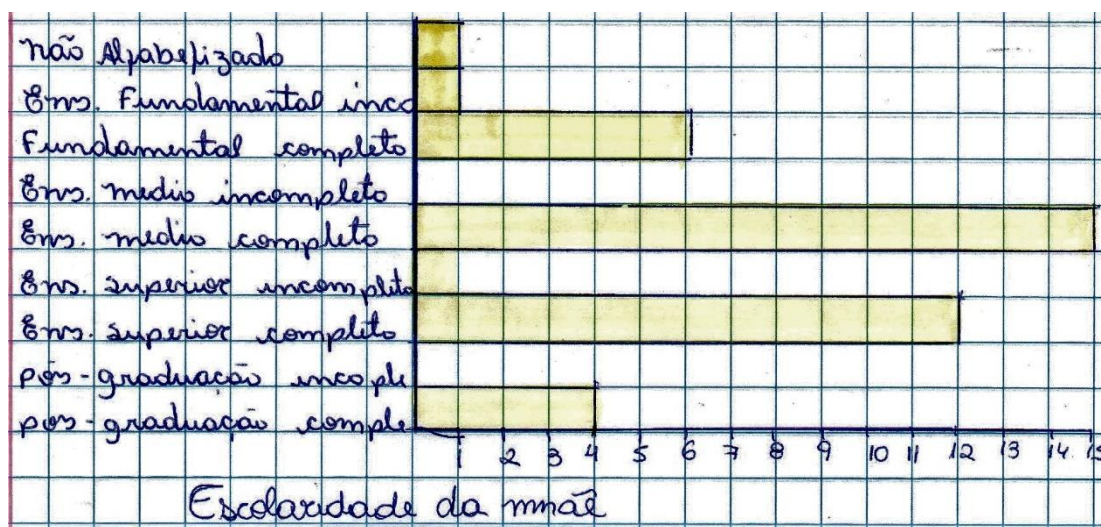
Fonte: o autor, 2020.

Novamente, Z mobiliza a concepção $\mathbb{C}_3$, deslocada de seu domínio de validade, segundo a qual a mediana não se encontra entre os dois efetivos intermediários extraídos da amostra, organizados em ordem crescente, mas entre as duas classes intermediárias. A mediana, tanto para homens quanto para mulheres, corresponde a

possuir o ensino médio completo, enquanto Z afirma que está entre o 'ensino médio completo' e o 'ensino superior'.

Ela também constrói um gráfico de barras horizontais, como solicitado no exercício, evidenciando a moda:

Figura 46 – Excerto VII do Apêndice 2, Protocolo 1 – Grupo I



Fonte: o autor, 2020.

Identificamos, por parte de Z, a mobilização da concepção Φ_5 , caracterizada pelos elementos:

- P – campo de problemas: determinação da possibilidade de obtenção das medidas de tendência central, moda, mediana e média aritmética simples, com base nos valores apresentados em uma tabela.
- R – operadores: identificação da natureza da variável.
- L – conjunto de representações: representação numérica, representação tabular e representação gráfica.
- Σ – estrutura de controle: distinção da natureza da variável como quantitativa (contínua ou discreta) e qualitativa (nominal ou ordinal).

Assim, podemos inferir a existência de uma possível concepção Φ_5 , mobilizada pela aluna Z, segundo a qual é possível determinar moda e mediana, mas não média aritmética, no caso de uma variável qualitativa ordinal. O reconhecimento da natureza da variável exerce o papel de controle nessa concepção.

Não discutiremos as possíveis concepções mobilizadas na construção de gráficos estatísticos pelas alunas do Grupo I, pois elas pouco falaram sobre isso. Os gráficos foram construídos em silêncio, sem grande interação notória.

Passemos, então, para o Grupo II. A dupla foi formada por 'B' (sexo feminino, 17 anos de idade) e 'P' (sexo masculino, 18 anos de idade), colegas de uma mesma turma.

Esta dupla demonstrou não compreender o conceito de variabilidade e, embora tenha determinado as medidas de tendência central, não conseguiu calcular as medidas de dispersão. Segundo Batanero (2001), é impossível realizar um estudo sobre uma distribuição de frequências limitando-se às medidas de tendência central. Um erro frequente é ignorar a dispersão de dados ao comparar duas ou mais amostras ou populações. A média, sem uma medida de dispersão associada, não pode caracterizar uma distribuição de frequência, e uma amostra bastante heterogênea pode apresentar a mesma média que uma amostra homogênea. Compreender essa variabilidade é fundamental para compreender a Estatística.

O conceito de variabilidade está no âmago da Estatística e da Probabilidade. Segundo Moore (1997), ela é onipresente, está em tudo aquilo que nos rodeia e permeia nossas atividades cotidianas, podendo ser facilmente observável, embora nem sempre bem compreendida. Para Gal (2002), o letramento estatístico implica, entre outras coisas, em compreender a variabilidade em diversos contextos.

A dupla de alunos em questão iniciou a atividade com as seguintes falas:

1. B: Variável?
2. P: É.
3. B: É que tem variância...
4. P: É para você explicar o que é variação das idades utilizando isso aqui, construir uma das representações do gráfico usando medidas, moda, média, mediana, quartis, desvio padrão e amplitude... O que que é que é variação?
5. B: Variância e desvio padrão é pra ver, tipo, o quão errado estão os cálculos. Tipo, é como se fosse uma margem de erro. Uma variância é, tipo, um X, e a diferença entre a variância é o desvio padrão. <LEND O LIVRO DIDÁTICO> Há duas medidas desse tipo: a variância e o desvio padrão, que são usadas sempre... Variância é a média dos quadrados dos desvios...

Para esses alunos, no entanto, a variabilidade não parece ser compreendida como um fenômeno natural em nossas vidas, pelo contrário, é vista como medida de

anomalia, associada ao erro, determinando "o quão errado estão os cálculos". Eles estendem esse entendimento aos conceitos de variação, variância e desvio padrão.

Silva (2007) afirma que os conceitos de variação e variabilidade não são completamente compreendidos pelos alunos nem pelos professores. Segundo a autora, no ensino de Estatística, devemos ir além do simples cálculo de medidas e elaboração de gráficos. Silva (ibid., p. 20) considera o termo "variabilidade" como uma característica da entidade que é observável, enquanto o termo "variação" refere-se mais à descrição ou medida dessa característica. Para ela, a variabilidade é uma forma substantiva do adjetivo variável. Assim, devemos aprofundar a discussão sobre o real significado e aplicabilidade das medidas e representações, relacionando tais conceitos. De acordo com Silva (ibid.), a variabilidade se encontra particularmente evidenciada nas medidas de dispersão, como variância e desvio padrão.

Novaes (2011) também identificou a frágil compreensão sobre a variabilidade em seus estudos com professores e alunos, identificando uma concepção deslocada de seu domínio de validade por parte dos estudantes: "C_{E3}: A variabilidade nos dados pode ser caracterizada apenas com o valor mais frequente na distribuição" (ibid., p. 157).

É justo inferir que esses alunos não estavam com a mesma motivação dos demais, uma vez que foi o grupo que terminou mais rapidamente as tarefas e, mesmo dispondo de até três sessões para concluí-la, o fez na primeira. Também não demonstraram a mesma atenção e concentração dos estudantes dos demais grupos, o que pode estar evidenciado nos risos e brincadeiras, como observamos na continuidade do diálogo entre os estudantes:

6. P: Aí tem, tipo, que contar todas essas medidas. <RISOS> Putz!
7. B: <LENDO UM EXEMPLO DO LIVRO DIDÁTICO> Mas por que que está aqui o desvio -2? Tá vendo aqui? Desvio -2!
8. P: É que aqui... Sei lá! Tipo, bota aqui média, moda e mediana, o resto eu não sei fazer, não!
9. B: Eu também não sei fazer. <RISOS>
10. P: É o mais simples, véio!
11. B: No trabalho eu é que fiz o cartaz, e o que mais pegou foi essa parte de variância e desvio padrão.
12. P: Deve ser por aqui, no '16', tipo, esse é a média, a moda e a mediana. Pode ser que tenha a ver com isso. <RISOS>
13. B: A média é difícil, porque é o que menos aparece...
14. P: Como assim? <RISOS>

A confusão entre os conceitos de variável, variação, variabilidade e medidas de dispersão continuam. Podemos atribuir o excesso de risos ao nervosismo dos alunos por saberem que estavam sendo gravados, como eles mesmos chegam a externalizar. No entanto, não houve perseverança para superar essas dificuldades. Vale ressaltar que eles já haviam determinado medidas de dispersão em seus projetos de pesquisa, ou seja, não era uma tarefa totalmente nova para os dois:

21. B: Ah, mas a variável já tá aqui, as idades que os alunos forneceram... Isso aqui já é a variável!
22. P: Então tem que fazer só... É só fazer o gráfico aqui.
23. B: <RELÊ O ENUNCIADO DO PROBLEMA> Dados, tem que fazer os dados.
24. P: É, tem que passar os dados junto com as contas. <RISOS>
25. B: Cê tá rindo de nervoso? <RISOS>
26. P: Tipo isso! <RISOS>
27. B: Hum...
28. P: Aqui já tá falando, tipo, da variação, né? Então, é só passar pro gráfico.
29. B: Mas tem que explicar, né?
30. P: É.

Identificamos, por parte de B, a mobilização da concepção Φ_6 , caracterizada pelos elementos:

- P – campo de problemas: determinação da variação dos dados por meio de medidas de posição e representações gráficas, com base nos valores apresentados em uma tabela.
- R – operadores: identificação da natureza da variável, reconhecimento de sua medida por meio dos valores representados em uma tabela, operacionalização dos conceitos por meio de algoritmos e construção de gráficos e tabelas estatísticas.
- L – conjunto de representações: representação numérica, representação tabular e representação gráfica.
- Σ – estrutura de controle: representação da variabilidade dos dados, apontando tendências centrais e desvios.

Assim, podemos inferir a existência de uma possível concepção Φ_6 , mobilizada pela aluna B, segundo a qual a variação, em uma distribuição de frequências, é um indicador do "quão errado estão os cálculos", determinando "uma margem de erro" deles, o que permite o controle da concepção.

Reconhecemos uma possível concepção sobre a mediana associada à maneira como os dados são representados na tabela, como observamos no diálogo apresentado a seguir:

65. P: Ah... <PAUSA> Não, a mediana também vai ser 16...
66. B: Quando eu fiz o cartaz, uma dessas... eu acho que era a mediana, tinha que ver se era ímpar ou par.
67. P: AI, CALMA AÍ QUE EU FIZ ERRADO! Tem que colocar em ordem crescente. Que bos*! <PALAVRÃO>
68. B: <RISOS> Mas já tá em ordem crescente. Não, peraí... Quando tá em ímpar, pega o do meio...
69. P: É, pega o central. Ou então, pega os dois do meio e divide por 2.
70. B: É.
71. P: 16, vai dar tudo 16. A média é 16, a moda é 16 e a mediana é 16. <RISOS> Na frequência.

Estes seriam os elementos dessa possível concepção:

- P – campo de problemas: determinação da mediana com base nos valores apresentados em uma tabela.
- R – operadores: localização das frequências associadas à variável 'idade' em uma tabela.
- L – conjunto de representações: representação numérica e representação tabular.
- Σ – estrutura de controle: reconhecimento da posição da classe intermediária em uma tabela com um número ímpar de classes.

Assim, podemos inferir a existência de uma possível concepção Φ_7 , mobilizada pelo aluno P, segundo a qual a mediana corresponde ao valor assumido pela classe intermediária em uma tabela com um número ímpar de classes. Isso fica evidenciado por meio do seguinte comentário:

79. P: Aqui tem que ter 11 intervalos... [...]

Também é apontado, pelo registro escrito, onde assinalaram com uma seta a possível mediana, a 6ª classe entre as 11, apresentadas na tabela:

Figura 47 – Excerto do Apêndice 2, Protocolo 2 – Grupo II

a) Com relação à variável "idade", os alunos forneceram as seguintes respostas:

Idade (anos)	11	12	13	14	15	16	17	18	19	20	21	Total
Freq. simples	3	4	4	6	5	8	4	3	2	0	1	40

Fonte: o autor, 2020.

O terceiro grupo foi também formado por dois alunos, denominado Grupo III, e demonstrou um nível de envolvimento maior com a atividade. Ela foi formada por alunas da mesma turma: 'N' (17 anos de idade) e 'L' (dezoito anos de idade).

Na primeira sessão, N trabalhou sozinha, pois L havia faltado, supostamente, por motivos de saúde. Solicitei à N que descrevesse oralmente sua atividade, passo a passo, à medida em que estivesse resolvendo as questões. Pedi, ainda, que ela apresentasse suas soluções parciais para L na sessão seguinte e, se L concordasse com tudo o que havia sido feito, bastava prosseguir e concluir as tarefas; caso L discordasse, deveriam rever as tarefas, fazendo as correções necessárias.

Como se estivesse falando consigo mesma, N iniciou as atividades lendo o enunciado da questão. Em seguida, determinou rapidamente a moda, a mediana e a média aritmética, partindo, então, para as medidas de dispersão. Essa aparente pressa pode ter sido a responsável pela mobilização da concepção de média aritmética simples que veremos a seguir.

2. N: Na questão 'a', primeiro, vou fazer a média, aí eu vou somar todos os valores e dividir pela quantidade... Média... <PAUSA MUITO LONGA>
3. N: Primeiro, eu fiz a média das idades. Daí, eu somei todas as idades e, aí, eu dividi pela quantidade. Total deu 176 dividido por 11, que dá 16. Então, a média, com relação à variável de idade dos alunos, é 16.
4. N: A moda... é... 16 também, porque é o número que mais aparece.
5. N: E a mediana também é 16.
6. N: <RELÊ O ENUNCIADO> Deixa eu fazer a variância, nesse caso... A variância eu vou montar uma tabelinha com todas as idades e vou fazer as idades menos a média. Então, eu pego o valor das idades e subtraio da média.

Identificamos, por parte de N, a mobilização de uma possível concepção Φ_8 , caracterizada pelos elementos:

- P – campo de problemas: determinação das medidas de tendência central, moda, mediana e média aritmética simples, com base nos valores apresentados em uma tabela.

- R – operadores: identificação da natureza da variável, reconhecimento de sua medida por meio dos valores representados em uma tabela e operacionalização dos conceitos por meio de algoritmos das medidas de tendência central.
- L – conjunto de representações: representação numérica e representação tabular.
- Σ – estrutura de controle: controle passo a passo dos algoritmos das respectivas medidas de tendência central.

Assim, podemos inferir a existência de uma possível concepção Φ_8 , mobilizada pela aluna N, segundo a qual a média aritmética simples, calculada com base em uma tabela sem intervalos de classe, é o quociente entre o somatório dos valores de todas as classes pelo número total de classes, no caso, $176 \div 11 = 16$ anos de idade, valor incorreto, mas próximo do esperado, que seria pouco maior que 15 anos de idade.

Figura 48 – Excerto I do Apêndice 2, Protocolo 3 – Grupo III

a) "Idade" variável

Média = $\frac{11+12+13+14+15+16+17+18+19+20+21}{11} = 16$

	Idades	Valor - Média	Quadrado
Média = 16	11	-5	25
Meda = 16	12	-4	16
Mediana = 16	13	-3	9
Variância = 10	14	-2	4
Desvio Padrão = 3,16	15	-1	1
	16	0	0
	17	1	1
	18	2	4
	19	3	9
	20	4	16
	21	5	25

Média = 16

→ Média = $\frac{25+16+9+4+1+0+1+4+9+16+25}{11} = 10$

$\sqrt{10} = 3,16$

• Explicação

É necessário entender que uma variação de dados pode ser explicada através de enunciados. Diante disso, resultados como a meda não nitidamente explicados com a análise dos dados transmitidos pelo texto.

Fonte: o autor, 2020.

Vale a pena ressaltar que N dispendeu algum tempo e esforço, inicialmente, organizando e explorando os dados do problema, como podemos ver, a seguir:

Figura 49 – Excerto II do Apêndice 2, Protocolo 3 – Grupo III

• Atividade de Estatística	
Questão 1	$Q1 = 26 + 31 + P + 11 + 1 + 0 + 1 + P + P + 31 + 26 = 101$
938 alunos	11
417 alunos	(estudantes do Ensino Fundamental)
521 alunos	(Ensino médio)
204 alunos	1.º ano
179 "	2.º ano
138 "	3.º ano
513 mulheres	
Investiguei 40 alunos	
20	do período matutino
20	do período vespertino

Fonte: o autor, 2020.

Tentativas de aproximação aos valores desejados foram realizadas vezes pelo Grupo I e a presença de colegas para discutir os resultados conduziu a uma reflexão sobre os mesmos e a eventuais ajustes. Mais uma vez, reafirmamos a importância do trabalho em grupo, como observado por Batanero e Díaz (2004, 211), Garfield (1993, 2013) e Díaz (2016).

A aluna N se deparou com uma situação um pouco diferente na resolução do item 'b', que denominou como 'questão 2':

Figura 50 – Questão b dos exercícios aplicados

b) Em relação à renda familiar, os alunos obtiveram as seguintes respostas

Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)
1	1108,38	11	4800	21	3000	31	Não sabe
2	1500	12	5100	22	1108,38	32	3328,14
3	2521	13	1108,38	23	Prefere não declarar	33	4623
4	Prefere não declarar	14	1108,38	24	3820	34	4100
5	2800	15	1500	25	2216,76	35	5000
6	1108,38	16	Não sabe	26	2600	36	2400
7	3500	17	2750	27	1200	37	1108,38
8	4714	18	3400	28	1108,38	38	1700
9	1108,38	19	Prefere não declarar	29	30000	39	6330
10	2200	20	1108,38	30	1300	40	Não sabe

Como você explica a variação das rendas declaradas utilizando as ferramentas estatísticas que aprendeu? Construa uma ou mais representações gráficas, explicando-as. Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de rendas familiares do grupo de 40 alunos pesquisados.

Fonte: o autor, 2020.

Quando algum valor da tabela se repetia, N escrevia mais uma vez e o considerava no cálculo, assim como é esperado no cálculo da média aritmética ponderada (somando n vezes ou multiplicando por n , quando havia n repetições). Isso não foi feito na resolução do item 'a'.

Ela também não cometeu esse erro ao determinar a mediana, pois considerou os valores repetidos, o que não fez no primeiro exercício. Uma possível explicação para isso reside na estrutura de ambas as tabelas: na primeira, há uma indicação clara do número de repetições de cada valor assumido pela variável, enquanto na segunda, cada valor repetido era reescrito.

Não podemos afirmar se houve distração no primeiro caso, já que N conhecia o algoritmo correto, ou se houve mudança de concepção do primeiro para o segundo caso, o que, na perspectiva de Balacheff (1995), seria considerado um sinal de aprendizagem. Entretanto, esta hipótese não pode ser descartada.

Figura 51 – Excerto III do Apêndice 2, Protocolo 3 – Grupo III

Questão 2
Renda Familiar
Média aritmética = $1.108,38 + 1500 + 2521 + 0 + 2800 + 1.108,38 + 3500$ $+ 4714 + 1.108,38 + 2200 + 4800 + 5100 + 1.108,38 + 1.108,38 + 1500 + 0 + 2750$ $+ 3400 + 0 + 1.108,38 + 3000 + 1.108,38 + 0 + 3820 + 2.216,76 + 2600$ $+ 1200 + 1.108,38 + 3000 + 1300 + 0 + 3.328,14 + 4623 + 4100 + 5000 + 240$ $+ 1.108,38 + 1700 + 6330 + 0$
40
$M = 116.378,32$
40
$M = 2.909.458$
média = 2.909.458
Meda = 1108,38
Mediana = 0 - 0 - 0 - 0 - 0 - 0 - 1.108,38 - 1.108,38 - 1.108,38 - 1.108,38 - 1.108,38 - 1.108,38 - 1.108,38 - 1.108,38 - 1.108,38 - 1.200 - 1300 - 1500 - 1500 - 1700 - 2200 - 2216,76 - 2400 - 2521 - 2600 - 2750 - 2800 - 3000 - 3328,14 - 3400 - 3500 - 3820 - 4100 - 4623 - 4714 - 4800 - 5000 - 5100 - 6330 - 30000
Mediana = $\frac{1700 + 2200}{2} = \frac{3900}{2} = 1950$

Fonte: o autor, 2020.

No entanto, agora temos uma situação nova: no item 'renda familiar' surgem as respostas 'prefere não declarar' e 'não sabe'. A aluna N poderia calcular as medidas de posição baseada apenas nos valores declarados e comentar o fato, em sua análise. Todavia, ela transforma as respostas de natureza qualitativa em respostas quantitativas, convertendo-as no número zero, o que permitiu o cálculo de maneira algo semelhante ao que já havia feito no exercício anterior e nas atividades do material didático, em sala de aula, nos quais essa situação nunca havia acontecido.

Apresentamos, então, os elementos dessa nova concepção $\mathbb{C}_9$, mobilizada por N, relativa ao cálculo da média aritmética:

- P – campo de problemas: determinação da média aritmética com base nos valores apresentados em uma tabela, porém com alguns dados quantitativos ausentes.
- R – operadores: localização dos valores numéricos referentes às rendas, mas também de respostas que indicam a não declaração das mesmas em uma tabela, associadas aos sujeitos da amostra, bem como a verificação de suas respectivas frequências. Determinação da razão entre o somatório dos valores correspondentes à variável ‘renda familiar’, atribuindo o valor zero às rendas familiares não declaradas, e o número total n de elementos da distribuição de frequências.
- L – conjunto de representações: representação numérica e representação tabular.
- Σ – estrutura de controle: controle passo a passo do algoritmo da média aritmética, com a transformação das respostas não fornecidas em zeros.

Podemos inferir, assim, a existência de uma possível concepção Φ_9 , segundo a qual o valor da média será o quociente entre o somatório dos valores x , assumidos pela variável em análise, e o número total n de elementos da amostra (controle nessa concepção), considerando como zeros todos os valores x não fornecidos (no caso, ‘não quer declarar’ e ‘não sabe’).

Isso nos remete ao modelo de letramento estatístico, apresentados por Gal (2002) – veja figura 1, p. 41. Constatamos que N apresenta dificuldades no domínio dos elementos cognitivos de letramento associados à própria língua materna, bem como aos conhecimentos estatísticos e ao conhecimento de contexto.

Além disso, falando de elementos disposicionais, ela expressa uma crença pessoal, segundo a qual toda a ausência (nesse caso, de resposta de natureza quantitativa) corresponde a um zero. Os motivos para interpretar tal ausência, bem como a autonomia para encontrar uma solução para enfrentar tal situação refletem o nível de criticidade, também presente no modelo de Gal (2002).

Mas não podemos atribuir toda a responsabilidade à aluna. Cabem críticas ao professor, por não proporcionar ao aluno situações de aprendizagem semelhantes, bem como ao material didático de apoio, por oferecer ao aluno sempre exercícios repetitivos, que propõem situações ideais, em detrimento dos problemas realistas e contextualizados.

Em outra de suas tarefas, N demonstra indícios de reconhecimento de uma das propriedades da média aritmética apresentada por Strauss e Bichler (1988): a soma dos desvios a partir da média é zero.

7. N: Com esse valor, eu vou colocar ele ao quadrado, e os resultados ao quadrado eu vou somar e dividir pela quantidade pra ter a variância. Então, vamo lá: 11, 12, 13, 14. <SUSSURROS> Aí meu Deus, é complicado... Idade... <PAUSA MUITO LONGA> Então, vamo lá.
8. N: Então, pra calcular a variância, eu peguei o valor das idades e tirei com o valor da média. Eu não posso 'somar todos os resultados', porque 'vai dar zero'. Então, eu preciso colocar, fazer todos os resultados ao quadrado, somar e, com o resultado ao quadrado, eu vou fazer a média deles, somar todos e dividir por 11, e essa vai ser a variância. <PAUSA MUITO LONGA> <SUSSURROS>

Identificamos, por parte de N, a mobilização de uma possível concepção Φ_{10} , caracterizada pelos elementos:

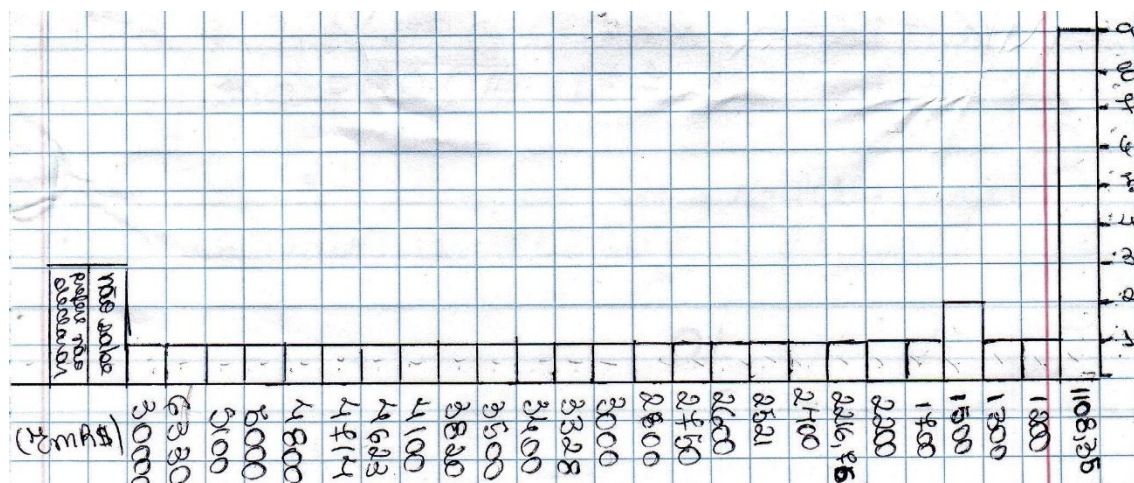
- P – campo de problemas: determinação das medidas de posição com base nos valores apresentados em uma tabela.
- R – operadores: identificação da natureza da variável, reconhecimento de sua medida por meio dos valores representados em uma tabela e operacionalização dos conceitos por meio de algoritmos das medidas de posição.
- L – conjunto de representações: representação numérica e representação tabular.
- Σ – estrutura de controle: controle passo a passo dos algoritmos das respectivas medidas de tendência central.

Dessa forma, podemos inferir a existência de uma possível concepção Φ_{10} , mobilizada pela aluna N, segundo a qual a soma dos desvios a partir da média aritmética simples é nula, o que também anularia, por consequência, a variância e o desvio padrão.

O quarto e último grupo analisado foi formado por três alunos do sexo masculino, que denominamos "A", "V" e "W", com 17 anos de idade completos.

O gráfico construído para o item 'b' chamou nossa atenção:

Figura 52 – Excerto do Apêndice 2, Protocolo 4 – Grupo IV



Fonte: o autor, 2020.

No eixo vertical, a representação foi facilitada pela estrutura quadriculada da folha de papel. Contudo, notamos que não há escala alguma no eixo horizontal. Diversas pesquisas recentes realizadas no Brasil acusaram os equívocos dos estudantes com escalas na construção gráfica, atribuindo-a, em parte, à sua dificuldade em lidar com conversões de unidades de medidas. Evangelista e Guimarães (2015, p. 1) destacam o "desconhecimento dos alunos, que demonstraram dificuldades em representar, localizar, analisar, comparar e construir escalas em gráficos".

Cavalcanti, Natrielli e Guimarães (2010) analisaram diversos gráficos veiculados na mídia impressa e constataram que 39% deles apresentavam escalas desproporcionais, induzindo o leitor a compreensões equivocadas da verdadeira informação. Em consonância com essas ideias, Albuquerque (2010, p. 110) afirma que "cabe à escola trabalhar de forma sistematizada a compreensão de escalas, representadas em gráficos ou não". Segundo a autora (ibid.), a despeito de toda a superexposição de gráficos na mídia, tanto crianças quanto adultos encontram grandes dificuldades na compreensão de escalas.

Embora não tenha sido expressa oralmente, apenas no registro por escrito, podemos inferir a mobilização de uma concepção Φ_{11} , caracterizada pelos seguintes elementos:

- P – campo de problemas: construção de um gráfico estatístico para a representação de uma variável quantitativa contínua com base nos valores apresentados em uma tabela.
- R – operadores: reconhecimento da natureza da variável, identificação dos valores a ela associados, leitura e estruturação de escalas.
- L – conjunto de representações: representação numérica, representação tabular e representação gráfica.
- Σ – estrutura de controle: visualização da proporcionalidade dos valores representativos dos salários e suas respectivas frequências.

Dessa forma, podemos inferir a existência de uma possível concepção Φ_{11} , mobilizada pelo aluno A, segundo a qual não é necessário elaborar uma escala para representar uma variável quantitativa contínua, no caso, o salário (R\$); basta ordená-la, como se fosse uma variável qualitativa ordinal.

Resumimos essas possíveis concepções no Quadro 7, a seguir:

Quadro 7 – Concepções mobilizadas por alunos do último ano do ensino médio

CAMPO DE PROBLEMAS (P)	OPERADORES (R)	SISTEMAS DE REPRESENTAÇÃO (L)	ESTRUTURAS DE CONTROLE (Σ)	CONCEPÇÕES (Φ)
Determinação da média aritmética simples com base nos valores apresentados em uma tabela.	Localização dos valores das idades associadas aos sujeitos da amostra em uma tabela e verificação de suas respectivas frequências. Estabelecimento da razão entre o somatório dos produtos desses dois valores e o número total de elementos da distribuição de frequências.	Representação numérica e representação tabular.	Controle passo a passo do algoritmo da média aritmética simples.	Φ_1 – O valor da média será o quociente entre o somatório dos produtos das frequências f pelos respectivos valores x , assumidos pela variável em análise, e o número total n de elementos da amostra.
Determinação da moda com base nos valores apresentados em uma tabela.	Localização das frequências associadas à variável 'idade' em uma tabela.	Representação numérica e representação tabular.	Comparação dos valores das frequências associadas à variável analisada e identificação do maior valor.	Φ_2 – Se a variável for quantitativa discreta ou qualitativa, a moda é o efetivo associado à frequência máxima.
Determinação da mediana com base nos valores apresentados em uma tabela.	Localização das frequências associadas à variável 'idade' em uma tabela.	Representação numérica e representação tabular.	Reconhecimento da posição da frequência intermediária.	Φ_3 – A mediana não se encontra entre os dois efetivos intermediários extraídos da amostra, organizados em ordem crescente, mas, sim, entre as duas classes intermediárias (a classe que contém o maior efetivo entre os menores e a que contém o menor efetivo entre os maiores).
Determinação do desvio padrão com base nos valores apresentados em uma tabela.	Localização das frequências associadas à variável 'idade' em uma tabela, desvio compreendido como a diferença entre um dado efetivo e a média aritmética simples da amostra e variância como desvio quadrático da média.	Representação numérica e representação tabular.	Controle passo a passo do algoritmo do desvio padrão, comparação dos procedimentos com registros dos cadernos de cada aluna e do material didático adotado pela escola.	Φ_4 – O desvio padrão corresponde à raiz quadrada da variância.

Determinação da possibilidade de obtenção das medidas de tendência central, moda, mediana e média aritmética simples, com base nos valores apresentados em uma tabela.	Identificação da natureza da variável.	Representação numérica, representação tabular e representação gráfica.	Distinção da natureza da variável como quantitativa (contínua ou discreta) e qualitativa (nominal ou ordinal).	Φ_5 – É possível determinar moda e mediana, mas não média aritmética, no caso de uma variável qualitativa ordinal.
Determinação da variação dos dados por meio de medidas de posição e representações gráficas, com base nos valores apresentados em uma tabela.	Identificação da natureza da variável, reconhecimento de sua medida por meio dos valores representados em uma tabela, operacionalização dos conceitos por meio de algoritmos e construção de gráficos e tabelas estatísticas.	Representação numérica, representação tabular e representação gráfica.	Representação da variabilidade dos dados, apontando tendências centrais e desvios.	Φ_6 – A variação, em uma distribuição de frequências, é um indicador do "quão errado estão os cálculos", determinando "uma margem de erro" deles.
Determinação da mediana com base nos valores apresentados em uma tabela.	Localização das frequências associadas à variável 'idade' em uma tabela.	Representação numérica e representação tabular.	Reconhecimento da posição da classe intermediária em uma tabela com um número ímpar de classes.	Φ_7 – A mediana corresponde ao valor assumido pela classe intermediária em uma tabela com um número ímpar de classes.
Determinação das medidas de tendência central, moda, mediana e média aritmética simples, com base nos valores apresentados em uma tabela.	Identificação da natureza da variável, reconhecimento de sua medida por meio dos valores representados em uma tabela e operacionalização dos conceitos por meio de algoritmos das medidas de tendência central.	Representação numérica e representação tabular.	Controle passo a passo dos algoritmos das respectivas medidas de tendência central.	Φ_8 – A média aritmética simples, calculada com base em uma tabela sem intervalos de classe, é o quociente entre o somatório dos valores de todas as classes pelo número total de classes.

Determinação da média aritmética com base nos valores apresentados em uma tabela, porém com alguns dados quantitativos ausentes.	Localização dos valores numéricos referentes às rendas familiares, mas também de respostas que indicam a não declaração das mesmas em uma tabela, associadas aos sujeitos da amostra, bem como a verificação de suas respectivas frequências. Determinação da razão entre o somatório dos valores correspondentes à variável 'renda familiar', atribuindo o valor zero às rendas familiares não declaradas, e o número total n de elementos da distribuição de frequências.	Conjunto de representações: representação numérica e representação tabular.	Controle passo a passo do algoritmo da média aritmética, com a transformação das respostas não fornecidas em zeros.	Φ_9 – O valor da média será o quociente entre o somatório dos valores x, assumidos pela variável em análise, e o número total n de elementos da amostra (controle nessa concepção), considerando como zeros todos os valores x não fornecidos.
Determinação das medidas de posição com base nos valores apresentados em uma tabela.	Identificação da natureza da variável, reconhecimento de sua medida por meio dos valores representados em uma tabela e operacionalização dos conceitos por meio de algoritmos das medidas de posição.	Representação numérica e representação tabular.	Controle passo a passo dos algoritmos das respectivas medidas de tendência central.	Φ_{10} – A soma dos desvios a partir da média aritmética simples é nula, o que também anulária, por consequência, a variância e o desvio padrão.
Construção de um gráfico estatístico para a representação de uma variável quantitativa contínua com base nos valores apresentados em uma tabela.	Reconhecimento da natureza da variável, identificação dos valores a ela associados, leitura e estruturação de escalas.	Representação numérica, representação tabular e representação gráfica.	Visualização da proporcionalidade dos valores representativos dos salários e suas respectivas frequências.	Φ_{11} – Não é necessário elaborar uma escala para representar uma variável quantitativa contínua, basta ordená-la, como se fosse uma variável qualitativa ordinal.

Fonte: o autor, 2020.

Buscamos, neste capítulo, identificar concepções mobilizadas pelos alunos ao resolver problemas de Estatística. Sem julgar nenhuma dessas concepções sob o critério dicotômico de certo ou errado, ao término desta etapa, destacamos alguns aspectos associados a elas:

- a média assumirá o valor do quociente entre o somatório dos produtos das frequências f pelos respectivos valores x , assumidos pela variável em análise, e o número total n de classes da amostra;
- em uma variável qualitativa ou quantitativa discreta, a moda é o efetivo associado à frequência máxima;
- em uma tabela, a mediana está associada à classe intermediária, não à classe que contém o elemento intermediário;
- o desvio padrão corresponde à raiz quadrada da variância;
- para uma variável qualitativa ordinal, é possível determinar moda e mediana, mas não média aritmética;
- a variabilidade é compreendida como indicador da margem de erro de uma variável;
- no cálculo da média aritmética simples, qualquer valor quantitativo não informado será considerado como zero;
- a soma dos desvios a partir da média aritmética simples é nula, o que também anularia, por consequência, a variância e o desvio padrão;
- não é necessário elaborar uma escala para representar uma variável quantitativa contínua, basta ordená-la, como se fosse uma variável qualitativa ordinal.

Ressaltamos o aspecto dinâmico da mobilização das concepções. Os aspectos supracitados aparecem e são transformados ou simplesmente são abandonados pelos alunos em seus processos cognitivos de busca por significados, sob o efeito das estruturas de controle.

Além de submeter suas ideias às suas próprias estruturas de controle, em um trabalho cooperativo, os alunos submetem suas hipóteses aos colegas e, mesmo quando o grupo está próximo de chegar a um consenso, não é raro recorrer ao livro didático como o detentor de uma verdade máxima. Muitas vezes, os estudantes abandonavam os resultados obtidos às custas de muito esforço; tentavam ajustar

suas respostas a exemplos do livro e inseriam seus dados em algoritmos que seriam adequados em outras situações ou, ainda, simplesmente transcreviam trechos do livro.

Nesta seção, inferimos a existência de algumas concepções, baseados nas interações entre os alunos, registradas em nossas gravações bem como em sua produção escrita. Lembramos que, para Balacheff (1995), a concepção é uma estrutura mental atribuída a um sujeito por um observador do seu comportamento, neste caso, o pesquisador. Essa foi, portanto, a nossa leitura de algumas concepções por eles mobilizadas, elaborada com base em nossas observações e, intersubjetivamente, nas contribuições de nossos colaboradores.

A seguir, apresentaremos nossas considerações finais.

7 CONSIDERAÇÕES FINAIS

Iniciamos as investigações apresentadas nesta tese motivados pelos resultados de nossa dissertação de Mestrado, a qual tratou do ensino e aprendizagem de Estatística em uma abordagem por meio de projetos. Além disso, teve papel importante nessa motivação o processo de implementação da BNCC, que criou a unidade de conhecimento "Estatística e Probabilidade" ao longo de toda a educação básica assim como um maior incentivo para a realização, pelos alunos, de pesquisas estatísticas por projetos.

Procuramos, em nossa tese, identificar as concepções mobilizadas por alunos do ensino médio ao resolverem problemas de Estatística, após participarem da realização de pesquisas estatísticas por meio do desenvolvimento de projetos. Adotamos o modelo cK Φ pelo caráter dinâmico e interativo que atribui às concepções, adequando-se satisfatoriamente a nossa perspectiva de trabalho com projetos.

Consideramos este estudo relevante, pela importância da compreensão do papel das concepções no ensino e aprendizagem de Estatística. Além disso, pretendemos contribuir no preenchimento da lacuna existente nessa área de investigação, de acordo com nossa revisão de literatura, uma vez que existem poucas pesquisas realizadas sobre Educação Estatística à luz da Teoria das Concepções.

Para alcançar nossos objetivos, formulamos, em nossa problematização, a seguinte questão de pesquisa: "Quais concepções são mobilizadas por alunos do ensino médio ao resolverem problemas relacionados à Estatística, após o desenvolvimento de projetos nessa área?".

Em nossa revisão de literatura, buscamos outras pesquisas que enfocassem os processos de ensino e/ou de aprendizagem de Estatística no ensino médio em uma abordagem por meio de projetos assim como pesquisas que explorassem o modelo cK Φ na análise das concepções, conhecimentos e conceitos envolvidos nesses processos.

Dada a natureza de nossos objetivos, optamos por realizar um estudo de caso visando conhecer os processos cognitivos envolvidos na aprendizagem de Estatística. Esse método qualitativo de investigação busca uma compreensão profunda sobre o como e os porquês do fenômeno estudado, evidenciando sua identidade e suas características peculiares. Para nortear nossa análise, recorreremos a três aportes

teórico-metodológicos, a saber: Análise Estatística Implicativa – ASI e Análise de Similaridades, Análise Exploratória de Dados – AED e Teoria das Concepções.

Primeiro, para um melhor entendimento sobre as concepções mobilizadas pelos alunos, consideramos necessário identificar o que eles já sabiam sobre Estatística. Assim, optamos pela abordagem metodológica da ASI e Análise de Similaridades, por permitirem, para o estudo de dados multivariados, partir do cruzamento de sujeitos e variáveis (propriedades ou atributos), interpretando os grafos implicativo, coesitivo e de similaridades, com o suporte tecnológico do *software* CHIC.

Em segundo lugar, recorreremos à AED, a fim de compreender a dinâmica construção do conhecimento estatístico em um trabalho colaborativo, em nosso caso, o desenvolvimento de projetos de investigação sobre temas de livre escolha pertencentes ao universo de interesses desses alunos.

Por último, utilizamos a Teoria das Concepções, cerne das nossas investigações, para compreender os processos cognitivos inerentes à mobilização das concepções dos alunos em ação ao resolverem problemas estatísticos, em um momento posterior à apresentação dos resultados finais de seus projetos.

Os problemas estatísticos propostos para os alunos do ensino médio, similares àqueles com os quais se defrontaram ao desenvolverem suas investigações estatísticas em ambiente escolar, envolviam os conteúdos presentes na proposta curricular da rede de ensino que frequentam: medidas de tendência central, medidas de dispersão, tabelas e gráficos estatísticos e, permeando esses objetos de estudo, a onipresente variabilidade. Portanto, esses foram os alvos em nosso estudo do objeto.

Em uma análise documental, observamos a quase ausência de conteúdos da Estatística na rede de ensino paulista, na qual a escola envolvida na pesquisa está inserida, que, por sua vez, reflete as limitações do alcance dos PCN (BRASIL, 1997, 1998, 2000) nessa área. Com a BNCC (BRASIL, 2018), vislumbramos uma possibilidade de mudança desse quadro, muito embora, de acordo com os dados coletados, ela ainda não tenha impactado a formação dos sujeitos de nossas investigações.

Nossa análise, no que diz respeito aos conhecimentos prévios dos alunos, aponta para a necessidade de aprofundamento da abordagem de conteúdos estatísticos (medidas de tendência central, medidas de dispersão, tabelas e gráficos estatísticos), indicando a importância do desenvolvimento do letramento estatístico

dos estudantes, elemento norteador das ações da fase experimental com a resolução de problemas após o desenvolvimento dos projetos pelos alunos.

No entanto, apesar de esse estudo prévio específico de Estatística dentro da disciplina de Matemática estar praticamente ausente, os estudantes apresentaram conhecimentos básicos sobre a ciência, resultado de um possível contato com conceitos estatísticos em outras disciplinas escolares e, até mesmo, de sua presença no dia a dia, nas mais diversas mídias.

Para compreender as ações dos alunos nas atividades subsequentes, foi importante avaliar o que já sabiam sobre Estatística, antes de iniciar o desenvolvimento de projetos.

Constatamos, ao final desta pesquisa, que a metodologia e fundamentação teórica escolhidas proporcionaram condições de responder à questão formulada, uma vez que conseguimos apontar algumas concepções mobilizadas pelos alunos. O modelo cK Φ , desenvolvido por Balacheff (1995, 2001, 2002) para a caracterização de uma concepção, possibilitou, por meio das interações entre os estudantes em atividade de resolução de problemas, a identificação de operadores e estruturas de controle que asseguraram a mobilização da concepção. Dessa maneira, foi possível apontar as validações que permitiram a mobilização de tal concepção.

Dentre as concepções identificadas, algumas trazem operadores e estruturas de controle que: consideram a variabilidade como indicador da margem de erro de uma variável; consideram que, em uma tabela, a mediana está associada à classe intermediária, não à classe que contém o elemento intermediário da distribuição de frequência; em uma variável qualitativa ou quantitativa discreta, a moda existe, é única e corresponde ao efetivo associado à frequência máxima; que ausências de respostas quantitativas podem ser tratadas como zeros nos cálculos de medidas de posição; e não é necessário elaborar uma escala para representar uma variável quantitativa contínua, basta ordená-la, como se fosse uma variável qualitativa ordinal.

Por outro lado, também identificamos concepções mobilizadas associadas a uma estruturação de conhecimentos considerados como corretos. Tais conhecimentos estavam relacionados à estruturação de conceitos considerados igualmente corretos, segundo as normas e regras da Estatística, associados à variabilidade, às medidas de tendência central e de dispersão, aos gráficos e às tabelas estatísticas.

Um dos resultados de nosso trabalho, o qual julgamos ser mais relevante, é de que as estruturas de controle manifestadas por um aluno mobilizam as estruturas cognitivas dos demais, promovendo o aprimoramento de suas ideias e a revisão de suas próprias concepções. Ao compartilhar com os colegas o esboço de um gráfico estatístico, rever as passagens do cálculo de uma medida de dispersão, comparar a determinação de uma medida de tendência central com outra obtida pelo colega por um método diferente e discutir, dentro do grupo, seu entendimento sobre variabilidade, os estudantes reafirmaram ou reviram suas concepções, proporcionando mudanças que podem ser consideradas como indicadores de aprendizagem, dentro do modelo CKC de Balacheff (1995, 2001, 2002).

Outro fator influenciador das decisões dos alunos está associado ao acesso que tiveram ao livro didático adotado pela sua unidade escolar. Por diversas vezes, observamos os estudantes abandonarem resultados e tentarem aproximá-los a exemplos desse livro; inserirem seus dados em algoritmos que seriam considerados corretos em outras circunstâncias, em outros problemas; ou transcreverem trechos do livro. Atribuímos essa conduta ao papel legitimador do livro didático (LAJOLO, 1996), à insegurança de muitos alunos diante de conteúdos estatísticos e matemáticos ou, ainda, às dificuldades relacionadas à leitura e interpretação dos problemas, lembrando que o domínio de linguagem é um dos elementos cognitivos de Gal (2002), necessários ao letramento estatístico.

Reconhecemos que o trabalho cooperativo em pequenos grupos, em atividades de resolução de problemas e na abordagem por meio de projetos, em contextos significativos para os alunos, permite a mobilização de estruturas de controle externas à Matemática e à Estatística, relacionadas ao conhecimento de contexto, fundamentais para o letramento estatístico apresentado por Gal (2002).

Diaz (2016) ressaltou a importância do desenvolvimento de atividades em pequenos grupos, nos momentos de trocar informações e compartilhar dúvidas, comportamentos que, segundo a autora, reduzem a ansiedade e elevam a autoestima e autoconfiança dos alunos, tornando-os mais engajados, responsáveis e produtivos, resultado semelhante àquele encontrado por nós mesmos, anteriormente (GIORDANO, 2016).

Garfield (1993, 2013) destacou a importância de atividades de Estatística em pequenos grupos cooperativos tanto no que diz respeito aos aspectos emocionais/afetivos como nos relacionados à estruturação do grupo, organização das

tarefas e determinação dos papéis de cada integrante, enfatizando que a produção do conhecimento em Estatística é um fenômeno mais coletivo do que individual.

Além disso, um estatístico raramente trabalha só, ele está sempre em interação com profissionais de outras áreas, e a comunicação e relacionamento interpessoal é essencial, transcendendo os aspectos estritamente cognitivos.

A heterogeneidade do grupo de alunos torna a discussão mais rica, na medida em que o capital cultural de cada um pode, em alguns momentos, colaborar de modo singular para a evolução do trabalho. Basta olhar a diversidade de temas escolhidos por eles em suas pesquisas (Quadro 5), para entender essa necessidade de interação.

Um dos objetivos abandonados ao longo de nossas investigações, em função da otimização de tempo e de recursos, foi o de identificar concepções de professores e membros da equipe de gestão escolar envolvidos no gerenciamento e desenvolvimento de projetos de Estatística. Pretendemos retomar esses estudos futuramente.

Consideramos igualmente importante a investigação sobre as possíveis concepções mobilizadas no ensino e na aprendizagem de Estatística no uso regular do livro didático. Destacamos, ainda, a necessidade de um estudo mais aprofundado sobre crenças e afetos ligados à mobilização de concepções, pois, talvez, complementem a perspectivas estritamente cognitivista da Teoria das Concepções.

Respondendo à nossa questão de pesquisa, identificamos onze concepções estatísticas mobilizadas pelos alunos, algumas delas deslocadas de seu domínio de validade. Consideramos que a sua compreensão possa nos auxiliar a diagnosticar tanto dificuldades quanto avanços no processo de aprendizagem, assim como a reconhecer falhas estruturais no ensino, além de apontar possíveis caminhos que assegurem a presença significativa da Estatística no ensino médio brasileiro, em tempos de BNCC (BRASIL, 2018)

REFERÊNCIAS

- ALBANI, P. **Estatística com projetos: uma alternativa de Ensino e de Aprendizagem**. 2015. 115 f. Dissertação (Mestrado Profissional em Matemática) - Universidade Tecnológica Federal do Paraná, Pato Branco, 2015.
- ALBUQUERQUE, M. R. G. C. **Como adultos e crianças compreendem a escala representada em gráficos**. 2010. 124 f. Dissertação (Mestrado em Educação Matemática e Tecnológica) - Universidade Federal de Pernambuco, Recife, 2010.
- ALMOULOUD, S. A. L'analyse statistique de données multidimensionnelles: outil révélateur des conceptions d'enseignants en formation. *In: RENCONTRES INTERNATIONALES A.S.I. ANALYSE STATISTIQUE IMPLICATIVE*, 3., 2005, Palermo, Itália. **Anais eletrônicos** [...], Palermo: Università degli Studi di Palermo, 2005. Disponível em: <http://dipmat.math.unipa.it/~grim/asi/asi_05_saddo_5.pdf>. Acesso em: 27 jun. 2020.
- ALMOULOUD, S. A. **Fundamentos da didática da matemática**. Curitiba: Editora UFPR, 2007.
- ALMOULOUD, S. A. O que está por detrás do CHIC. *In: GRAS, R. et al. O uso do CHIC na formação de Educadores: à guisa de apresentação dos fundamentos e das pesquisas e foco*. Rio de Janeiro: Letra Capital, 2015.
- ALMOULOUD, S. A.; GRAS, R.; RÉGNIER, J. A.S.I. – Análise estatística implicativa: mais uma vez, o que é? **Educação Matemática Pesquisa**, São Paulo, v. 16, n. 3, p. 623-1087, 2014.
- ANDERMAN, L. H.; MIDGLEY, C. **Motivation and Middle School Students**. Champaign: ERIC Digest, 1998.
- ANDRADE, V. L. V. X. **Os conceitos de medidas de tendência central e de dispersão na formação estatística no Ensino Médio no Brasil e na França: abordagem exploratória no quadro da Teoria Antropológica do Didático e da Teoria dos Campos Conceituais**. Tese (Doutorado em Ensino das Ciências e Matemática) - Universidade Federal Rural de Pernambuco, Recife, 2013.
- ARTEAGA, P. et al. El lenguaje de los gráficos estadísticos. **Unión: Revista Iberoamericana de Educación Matemática**, [S. l.], v. 18, p. 93-104, 2009.
- ARTEAGA, P. et al. Las tablas y gráficos estadísticos como objetos culturales. **Números, Revista de Didáctica de las Matemáticas**, [S. l.], v. 76, p. 55-67, mar. 2011.
- ARTIGUE, M. Ingénierie didactique. **Recherches en didactique des mathématiques**, Grenoble, v. 13, n. 1, p. 5-64, 1990.
- AZCÁRATE, P. Los escenarios de aprendizaje. Una estrategia para tratar los conocimientos estocásticos en las aulas. **Didáctica de la Estadística, Probabilidad y Combinatoria**, Granada, n. 2, p. 69-86, 1995.

AZCÁRATE, P. **Estudio de las concepciones disciplinares de futuros profesores de primaria em torno a las nociones de Aleatoriedad y Probabilidad**. Granada: Comares, 1996.

BALACHEFF, N. Conception, connaissance et concept. *In*: GRENIER, D. (Ed.). **Séminaire de l'équipe DidaTech**. Grenoble: IMAG, 1995. p. 219-244.

BALACHEFF, N. Les connaissances, pluralité de conceptionsle cas des mathématiques. *In*: CONFERENCE INGENIERIE DE LA CONNAISSANCE. 2001. Toulouse. **Anais [...]**. Toulouse: [S. n.], 2001, p. 83-90.

BALACHEFF, N. Cadre, registre et conception. **Les cahiers du laboratoire Leibniz**, Grenoble, n. 58, p. 1-19, 2002.

BALACHEFF, N.; GAUDIN, N. Students conceptions: an introduction to a formal characterization. **Les cahiers du laboratoire Leibniz**, [S. l.], n. 65, 2002.

BALACHEFF, N.; MARGOLINAS, C. CK ϕ : modèle de connaissances pour le calcul des situations didactiques. *In*: MERCIER, A; MARGOLINAS, C. (Eds.). **Balises pour la didactique des mathématiques**. Grenoble: La Pensée Sauvage Editions, 2005. p. 75-106.

BARBERINO, M. R. B. **Ensino de Estatística através de Projetos**. 2016. 126 f. Dissertação (Mestrado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2016.

BATANERO, C. Taller sobre análisis exploratorio de datos en la enseñanza secundaria. *In*: CONFERÊNCIA INTERNACIONAL: EXPERIÊNCIAS E EXPECTATIVAS DO ENSINO DE ESTATÍSTICA - DESAFIOS PARA O SÉCULO XXI, 1., 1999, Florianópolis. **Anais [...]**. Florianópolis: Universidade Federal de Santa Catarina, 1999.

BATANERO, C. **Didáctica de la Estadística**. Granada: Departamento de Didáctica de la Matemática de la Universidad de Granada, 2001.

BATANERO, C.; DÍAZ, C. El papel de los proyectos en la enseñanza y aprendizaje de la estadística. *In*: ROYO, J. P. (Ed.). **Aspectos didácticos de las matemáticas**. Zaragoza (España): ICE, 2004. p. 125-164.

BATANERO, C.; DÍAZ, C. **Estadística con proyectos**. Granada: Departamento de Didáctica de la Matemática de la Universidad de Granada, 2011.

BATANERO, C.; DÍAZ, C.; CONTRERAS, J. M.; ROA, R. El sentido estadístico y su desarrollo. **Números, Revista de Didáctica de las Matemáticas**, [S. l.], v. 83, p. 7-18, jul. 2013.

BATANERO, C.; ESTEPA, A.; GODINO, J. D. Análisis exploratorio de datos: sus posibilidades en la enseñanza secundaria. **Suma, Huelva**, n. 9, p. 25-31, set. 1991.

BATANERO, C.; GODINO, J. **Análisis de datos y su didáctica**. Granada: Departamento de Didáctica de la Matemática de la Universidad de Granada, 2001.

BATANERO, C.; GODINO, J.; GREEN, D.; HOLMES, P.; & VALLECILLOS, A. Errores y dificultades en la comprensión de los conceptos estadísticos elementales. **International Journal of Mathematics Education in Science and Technology**, Reino Unido, v. 25, n. 4, p. 527-547, 1994.

BRASIL. **Parâmetros curriculares nacionais**: Matemática (1.º e 2.º ciclos do ensino fundamental). Brasília: MEC/SEF, 1997.

BRASIL. **Parâmetros curriculares nacionais**: Matemática (3.º e 4.º ciclos do ensino fundamental). Brasília: MEC/SEF, 1998.

BRASIL. **Parâmetros curriculares nacionais (ensino médio)**. Brasília: MEC, 2000.

BRASIL. **Base Nacional Comum Curricular - Educação é a Base**. Brasília: MEC, 2018.

BROUSSEAU, G. Fondements et méthodes de la didactique des mathématiques. **Recherches en didactique des mathématiques**, Grenoble, v. 7, n. 2, p. 33-115, 1986.

BROUSSEAU, G. Os diferentes papéis do professor. In: PARRA, C.; SAIZ, I. **Didática da matemática**: reflexões psicopedagógicas. Porto Alegre: Artmed, 2007. p. 48-72.

CAMPOS, C. R. **A educação estatística**: uma investigação acerca dos aspectos relevantes à didática da estatística em cursos de graduação. 2007. 242 f. Tese (Doutorado em Educação Matemática) - Instituto de Geociências e Ciências Exatas, Universidade Estadual Paulista, Rio Claro, 2007.

CAMPOS, C. R.; WODEWOTZKI, M. L. L.; JACOBINI, O. R. **Educação estatística**: teoria e prática em ambientes de modelagem matemática. 2. ed. Belo Horizonte: Autêntica, 2013.

CARVALHO, C. Literacia estatística. In: SEMINÁRIO DE ENSINO DE MATEMÁTICA, 1., 2003, Campinas. **Comunicações apresentadas no Seminário de Ensino de Matemática**. Campinas: [S. n.], 2003. (Mesa-redonda 'Literacia estatística'). Disponível em: <<http://www.miniweb.com.br/Educadores/artigos/pdf/estatica-portugal.pdf>>. Acesso em: 7 maio 2016.

CAVALCANTI, M. R. G.; NATRIELLI, K. R. B.; GUIMARÃES, G. L. Gráficos na mídia impressa. **Bolema**, Rio Claro, v. 23, n. 36, p. 733-751, set. 2010.

CAZORLA, I. M. **A relação entre a habilidade viso-pictórica e o domínio de conceitos estatísticos na leitura de gráficos**. 2002. 335 f. Tese (Doutorado) - Faculdade de Educação, Universidade Estadual de Campinas, Campinas, 2002.

COBELLO, L. S. **Letramento estatístico: análise e reflexões sobre as tarefas contidas no material didático da secretária estadual de educação de São Paulo para o ensino médio**. 2018. 135 f. Dissertação (Mestrado em Ensino de Ciências Exatas) - Universidade Federal de São Carlos, Sorocaba, 2018.

COBO, B. **Significado de las medidas de posición central para lo estudiantes de secundaria**. 2003. 301 f. Tese (Doutorado), Universidad de Granada, Granada, 2003.

COSTA, G. D. F. **A metodologia de projetos como alternativa para ensinar estatística no ensino superior**. 2012. 261 f. Tese (Doutorado) - Faculdade de Educação, Universidade Estadual de Campinas, Campinas, 2012.

COUTANSON, B. **La question de l'éducation statistique et de la formation de l'esprit statistique à l'école primaire en France: Étude exploratoire de quelques caractéristiques de situations inductrices d'un enseignement de la statistique au cycle III**. 2010. 283 f. Tese (Doutorado em Ciência da Educação) - Université Lumière Lyon 2, Lyon, 2010.

COUTINHO, C. Q. S. Educação estatística e os livros didáticos para ensino médio. **Educação Matemática em Foco**, [S. l.], v. 1, p. 68-86, 2013.

COUTINHO, C. Q. S. Transnumeração: o uso do Geogebra na transformação de representações dos dados. **Unión, Revista Iberoamericana de Educación Matemática**, [S. l.], v. 13, n. 49, p. 11-25, abr. 2017.

COUTINHO, C. Q. S.; MIGUEL, M. I. R. Análise Exploratória de Dados: um estudo diagnóstico sobre concepções de professores. *In*: REUNIÃO ANUAL DA ANPED, 30., 2007, Caxambu-MG. **Anais [...]**, Rio de Janeiro: ANPEd, 2007.

COUTINHO, C. Q. S.; SILVA, M. J. F.; ALMOULOU, S. A. Desenvolvimento do pensamento estatístico e sua articulação com a mobilização de registros de representação semiótica. **Bolema**, Rio Claro, v. 24, n. 39, p. 495-514, ago. 2011.

COUTURIER, R.; BODIN, A.; GRAS, R. **A Classificação Hierárquica Implicativa e Coesiva**. 2004. Manual do Curso CHIC - Versão 2.3. Disponível em: <http://math.unipa.it/~grim/asi/asi_03_gras_bodin_cout.pdf>. Acesso em: 06 jan. 2020.

COUTURIER, R.; GRAS, R. CHIC: traitement de données avec l'analyse implicative. *In*: EXTRACTION ET GESTION DES CONNAISSANCES, 2005, Paris. **Anais [...]**. Paris: [S. n.], 2005, p. 679-684.

CRESWELL, J. W. **Projeto de pesquisa: métodos qualitativo, quantitativo e misto**. Porto Alegre: Artmed, 2010.

CURCIO, F. R. **Developing Graph Comprehension: Elementary and Middle School Activities**. Reston: National Council of Teachers of Mathematics, 1989.

DAMM, R. F. Registros de representação. *In*: MACHADO, S. D. A. et al. **Educação Matemática: Uma (nova) introdução**. 3. ed. São Paulo: Educ, 2012. p. 11-48.

DANGIÓ, E. G. Z. **O Ensino de Estatística no Ensino Médio através de projetos**. 2014. 98 f. Dissertação (Mestrado em Matemática) - Universidade Federal de São Carlos, São Carlos, 2014.

DIAZ, D. **Les facteurs influençant la réussite des activités collaboratives médiées par les TICE dans une situation de formation universitaire à la statistique**. 2016. 328 f. Tese (Doutorado em Ciências da Educação) - Université Lumière Lyon 2, Lyon, 2016.

DÍAZ-LEVICOY, D. et al. Lectura de tablas estadísticas por futuras maestras de Educación Infantil. **Educação Matemática Pesquisa**, São Paulo, v. 18, n. 3, 2016.

DUVAL, R. Registros de representação semiótica e funcionamento cognitivo da compreensão em matemática. *In*: MACHADO, S. D. A. (Org.). **Aprendizagem em Matemática: Registros de representação semiótica**. Campinas: Papirus, 2003. p. 11-33.

ESTEPA, A. **Concepciones iniciales sobre la asociación estadística y su evolución como consecuencia de una enseñanza basada en el uso de ordenadores**. 1993. 385 f. Tese (Doutorado em Didática da Matemática) - Universidad de Granada, Granada, 1993.

EVANGELISTA, M. B.; GUIMARÃES, G. L. Escalas representadas em gráficos: Um estudo de intervenção com alunos do 5º ano. **Revista Portuguesa de Educação**, Braga, v. 28, n. 1, p. 117-138, 2015.

FIORENTINI, D.; LORENZATO, S. **Investigação em educação matemática: percursos teóricos e metodológicos**. 2. ed. rev. Campinas: Autores Associados, 2007.

FISCHBEIN, E. Intuition and proof. **For the learning of mathematics**, New Westminster, v. 3, n. 2, p. 9-24, nov. 1982.

FREITAS, J. L. M. Teoria das Situações Didáticas. *In*: MACHADO, S. D. A. (Org.). **Educação matemática: uma (nova) introdução**. 3. ed. São Paulo: Educ, 2012. p. 77-111.

FREITAS, M. M. **Praticando Estatística no Ensino Médio**. 2016. 64 f. Dissertação (Mestrado em Matemática) - Universidade de Brasília, Brasília, 2016.

FRIEL, S. N.; CURCIO, F. R.; BRIGHT, G. W. Making sense of graphs: Critical factors influencing comprehension and instructional implications. **Journal for Research in mathematics Education**, p. 124-158, 2001.

GAL, I. Adults' statistical literacy: Meanings, components, responsibilities. **International statistical review**, v. 70, n. 1, p. 1-25, 2002.

GALICIA, S. A. M. **Comprensión de las medidas de tendencia central en estudiantes mexicanos de educación secundaria y bachillerato**. 2009. 375 f. Tese (Doutorado em Didática da Matemática) - Universidad de Granada, Granada, 2009.

GARFIELD, J. Teaching statistics using small-group cooperative learning. **Journal of Statistics Education**, v. 1, n. 1, p. 1-9, jul. 1993.

GARFIELD, J. Cooperative learning revisited: from an instructional method to a way of life. **Journal of Statistics Education**, v. 21, n. 2, p. 1-8, jul. 2013.

GARFIELD, J.; GAL, I. Teaching and assessing statistical reasoning. *In*: STIFF, L; CURCIO, F. R. **Developing Mathematical Reasoning in Grades K-12**. Reston: National Council of Teachers of Mathematics, 1999. p. 207-219.

GIORDANO, C. C. **O desenvolvimento do letramento estatístico por meio de projetos**: um estudo com alunos do Ensino Médio. 2016. 153 f. Dissertação (Mestrado em Educação Matemática) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2016.

GOULART, A. **O discurso sobre os conceitos probabilísticos para a escola básica**. 2007. 88 f. Dissertação (Mestrado em Educação Matemática) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2007.

GOULART, A. **Um estudo sobre a abordagem dos conteúdos estatísticos em cursos de Licenciatura em Matemática**: uma proposta sob a ótica da ecologia do didático. 2015. 167 f. Tese (Doutorado em Educação Matemática) - Pontifícia Universidade Católica de São Paulo. São Paulo, 2015.

GRAS R.; ALMOULOU S. A. A implicação estatística usada como ferramenta em um exemplo de análise de dados multidimensionais. *In*: COLÓQUIO "O MÉTODO ESTATÍSTICO IMPLICATIVO UTILIZADO EM ESTUDOS QUALITATIVOS DE REGRAS DE ASSOCIAÇÃO CONTRIBUIÇÃO À PESQUISA EM EDUCAÇÃO". 1., 2003. **Anais** [...]. São Paulo: [S. n.], 2003.

GRAS, R. ; REGNIER, J. C. ; MARINICA, C.; GUILLET, F. (Eds.). **L'analyse statistique implicative Méthode exploratoire et confirmatoire à la recherche de causalités**. Toulouse: Cépaduès Éditions, 2013.

GRAS, R. **Uso do CHIC na Formação de Educadores**: à guisa de apresentação dos fundamentos e das pesquisas e foco. Rio de Janeiro: Letra Capital, 2015.

KADER, G. D.; PERRY, M. A framework for teaching statistics within the K-12 Mathematics curriculum. *In*: INTERNATIONAL CONFERENCE ON TEACHING STATISTICS. 7., 2006, Salvador. **Anais** [...]. Salvador: International Association for Statistical Education and International Statistical Institute, 2006. [CD-ROM].

KENDALL, M.G.; YULE, G. U. **Introdução à Teoria Estatística**. Rio de Janeiro: IBGE, 1948.

LAJOLO, M. Livro didático: um (quase) manual de usuário. **Em aberto**, Brasília, v. 16, n. 69, p. 3-9, jan./mar. 1996.

LANGFORD, E. Quartiles in Elementary Statistics. **Journal of Statistics Education**, [S. l.], v. 14, n. 3, nov. 2006.

LERMAN, I. C.; GRAS, R.; ROSTAM, H. Élaboration et évaluation d'un indice d'implication pour des données binaires. 2. **Mathématiques et sciences humaines**, Paris, v. 75, p. 5-47, 1981.

LIMA, E. L. et al. **A matemática do ensino médio**. 6. ed. Rio de Janeiro: SBM, 2006. v. 2.

LIMA, I. **De la modélisation de connaissances des élèves aux décisions didactiques des professeurs**: étude didactique dans le cas de la symétrie orthogonale. 2006. 390 f. Tese (Doutorado em Didática da Matemática) - Université Joseph Fourier – Grenoble 1, Grenoble, 2006.

LIMA, I. Concepções de alunos do Ensino Fundamental na resolução de problemas de simetria de reflexão. *In*: FIORENTINI, D. **Educação Matemática na Atualidade**. Passo Fundo: Editora da UPF, 2008. p. 1-12. (Anais da II Jornada de Educação Matemática)

LOPES, C. E.; COUTINHO, C. Q. S.; ALMOULOU, S. A. **Estudos e Reflexões em Educação Estatística**. Campinas: Mercado das letras, 2010.

LOPES, C. E.; MEIRELLES, E. O Desenvolvimento da Probabilidade e da Estatística. **VIII Encontro Regional de Professores de Matemática LEM/IMECC/UNICAMP**. Campinas, 2005.

LUMSDEN, L. S. Student Motivation. **Research Roundup**, Alexandria (EUA), v. 10, n. 3, 1994.

MARQUES, M.; GUIMARÃES, G.; GITIRANA, V. Compreensões de alunos e professores sobre média aritmética. **Bolema**, Rio Claro, v. 24, n. 40, p. 725-745, dez. 2011.

MELO, D. M. B. de. **A simetria de reflexão**: elementos de concepções mobilizadas por alunos do Ensino Fundamental. 2010. 112 f. Dissertação (Mestrado em Educação Matemática e Tecnológica) - Universidade Federal de Pernambuco, Recife, 2010.

MELO, D. M. B. de; LIMA, I. M. A simetria de reflexão: concepções mobilizadas por alunos brasileiros. *In*: CONFERÊNCIA INTERAMERICANA DE EDUCAÇÃO MATEMÁTICA. 13., 2011, Recife. **Anais [...]**. Recife: UFPE, 2011.

MENDONÇA, L. O. **A educação estatística em um ambiente de modelagem matemática no ensino médio**. 2008. 236 f. Dissertação (Mestrado em Ensino de Ciências e Matemática) - Universidade Cruzeiro do Sul, São Paulo, 2008.

MIYAKAWA, T. **Une étude du rapport entre connaissance et preuve: le cas de la notion de symétrie orthogonale.** 2005. 304 f. Tese (Doutorado em Didática da Matemática) - Université Joseph Fourier – Grenoble 1, Grenoble, 2005.

MONTEIRO, M. A. Medidas estatísticas de posição com dados agrupados. **Revista do Professor de Matemática**, São Paulo, n. 85, 2014.

MOORE, D. S. **The basic practice of statistics.** Nova Iorque: Freeman, 1995.

MOORE, D. S. New Pedagogy and New Content: The Case of Statistics. **International Statistical Review**, Haia, v. 65, n. 2, p. 123-137, 1997.

MORETTIN, P. A.; BUSSAB, W. O. **Estatística Básica.** 8. ed. São Paulo: Saraiva, 2013.

NOVAES, D. V. **Concepções de professores da educação básica sobre variabilidade estatística.** 2011. 209 f. Tese (Doutorado em Educação Matemática) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2011.

NOVAES, D. V.; COUTINHO, C. Q. S. **Estatística para a educação profissional.** São Paulo: Atlas, 2009.

NOVAES, D. V.; COUTINHO, C. Q. S. Estudo do Ecossistema formado pelos saberes necessários para a compreensão da variabilidade. *In: CONFERÊNCIA INTERAMERICANA DE EDUCAÇÃO MATEMÁTICA.* 13., 2011, Recife. **Anais [...].** Recife: UFPE, 2011.

OLIVEIRA, A. F. **Práticas pedagógicas no Ensino Médio: por uma Estatística crítica e contextualizada.** 2019. 242 f. Dissertação (Mestrado em Educação em Ciências e Matemática) - Universidade Federal de Goiás, Goiânia, 2019.

OLIVEIRA, P. G. **Probabilidade: Concepções construídas e mobilizadas por alunos do Ensino Médio à luz da Teoria das Concepções (cK Φ).** 2010. 196 f. Dissertação (Mestrado em Educação Matemática) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2010.

OLIVEIRA, P. G.; COUTINHO, C. Q. S. Concepções probabilísticas à luz da Teoria cK Φ . *In: CONFERÊNCIA INTERAMERICANA DE EDUCAÇÃO MATEMÁTICA.* 13., 2011, Recife. **Anais [...].** Recife: UFPE, 2011.

ORIOL, J. C. **Formation à la statistique par la pratique d'enquêtes par questionnaires et la simulation: étude didactique d'une expérience d'enseignement dans un département d'IUT.** 2007. 277 f. Tese (Doutorado em Ciência da Educação) - Université Lumière Lyon 2, Lyon, 2007.

ORÚS, P.; ZAMORA, L.; GREGORI, P. (Eds.). **Teoría y aplicaciones del Análisis Estadístico Implicativo: primera aproximación en lengua hispana.** Castelló de la Plana: Innovació Digital Castelló, 2009.

PAGAN, A. et al. A leitura e interpretação de gráficos e tabelas no ensino fundamental e médio. *In: SIMPÓSIO INTERNACIONAL DE PESQUISA EM EDUCAÇÃO MATEMÁTICA*. 2., 2008, Recife. **Anais [...]**. Recife: [S. n.], 2008. Universidade Federal Rural de Pernambuco, 2008. p. 01-10. (CD-ROM). ISBN: 978-85-87459-81-7.

PALLAUTA, J. D.; GEA, M. M. S.; VENEGAS, A. G. Las actividades sobre tablas estadísticas en textos escolares chilenos de educación básica. *In: CONGRESO INTERNACIONAL VIRTUAL DE EDUCACIÓN ESTADÍSTICA*. 3., 2019. **Anais [...]**. [S. l.], [S. n.], 2019.

PFANNKUCH, M.; BEN-ZVI, D. Developing teachers' statistical thinking. *In: BATANERO, C.; BURRILL, C.; READING, C. (Eds.). Teaching Statistics in School Mathematics: Challenges for Teaching and Teacher Education*. Dordrecht: Springer, 2011. p. 323-333. A Joint ICMI/IASE Study: The 18th ICMI Study.

PONTE, J. P. Estudos de caso em Educação Matemática. **Bolema**, Rio Claro, v. 19, n. 25, p. 105-132, 2006.

PORCIÚNCULA, M.; SAMÁ, S. Teaching Statistics Through Learning Projects. **Statistics Education Research Journal** [S. l.], v. 13, n. 2, p. 177-186, nov. 2014.

PORCIÚNCULA, M.; SAMÁ, S. Projetos de aprendizagem: uma proposta pedagógica para a sala de aula de estatística. *In: SAMÁ, S.; PORCIÚNCULA, M. (Orgs.). Educação estatística: ações e estratégias pedagógicas no ensino básico e superior*. Curitiba: CRV, 2015.

PRADO, M. O uso do CHIC na análise de registros textuais em ambiente virtual de formação de professores. **Educação Matemática Pesquisa**, São Paulo, v. 4, n. 2, p. 103-123, 2002.

QUIRÓS, M. V. **El aprendizaje estadístico en la educación secundaria obligatoria a través de una metodología por proyectos**: estudio de caso en un aula inclusiva. 2012. 621 f. Tese (Doutorado em Didática da Matemática) - Universidad de Granada, Granada, 2012.

READING, C.; SHAUGHNESSY, J. M. Reasoning about variation. *In: The challenge of developing statistical literacy, reasoning and thinking*. Dordrecht: Springer, 2004. p. 201-226.

RÉGNIER, J. C. **Histogramme**: Enseigner la Statistique du CM à la Seconde Pourquoi? Comment? Villeurbanne: IREM / Université Claude Bernard, 1998. p. 21-42.

RÉGNIER, J. C. **Vocabulaire de Statistique**. Lyon: ISPEF – Université Lumière Lyon 2, 2007.

RIDGWAY, J.; NICHOLSON, J.; McCUSKER, S. Reconceptualising 'statistics' and 'education'. Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study, p. 311-322, 2008.

SAMÁ, S.; PORCIÚNCULA, M.; SILVA, J. A. Modelo Pedagógico Relacional na Educação Estatística. *In*: CONFERÊNCIA INTERAMERICANA DE EDUCAÇÃO MATEMÁTICA. 13., 2011, Recife. **Anais [...]**. Recife: UFPE, 2011.

SANGARÉ, M. **La rotation**: Approche cognitive et didactique – une étude de cas au Mali. 2000. 342 f. Tese (Doutorado em Didática da Matemática) - Université du Mali, Bamaco, 2000.

SANTANA, M. S. **A educação estatística com base num ciclo investigativo**: um estudo do desenvolvimento do letramento estatístico de estudantes de uma turma do 3.º ano do Ensino Médio. 2011. 197 f. Dissertação (Mestrado em Educação Matemática) - Universidade Federal de Ouro Preto, Ouro Preto, 2011.

SÃO PAULO. **Currículo do Estado de São Paulo - Matemática e suas tecnologias**: Ensino Fundamental ciclo II e Ensino Médio. São Paulo: SE/CENP, 2012.

SÃO PAULO. **Proposta curricular**: Caderno do Aluno – Ensino Médio: Matemática. São Paulo: IMESP, 2014. v. 2.

SCHREIBER, K. P.; BRAZ, G.; BERTOLUCCI, C. C.; SILVA, J. A.; PORCIÚNCULA, M. Níveis de Compreensão do Conceito de Média Aritmética de Adolescentes a partir do Método Clínico-Crítico Piagetiano. **Bolema**, Rio Claro, v. 33, n. 64, p. 491-512, mai./ago. 2019

SILVA, C. B. da. **Pensamento estatístico e raciocínio sobre variação**: um estudo com professores de Matemática. 2007. 354 f. Tese (Doutorado em Educação Matemática) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2007.

SILVA, C. B. da; KATAOKA, V. Y.; CAZORLA, I. M. Linguagem, Estratégia e Nível de Raciocínio de Variação dos Alunos do Ensino Fundamental II. **Bolema**, v. 24, n. 39, ago. 2011.

STRAUSS, S.; BICHLER, E. The development of children's concepts of the arithmetic average. **Journal for Research in Mathematics Education**, Reston, v. 19, n. 1, p. 64-80, jan. 1988.

TOMAZ, V. S.; DAVID, M. M. M. S. **Interdisciplinaridade e aprendizagem da matemática em sala de aula**. Belo Horizonte: Autêntica, 2012.

TUKEY, J. W. The future of data analysis. **Annals of Mathematical Statistics**, [S. l.], v. 33, n. 1, p. 1-67, mar. 1962.

TUKEY, J. W. **Exploratory data analysis**. Reading: Addison-Wesley, 1970.

VADCARD, L. **Etude de la notion d'angle sous le point de vue des conceptions**. 2000. 446 f. Tese (Doutorado em Didática da Matemática) - Université Joseph Fourier – Grenoble 1, Grenoble, 2000.

VERGNAUD, G. La théorie des champs conceptuels. **Recherche en didactique des mathématiques**. Vol 10 no 2. La Pensée Sauvage, Grenoble, 1991.

VERGNAUD, G. Repères pour une théorie psychologique de la connaissance. *In*: MECIER, A.; MARGOLINAS, C. (Coord.). **Balises en didactique des mathématiques**: Cours de la XIIe École d'été de didactique de mathématiques, 2003. Grenoble: La Pensée Sauvage Editions, 2005. p. 123-136.

VIALI, L. **Estatística Descritiva**. [S. l.]: UFRGS, 2016. (Série Estatística Básica). Material didático. Disponível em <http://www.mat.ufrgs.br/~viali/sociais/mat02214/material/apostilas/EDSociais.pdf>. Acesso em: 26 jun. 2020.

VIEIRA, M. **Análise exploratória de dados**: uma abordagem com alunos do Ensino Médio. 2008. 184 f. Dissertação (Mestrado em Educação Matemática) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2008.

WAINER, H. A study of display methods for NAEP results: I. Tables. **ETS Research Report Series**, v. 1995, n. 1, p. i-47, mar. 1995.

WILD, C. J.; PFANNKUCH, M. Statistical Thinking in Empirical Enquiry. **International Statistical Review**, Haia, v. 67, n. 3, p. 223-248, dez. 1999.

YIN, R. K. **Estudo de Caso**: Planejamento e métodos. Porto Alegre: Bookman Editora, 2015.

APÊNDICE 1 – Questionário

Disciplina: Matemática		2º bim/2018	Data: / /
Nome:		Nº	Ano/Turma:
Objetivos: Avaliar conhecimentos, habilidades e competências envolvendo Estatística e Probabilidade			

1 - O que é Estatística? Ela é uma ciência exata ou uma ciência social? É uma área da Matemática?

2 - Você já estudou Estatística no Ensino Fundamental? Em que ano?

() 1º () 2º () 3º () 4º () 5º () 6º () 7º () 8º () 9º () não lembra
() nunca estudou

3 - Você já estudou Estatística no Ensino Médio? Em que ano?

() 1º () 2º () 3º () não lembra () nunca estudou

4 - Você já estudou Estatística em outro curso, fora do ensino regular?

() Não () Sim, estudei no(s) curso(s):

5 - Classifique o seu nível de conhecimentos em Estatística.

1 2 3 4

Ausência de conhecimento ☐ ☐ ☐ ☐ Domínio perfeito

6 - A Estatística é importante para compreender as notícias do rádio, da tv, da internet, dos jornais?

1 2 3 4

Discordo totalmente ☐ ☐ ☐ ☐ Concordo totalmente

7 - Você concorda que a Estatística é importante para tomar decisões em seu dia a dia?

1 2 3 4

Discordo totalmente ☐ ☐ ☐ ☐ Concordo totalmente

8 - Você acredita que irá utilizar a Estatística no(s) curso(s) superior(es) que pretende fazer?

1 2 3 4

Não acredito ☐ ☐ ☐ ☐ Acredito totalmente

9 - Você acredita que Estatística é uma área da Matemática, portando uma ciência exata?

1 2 3 4

Não acredito ☐ ☐ ☐ ☐ Acredito totalmente

10 – Você confia em pesquisas estatísticas?

1 2 3 4

Não confio ☐ ☐ ☐ ☐ Confio totalmente

11 – Você sabe como são realizadas as pesquisas estatísticas?

1 2 3 4

Não sei ☐ ☐ ☐ ☐ Sei muito bem

12 - Você já estudou Probabilidade no Ensino Fundamental? Em que ano?

() 1º () 2º () 3º () 4º () 5º () 6º () 7º () 8º () 9º () não lembra
() nunca estudou

13 - Você já estudou Probabilidade no Ensino Médio? Em que ano?

() 1º () 2º () 3º () não lembra () nunca estudou

14 - Você já estudou Probabilidade em outro curso, fora do ensino regular?

() Não () Sim, estudei no(s) curso(s):

15 - Classifique seu nível de conhecimentos sobre Probabilidade:

1 2 3 4

Ausência de conhecimento ☐ ☐ ☐ ☐ Domínio perfeito

Nas afirmativas a seguir, indique seu nível de concordância, seguindo a escada abaixo:

1. Discordo totalmente

2- Discordo parcialmente

3- Concordo parcialmente

4- Concordo plenamente

16 - Em um lançamento de dois dados cúbicos comuns, numerados de 1 a 6, não viciados, as chances de obtermos soma dos valores das duas faces voltadas para cima igual a 9 é a mesma da soma 10? _____

17 - Em um sorteio da Mega-Sena, as chances de alguém ganhar com a aposta: 1, 2, 3, 4, 5, 6 é menor do que as chances de ganhar com a aposta: 5, 12, 23, 38, 45, 56? _____

18 - No caso anterior, essa seria uma boa aposta, pois sendo um resultado incomum, se alguém ganhar, dividirá o prêmio com um número menor de apostadores, concorda? _____

19 - Um raio não cai duas vezes no mesmo lugar? _____

20 - A chance de uma pessoa morrer ao pular do 12º andar de um prédio diretamente para o asfalto é o dobro da chance de morrer se pular do 6º andar? _____

21 - Quando nossa fila no caixa de um supermercado ou em um pedágio de estrada fica muito lenta é sempre uma boa escolha mudar para a fila do lado, que está mais rápida? _____

22 - Embaralhei as cartas de um baralho tradicional e retirei ao acaso um valete de copas, recolocando-o no monte em seguida. Em uma segunda retirada, a chance de sair outro valete de copas é menor do que a de retirar qualquer outra carta? _____

23 - No sábado, na previsão do tempo, foi anunciado que a probabilidade de não haver chuvas na cidade de São Paulo no dia seguinte era de apenas 90%. No entanto, choveu no domingo. A previsão do tempo estava errada? _____

24 - Valquíria deseja emplacar seu carro com as iniciais de seu nome e os dias de aniversário dela (22) e da filha (13), mas ficou sabendo que a placa VAL2213 é rara, sendo essa opção mais difícil do que a maioria das outras. Você concorda com essa informação? _____

ESTATÍSTICA NO ENEM

25 - (ENEM/2012) A tabela a seguir mostra a evolução da receita bruta anual nos três últimos anos de cinco microempresas (ME) que se encontram à venda.

ME	2009 (em milhares de reais)	2010 (em milhares de reais)	2011 (em milhares de reais)
<i>Alfinetes V</i>	200	220	240
<i>Balas W</i>	200	230	200
<i>Chocolates X</i>	250	210	215
<i>Pizzaria Y</i>	230	230	230
<i>Tecelagem Z</i>	160	210	245

Fonte: Dados fictícios

Um investidor deseja comprar duas das empresas listadas na tabela. Para tal, ele calcula a média da receita bruta anual dos últimos três anos (de 2009 até 2011) e escolhe as duas empresas de maior média anual. As empresas que este investidor escolhe comprar são:

- a) Balas W e Pizzaria Y.
- b) Chocolates X e Tecelagem Z.
- c) Pizzaria Y e Alfinetes V.
- d) Pizzaria Y e Chocolates X.
- e) Tecelagem Z e Alfinetes V.

26 - (ENEM/2010) O quadro seguinte mostra o desempenho de um time de futebol no último campeonato. A coluna da esquerda mostra o número de gols marcados e a coluna da direita informa em quantos jogos o time marcou aquele número de gols.

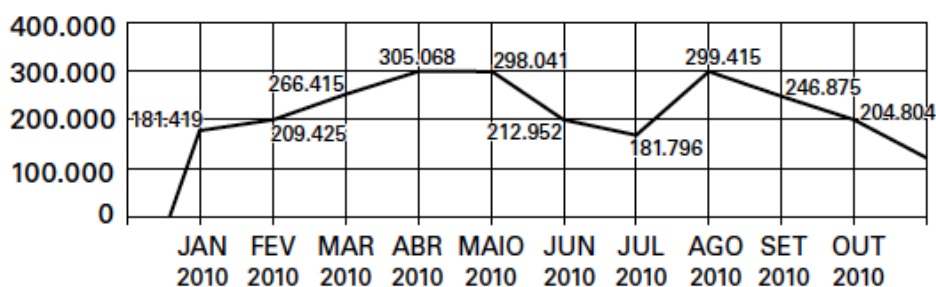
Gols marcados	Quantidade de partidas
0	5
1	3
2	4
3	3
4	2
5	2
7	1

Fonte: Dados fictícios

Se X, Y e Z são, respectivamente, a média, a mediana e a moda desta distribuição, então:

- a) $X = Y < Z$.
- b) $Z < X = Y$.
- c) $Y < Z < X$.
- d) $Z < X < Y$.
- e) $Z < Y < X$.

27 - (ENEM/2012) O gráfico apresenta o comportamento de emprego formal surgido, segundo o CAGED, no período de janeiro de 2010 a outubro de 2010.



Disponível em: www.mte.gov.br. Acesso em: 28 fev. 2012 (adaptado).

Pelo gráfico, o valor da parte inteira da mediana dos empregos formais surgidos no período é:

- a) 212.952.
- b) 229.913.
- c) 240.621.
- d) 255.496.
- e) 298.041.

28 - (ENEM/2010) Marco e Paulo foram classificados em um concurso. Para classificação no concurso, o candidato deveria obter média aritmética na pontuação igual ou superior a 14. Em caso de empate na média, o desempate seria em favor da pontuação mais regular. No quadro a seguir são apresentados os pontos obtidos nas provas de Matemática, Português e Conhecimentos Gerais, a média, a mediana e o desvio padrão dos dois candidatos.

Dados dos candidatos no concurso

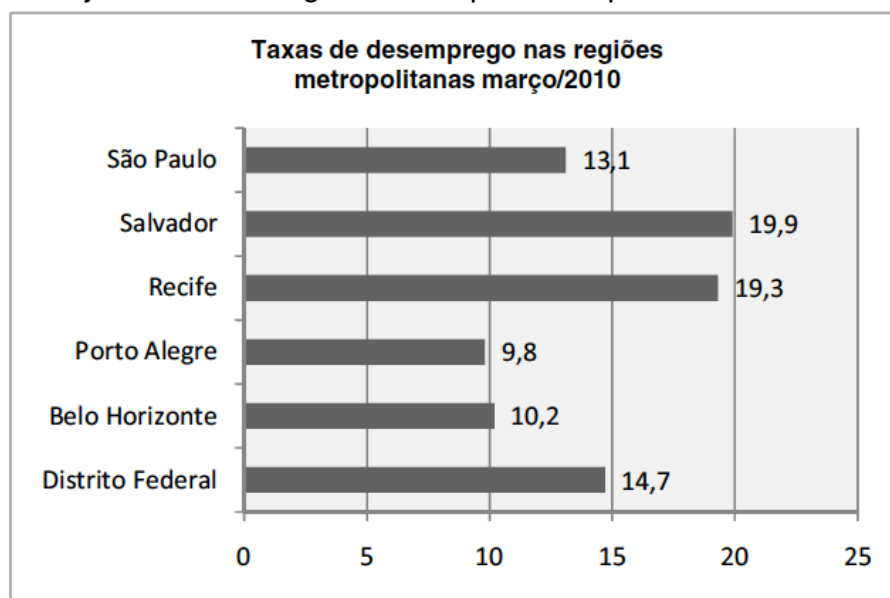
	Matemática	Português	Conhecimentos Gerais	Média	Mediana	Desvio Padrão
Marco	14	15	16	15	15	0,32
Paulo	8	19	18	15	18	4,97

Fonte: Dados fictícios

O candidato com pontuação mais regular, portanto mais bem classificado no concurso, é

- a) Marco, pois a média e a mediana são iguais.
- b) Marco, pois obteve menor desvio padrão.
- c) Paulo, pois obteve a maior pontuação da tabela, 19 em Português.
- d) Paulo, pois obteve maior mediana.
- e) Paulo, pois obteve maior desvio padrão.

29 - (ENEM/2010) Os dados do gráfico seguinte foram gerados a partir de dados colhidos no conjunto de seis regiões metropolitanas pelo Dieese.



Disponível em: <http://g1.globo.com>. Acesso em: 28 abr. 2010 (adaptado).

Supondo que o total de pessoas pesquisadas na região metropolitana de Porto Alegre equivale a 250 000, o número de desempregados em março de 2010, nessa região, foi:

- a) 24 500.
- b) 25 000.
- c) 220 500.
- d) 223 000.
- e) 227 500

APÊNDICE 2 – Protocolos de quatro grupos de alunos

Protocolo 1 – Grupo I

D: SEXO FEMININO, 17 ANOS
S: SEXO FEMININO, 19 ANOS
Z: SEXO FEMININO, 17 ANOS

PRIMEIRA SESSÃO [01:26:58] – 29/11/2018	
Produção oral	Produção escrita
APÓS A LEITURA DO EXERCÍCIO E DO ENUNCIADO DO ITEM A 1. S e Z: <LEEM O ENUNCIADO GERAL> 2. Z: <LÊ O ENUNCIADO DO ITEM A> 3. Z: Vamos achar primeiro a média, né? 4. S: Mas tem que fazer esse aqui... 5. Z: Não, é 'a', tem que fazer primeiro a 'a'. 6. S: Tá. 7. Z: Tá, primeiro a gente vai achar a média. Que que é a média? 8. S: Da idade. A média é o que tem mais? 9. Z: Não, essa aí é a moda. A média é que tem que pegar o número, fazer vezes e depois dividir, não é? 10. Z: Tá, primeiro a gente vai achar a moda. A moda é o que tem mais. 11. S: Moda... 16 anos, não é? 12. Z: É. 13. S: Não, mas tinha que achar esse primeiro, não é? 14. Z: Tá, termina de ler. 15. S: <LÊ O ENUNCIADO DO ITEM A> 16. Z: Tá, primeiro a gente vai achar a moda, então. A moda a gente já achou... A moda é 16 anos... Peraí... Deixa eu anotar aqui. 17. S: Acho que a mediana tem que fazer esse vezes esse... 18. Z: Não fala assim <RISOS>. 19. S: Tá, primeiro tem que fazer 11×3... 20. Z: Faz as contas aí... 21. S: Primeiro tem que fazer 11×3... que vai dar 33. Aí depois a gente tem que fazer... 22. Z: A moda é 16 anos.	

Atividade de Estatística

Em uma escola pública de Santo André, um grupo de alunos do terceiro ano do Ensino Médio, realizando uma pesquisa escolar, resolveu investigar a opinião dos colegas sobre a política de cotas para o ingresso em universidades federais. Essa escola possuía 938 alunos, sendo 417 estudantes do Ensino Fundamental, estudando no período vespertino (6º ao 9º ano) e 521 no Ensino Médio, estudando no período matutino (204 no 1º ano, 179 no 2º ano e 138 no 3º ano). Desse 938 alunos, 513 são mulheres. Esse grupo decidiu entrevistar 40 alunos, sendo 20 (dez homens e dez mulheres) do período matutino e 20 (dez homens e dez mulheres) do período vespertino. Elaborou um questionário com doze perguntas, cujas respostas podem ser observadas no banco de dados que segue. Analise os resultados utilizando as representações que julgar adequadas e as medidas que julgar adequadas. Lembre de sempre explicar ao máximo seu pensamento.

a) Com relação à variável "idade", os alunos forneceram as seguintes respostas:

Idade (anos)	11	12	13	14	15	16	17	18	19	20	21	Total
Freq. simples	3	4	4	6	5	8	4	3	2	0	1	40

Como você explica a variação das idades utilizando as ferramentas estatísticas que aprendeu: construa uma ou mais representações gráficas, explicando-as. Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de idades do grupo de 140 alunos pesquisados.

b) Em relação à renda familiar, os alunos obtiveram as seguintes respostas

Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)
1	1108,38	11	4800	21	3000	31	Não sabe
2	1500	12	5100	22	1108,38	32	3328,14
3	2521	13	1108,38	23	Preferência não declarar	33	4623
4	Preferência não declarar	14	1108,38	24	3820	34	4100
5	2800	15	1500	25	2216,76	35	5000
6	1108,38	16	Não sabe	26	2600	36	2400
7	3500	17	2750	27	1200	37	1108,38
8	4714	18	3400	28	1108,38	38	1700
9	1108,38	19	Preferência não declarar	29	30000	39	6330
10	2200	20	1108,38	30	1300	40	Não sabe

23. S: ... 12×4 ... Eu acho que tem que pegar...
24. Z: Fala baixo!
25. S: Eu acho, ué? Eu acho que tem que pegar 33 somar com 48 e dividir por 2, né?
26. D: Dividir por 40!
27. S: Pegar todo o total e dividir por 40. É a mediana.
28. Z: A mediana não, esse é a média.
29. S: É, essa é a média, e a mediana?
30. Z: Vamo achar primeiro a moda, a moda é qual? 16, né? 16 anos.
31. D: 16 anos!
32. Z: E a média é o que?
33. S: Não sei, eu tava falando, aí você começou...
34. Z: Taqui ó...
35. S: Somar tudo e dividir por 4.
36. D: Por 40!
37. S: É, somar tudo e dividir por 40.
38. Z: Vamo falar a média, vamos fazer a média, vai.
39. S: Vai fazer 11×3 ... que vai dar 33. Aí depois coloca aqui embaixo.
40. Z: Tem calculadora aí, faz aí.
41. S: Ah, tá...
42. Z: 11×3 , faz aí... <PAUSA LONGA> <LÊ A TABELA> Dá quanto?
43. S: 606.
44. D: Divide por 40.
45. S: 15,15.
46. Z: O soma de novo! <RELÊ A TABELA>
47. S: Acho que eu errei.
48. Z: Deu quanto?
49. S: 15,25.
50. Z: Deu quanto?
51. D: 601.
52. Z: O seu também, né? Erramos na primeira! Tá vendo? Então dá quanto?
53. S: Dividido por 40? 15,025.
54. D: Tá no '15' pro '16'.
55. Z: Deixa eu ver.
56. S: 15,025.
57. Z: Então a média tá...
58. S: Tá no '16 anos'.

59. Z: Então a gente achou a média e a moda.

60. D: Coloca aí.

61. Z: Entre o '15' e o '16' está a média.

62. S: O que é agora? A mediana?

63. Z: A mediana...

64. S: Tem que passar à caneta.

65. Z: Depois a gente passa. Como é que era a mediana, mesmo?

<PAUSA> Ah! A mediana é metade. O total é 40. A metade de 40 é

20. A gente tem que achar quem é que tá no '20'. Quem é que tá dentro do '20'.

66. D: 19?

67. Z: Não.

68. D: 21 e 20?

69. Z: Não! 40. Quem tá dentro do '20'. É a metade. Tipo, qual idade tá dentro do '20', entendeu?

70. D: Então é 10?

71. Z: Não.

72. D: A metade de 20 é 10.

73. Z: Não. <RISOS> Não tem nem 10 aqui!

74. S: Não tem nem 10! <RISOS> Começa no '11', amiga.

75. Z: Tem que achar a metade! Tá vendo isso aqui tudo?

76. D: Aham.

77. Z: Deu 40, certo? A metade de 40? 20.

78. S: 20.

79. Z: Mais ou menos aqui. Deixa eu ver... Então a gente vai ter que fazer as contas pra ver quem tá no '20'. Deixa eu ver... <PAUSA> O total deu o quê?

80. S: Deu 40.

81. Z: 40 pessoas?

82. S: É.

83. Z: É, 40 alunos. O 20º aluno é o que eu quero.

84. S: Tá, mas onde tá o 20º aluno?

85. Z: A gente vai ter que fazer as contas. <RISOS> <PAUSA> Abre o livro aí, vê se você acha alguma coisa, um exemplo pra gente ver. <PAUSA> Abre na página 46. <SUSSURROS>

86. S: Nós tem que mudar aqui, você colocou 106, é 101. Empresta a borracha. <PAUSA LONGA>

Como você explica a variação das rendas declaradas utilizando as ferramentas estatísticas que aprendeu: construa uma ou mais representações gráficas, explicando-as.

Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de dados do grupo de 40 alunos pesquisados.

L7 Fundamentos familiares

- c) Em relação à escolaridade dos pais, os alunos elaboraram a tabela a seguir. A tabela construída simplifica o banco de dados. Como você evidenciaria essa comparação se quisesse utilizar gráficos? Explique seu pensamento. É possível determinar uma medida-resumo que contribua com a análise desses dados? Qual? Como determinar? Explique seu pensamento.

Escolaridade	do pai	da mãe
não alfabetizado	2	1
ens. fundamental incompleto	2	1
ens. fundamental completo	4	6
ens. médio incompleto	4	0
ens. médio completo	9	15
ens. superior incompleto	6	1
ens. superior completo	8	12
pós-graduação incompleta	3	0
pós-graduação completa	2	4
Total	40	40

87. Z: Deixa eu ver, mano, deixa eu ver aqui rapidinho. Que coisa que é essa? 12? <PAUSA> Certo! A mediana é um negócio complicado!
88. S: Eu sei, mas você anotou aqui!
89. Z: Eu não anotei nada... Anotei as coisa que ele tinha falado... 40, 40 pessoas, certo... <RELÊ A TABELA> Até onde já deu 20? Deixa eu ver... 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22... Aqui, mais ou menos, deu 20. <PAUSA> 22, certo? Deixa eu ver: 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40. Aqui, ó! É aqui. A mediana. Espero que seja isso. Entre 14 e 16 anos. Não, entre 14 e 15 anos. A mediana está entre 14 e 15 anos. Qual que era a próxima? O desvio. <PAUSA>
90. S: -4, né?
91. Z: 11 anos, -4; 12, -3; 13, -2; 14 -1; agora... pra cima. 1, 2, 3, 4, 5, 6. Pra cima agora, 21, 6.
92. S: Vai parecer que é 11 - 4, 12 - 3...
93. Z: Põe como?
94. S: Põe assim, tipo, 'igual', não põe 'menos'.
95. Z: Ué, mas não é igual. 11 não é igual a -4. É o desvio.
96. S: Sim, mas é como se fosse... Mas é resposta, não é?
97. Z: É. Aqui no meu caderno...
98. S: Então, mas coloca igual. Pra não confundir.
99. Z: 11 = -4 é a mediana, 12 = -3 negativo. Agora pra cima, positivo. 21 = 6, seis casas acima, 20 = +5, 19 = +4, 18 = +3, 17 = +2, 16 = 1. Agora a variância. A variância é o que, mesmo?
100. S: Pega todos e divide por...
101. Z: Olha aqui, ó. É pra somar.
102. D: Dividir, ele falou que era dividir.
103. Z: Soma todos... eleva ao quadrado.
104. S: Acho que é soma.
105. Z: Eu tinha anotado mais. Eu acho que isso, gente... A variância é a multiplicação da idade com a frequência. Ao quadrado. Mais 40.
106. D: Pra mim ele falou dividir.
107. Z: Quer ir lá perguntar pra ele? Se vai dividir ou somar por 40 <D SAI DA SALA> <LONGA PAUSA> O desvio padrão é só elevar ao quadrado. Enquanto ela vê a variância, a gente vê o desvio padrão.
108. S: Mas tem que fazer os dois? O desvio e o desvio padrão?
109. Z: É esse aqui, o desvio padrão é ao quadrado. Quem sobra é quem aumenta da mediana. Quem sobra é quem aumenta, que é os positivos

1- Meda = 16 anos
Moda = 15 e 16 anos
Mediana = 14 e 15 anos
Desvio =

11	12	13	14	15	16	17	18	19	20	21	total
3	4	4	6	5	8	4	3	2	0	1	40

11 = -4² = 16
12 = -3² = 9
13 = -2² = 4
14 = -1² = 1
15 = 0² = 0
16 = 1² = 1
17 = 2² = 4
18 = 3² = 9
19 = 4² = 16
20 = 5² = 25
21 = 6² = 36

Variancia =

$$V = 3 \cdot 16 + 4 \cdot 9 + 4 \cdot 4 + 6 \cdot 1 + 5 \cdot 0 + 8 \cdot 1 + 4 \cdot 4 + 3 \cdot 9 + 2 \cdot 16 + 0 \cdot 25 + 1 \cdot 36 = 125$$

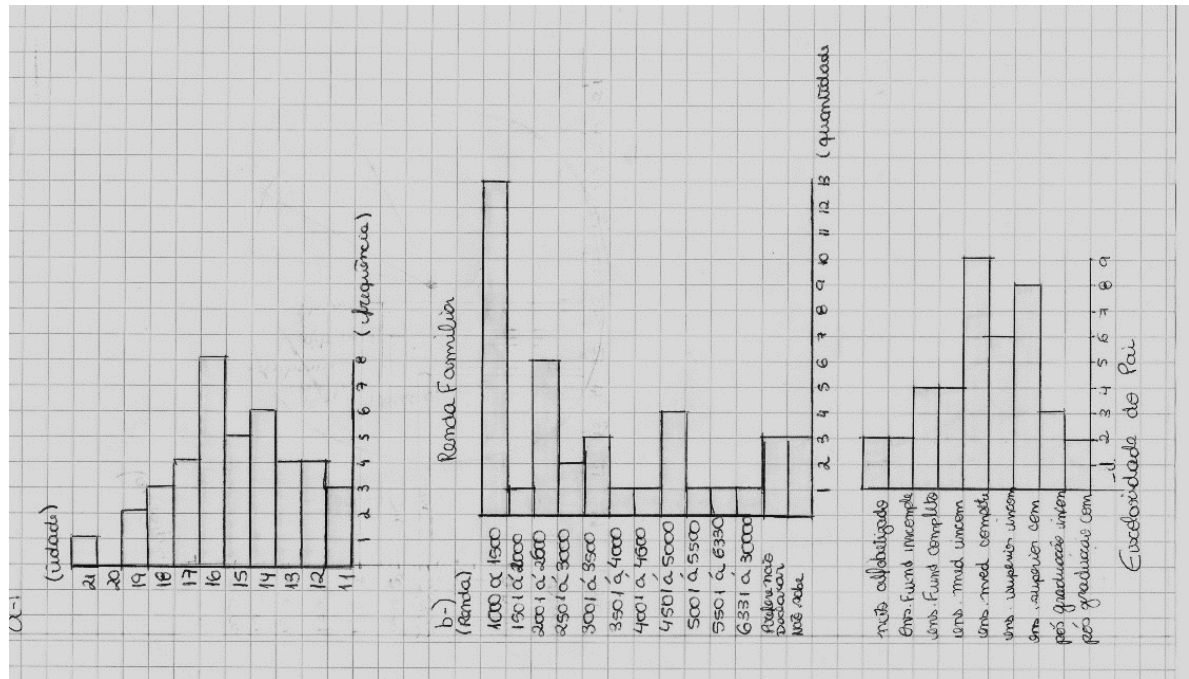
$$V = \frac{125}{40} = 3,125$$

Desvio padrão = $\sqrt{3,125} \approx 1,77$

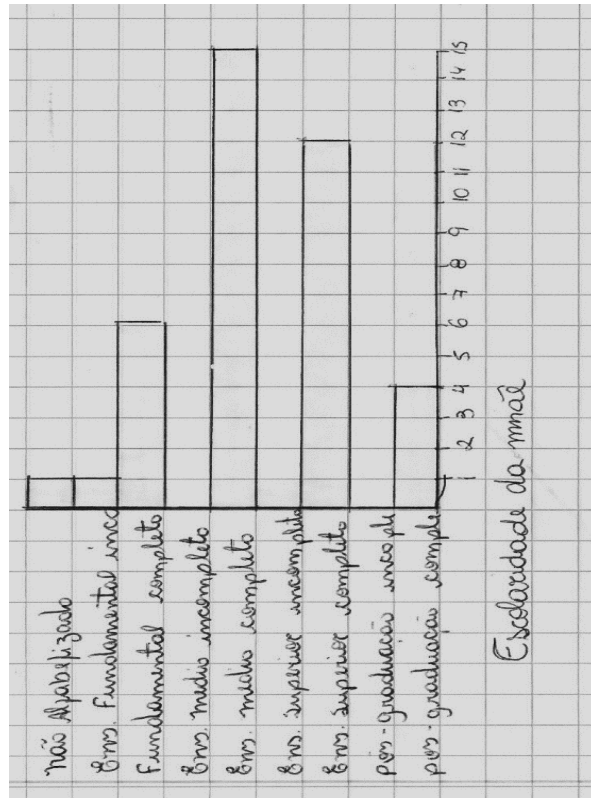
2- Meda = 1108,38
Moda = 2909,45
Mediana = 2501 e 3000
Desvio =

e os negativos. A variância, que são os números ao quadrado mais ou dividido por 40. A gente responde. O desvio padrão... ele é... ele é ao quadrado, da variância. <SUSSURROS> <PAUSA> Fazer os gráficos, enquanto isso. 110. D: <RETORNA> É dividido mesmo, e ela anotou errado. 111. Z: Ah... 112. S: E tem um que tem que fazer a raiz quadrada, não é? 113. D: É essa. 114. Z: A variância? 115. D: Não é? Aqui. Ele falou uma coisa assim... que, tipo, é 5, quando elevar ao quadrado, vai ficar 25. Aí vai ser 25 dividido por 40. 116. Z: Ah! É mesmo! Aqui, ó! Só que não é 5, mano. É 4, porque como tá aqui no meio. 117. S: Aí é 4 dividido por 40? 118. Z: 4 elevado ao quadrado. Vamos elevar logo aqui pra ver o resultado que vai dar. 119. S: Vai devagar, gente... 120. Z: Então vamos voltar para o desvio. 121. D: E ele falou que se tiver empacadinho, não é para ir lá perguntar, pula pra próxima. 122. S: Tá bom. <RISOS> 123. Z: Ó, vou passar aqui. 124. D: 2. É 8, né? 125. Z: 4 é 2^2. É 4^2. 126. S: 16. 127. Z: É, $4^2 = 16$. Agora 3... 128. S: 9. 129. Z: 2×2. 130. S e D: 4. 131. Z: 1×1. 132. S: 1. 133. Z: 6×6? 134. D: 12. 135. S: 36. <RISOS> 136. Z: 5×5? 137. S: 25. 138. Z: 4×4? 139. S: 16.	2001 de 2500 = - 9) moda = 5 Ensino médio Completo mediana = 3 esta entre os ensino médio completo e o ensino Superior incompleto.
---	--

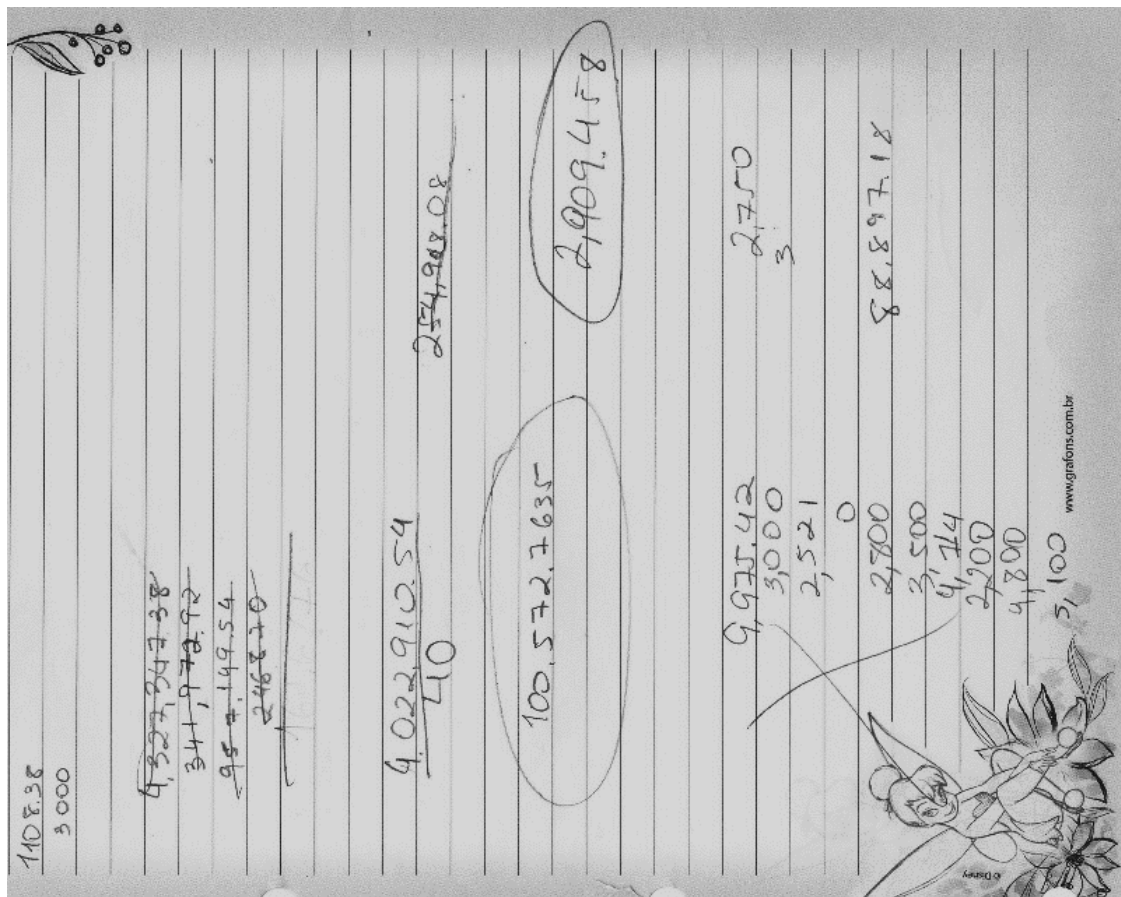
140. Z: $3 \times 3 \times 9 \times 2 \times 2 \dots$
 141. D: 4.
 142. Z: 4. $1 \times 1 \times 1$. Esse aqui é o desvio, certo? Agora a variância.
 143. S: É, mas não tem que pegar esse resultado e dividir por 40?
 144. Z: Soma tudo?
 145. S: É, mas não tem que pegar esse resultado e dividir por 40? Tipo, 36 dividido por 40?
 146. Z: Nem dá pra dividir por 40.
 147. D: Vai dar, tipo, 0,1.
 148. S: É, nem dá.
 149. Z: Agora a gente vai fazer a variância: o número multiplicado pelo outro ao quadrado, dividido por 40.
 150. D: Huuuuuuuummm!
 151. Z: É, professor, no final ela explica tudo. <RISOS> Vai, 11×3 , faz aí.
 152. S: $11 \times 3 = 33$, ao quadrado... 33×33
 153. D: Dá quebrado, 1,0...
 154. S: MIL...
 155. Z: 1.089.
 156. D: Deu número quebrado. É 1.089.
 157. Z: Dá um muro nela...
 158. D: <RISOS> Ah, é verdade, é 1.089.
 159. Z: 14, oh, 12×4 .
 160. D: 48.
 161. Z: $48^2 \dots$ 2.304. Vai, 13...
 162. S: 13... dividido...
 163. D: Vezes 4.
 164. Z: $13 \times 4 \dots$
 165. S: 52.
 166. Z: $52^2 \dots$ 2.304. $14 \times 6 \dots$
 167. D: 84.
 168. Z: 84^2 .
 169. D: Não é esse aqui? Não é esse resultado aqui?
 170. Z: $75^2 \dots$ 5.625. $128^2 = 16.384 \dots$ 17... 68^2 .
 171. S: 68?
 172. Z: $4.624 \dots$ 17×4 . Quanto é que é?
 173. D: 68.
 174. Z: 68 dividido por... oh, 68×68 ? Certo? 4.624 . $18 \times 3 \dots$ 54... $54^2 \dots$



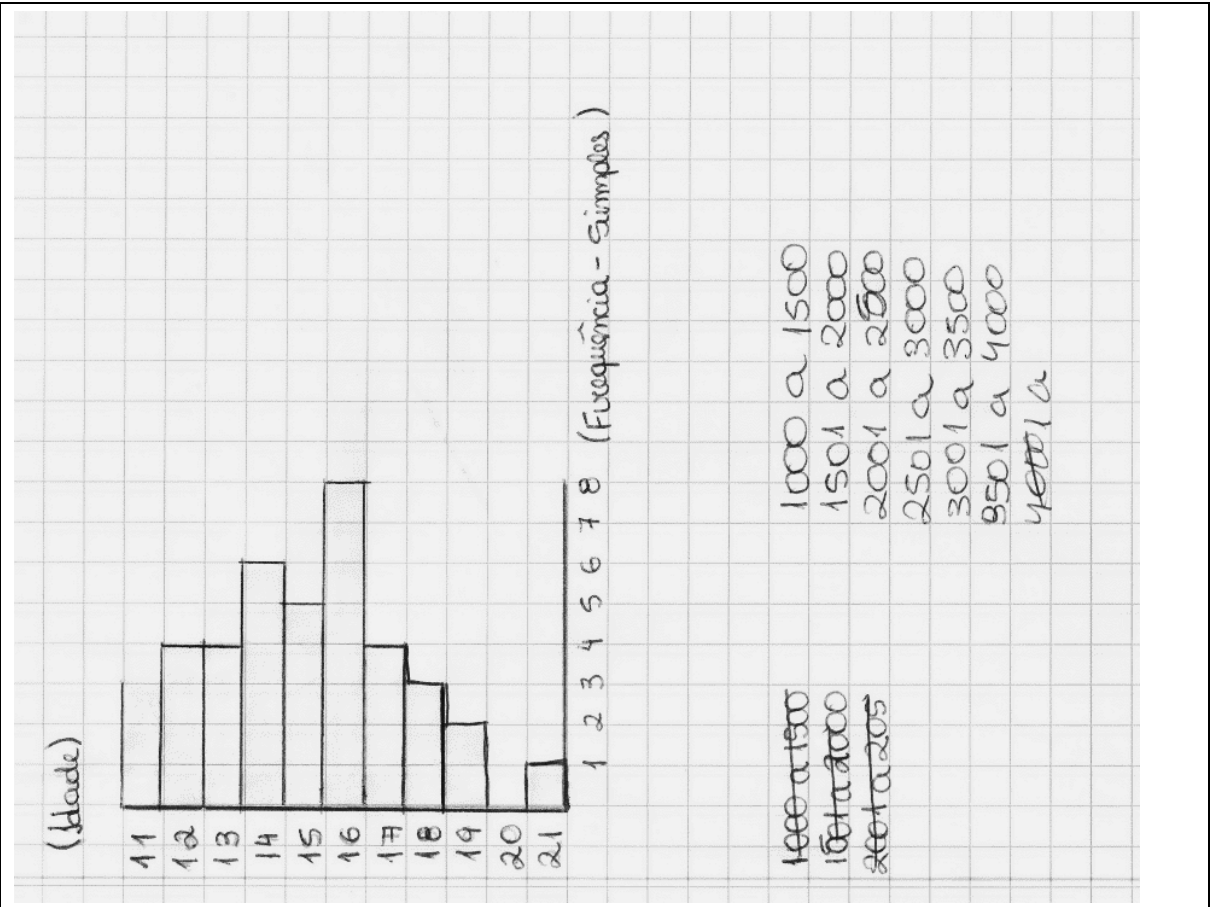
- 2.916... 19×2 , dá 38, 38^2 dá 1.444, 20×0 é 0, 0^2 é 0, 21×1 é 21, 21^2 é 21×21 ...
175. D: 441.
176. Z: 441. Pronto, essa é a variância. Agora vai dividir por... Somar tudo e dividir por... 40? Acho que a gente vai dividir cada uma por 40.
177. D: O resultado, né?
178. Z: Cada um, pra dar o resultado... 1.089 sobre 40.
179. S: O meu deu esse resultado aqui, eu vou pôr sem ponto, gente.
180. D: Ai, deu esse, agora.
181. Z: Não fala "deu esse"! Qual que é? 27.225. 2.304 sobre 40... 57,6. 2.704 sobre 40... 67,6. 7.056 sobre 40?
182. D: 176,4.
183. Z: 5.625 sobre 40.
184. D: 140,62.
185. Z: 16.374 sobre 40... 408,6.
186. Z: 84...
187. D: Aqui deu 409,6.
188. Z: 16.384.
189. S: Eu tinha invertido a ordem.
190. Z: Dividido por 40... 409,6... 4.624 sobre 40... 115,6... 2.916 sobre 40... 72,9... 1.444 sobre 40... 36,1... 441 sobre 40... 11,025. Pronto, a gente achou a variância. Agora é desvio padrão.
191. S: É nesse aí que a gente tira a raiz quadrada, né?
192. D: Tem a fórmula aqui.
193. S: Falei.
194. Z: O desvio padrão é esse.
195. D: O desvio padrão é aquele que tira a raiz quadrada.
196. Z: De 40, né?
197. S: Do número! Do que deu. Do resultado.
198. Z: Que resultado?
199. D: Esse aí, ó!
200. S: É esse resultado aqui. Não dá nem pra tirar a raiz quadrada!
201. Z: Desvio padrão é o da variância. Será que a gente não tinha que somar tudo e dividir por 40?
202. S: Acho que a gente fez alguma coisa errada aí.
203. Z: Acho que a gente não tinha que somar tudo e dividir por 40.
204. D: Qual foi aquela que ele falou assim, que tem que dar uma idade que tá aqui?



205. Z: É a moda. Acho que a gente não tinha que somar tudo e dividir por 40. Não dividir cada um. Porque a variância é uma só, não é 300!
206. D: Olha aqui, uma variância aqui, ó!
207. Z: <LENDO UM EXEMPLO DO LIVRO> A gente fez de cada um. Agora tem que somar tudo e dividir por 40. Vamos pra dois, vai, qual que é? Tem que achar o desvio padrão.
208. S: A gente fez errado!
209. Z: <LENDO UM EXEMPLO DO LIVRO>
210. D: A gente erramos porque esquecemos de colocar o mais.
211. Z: Vamos fazer igual tá aqui.
212. A: A gente tem que fazer com esses números. Não com os números que tão aí.
213. Z: <LENDO UM EXEMPLO DO LIVRO> <REFAZ MULTIPLICANDO CADA IDADE PELA FREQUÊNCIA E ELEVANDO AO QUADRADO> Deu quanto o resultado agora?
214. D: 8.848 sobre 40... 211,2.
215. S: Coloca aproximadamente.
216. Z: Agora, desvio padrão. Raiz quadrada de 211,2... Está entre 14 e 15 anos. Agora vamos fazer o gráfico. Eu vou fazendo enquanto vocês leem a questão.
217. S e D: <LEEM O ITEM B>
218. Z: Já vou chamar a moda, aqui... <PAUSA LONGA> A moda é o salário mínimo. Fala aí pra gravar: A moda é 1.108,38.
219. S: Pra tirar a média é esse vezes esse?
220. Z: É, no caso é renda familiar vezes o sujeito. <SUSSURROS>
221. S: Vai achando a moda aí... <SUSSURROS> <RISOS>
222. Z: Como é que eu vou montar o gráfico?
223. S: Esse aqui entra como 'zero', né? Ele 'Prefere não declarar'... <LEEM A TABELA> <CONVERSAS PARALELA> <SUSSURROS> <RISOS>
224. Z: Não estou conseguindo fazer esse gráfico...
225. S: Será que ele vai deixar fazer na segunda-feira? Não vai dar tempo...
226. Z: Ele não vai deixar levar pra casa! <PAUSA MUITÍSSIMO LONGA> Deixa eu ver essa variância... Essa variância está errada.
227. D: Acho que essa tá errada e a outra tá certa.
228. Z: Acho que as duas tão erradas, tem que estar dentro do intervalo...



<TERMINA O TEMPO DA AULA>



SEGUNDA SESSÃO [01:00:01] – 26/11/2018	
Produção oral	Produção escrita
APÓS A RELEITURA DO EXERCÍCIO E DO ENUNCIADO DO ITEM B 229. Z: Vai, termina de fazer o negócio. Ela vai fazendo a variância, eu vou arrumando aqui e você vai fazendo a média do... sei lá. 230. S: 1.750×17. 231. Z: Põe o seu salário, aí. 232. S: O meu salário não é muito confiável, não... $3.400 \times 18...$ <SUSSUROS> <PAUSA MUITÍSSIMO LONGA> 233. Z: Agora eu vou fazer o gráfico da primeira tabela. <PAUSA MUITÍSSIMO LONGA> 234. S: Pronto, calculei a média. 235. Z: Deu quanto? 236. S: Cadê? 237. Z: Como é que você vai calcular a média, mesmo? 238. S: Eu vou somar tudo, depois dividir por 40. 239. Z: Cê acha que aqui tem que colocar mais alguma coisa? Tem? 240. S: Não. 241. Z: Vai riscando com o lápis, senão cê vai se perder. <PAUSA LONGA> Você não está entendendo nada, você vai ter que fazer tudo de novo! <PAUSA MUITÍSSIMO LONGA> 242. Z: Já achou a média? 243. S: Tô terminando. <PAUSA MUITÍSSIMO LONGA> E agora divide isso aqui por quanto? 244. Z: Quantos são? 245. S: 40. 246. Z: Então é por 40, ué! <PAUSA LONGA> 247. S: Vai dar esse valor aqui... 248. Z: Fala mais alto! Não, nem tem esse valor aí! Como é que vai dar isso? 249. S: Não sei. 250. Z: Não, tá errado! Cê é loca, não tem 100 mil aí não! Tem que fazer TUDO DE NOVO! 251. S: Eu tô fazendo tudo de novo! 252. Z: Então tem que apagar. Se você deixar essa bagunça aí, o professor não vai entender nada! <SOM DE RESPIRAÇÃO PESADA>	

	253. S: Você já terminou com esse negócio aí que você riscou? 254. Z: Se você não achar a média e a mediana, eu não faço nada! Por que você não faz assim que nem eu? Aqui ó, você vai fazer assim... <SUSSURROS> 255. S: Salário de 1.800 tem dois... 256. Z: Então põe 2, ué! <SOM DE RESPIRAÇÃO PESADA> 257. S: <SUSSURROS> 258. Z: Conseguiu? 259. S: Tô conseguindo. <SOM DE CHORO> Que foi, que tá acontecendo, você tá bem? Me fala... <SOM DE RESPIRAÇÃO PESADA> 260. Z: Deixa eu fazer uma conta, aqui... <PAUSA MUITÍSSIMO LONGA> Ah... tem que dividir por 2 primeiro... não... 261. S: Parece que vai dar o mesmo resultado. 262. Z: Vai? 263. S: Deixa eu terminar a conta, depois eu te falo. <SOM DE RESPIRAÇÃO PESADA> 264. Z: Eu acho que a '1' já tá feita. Agora a '2'. 265. S: Deu um valor ainda maior. 266. Z: Então cê tinha errado. 267. S: Não deu esse valor... 268. Z: Divide ó... 269. S: Ah, é, divide por 40... 2.000... Não... É... eu tinha feito errado. Aqui... 2.909,458 270. Z: A moda é qual? 271. S: A moda é... 1.108,38. É a que tem mais. 272. Z: 1.108... 273. S: Vírgula 38. 274. Z: A média é... 275. S: 2.909,458. 276. Z: Põe 4, melhor fazer assim. 277. S: Agora tem que tirar a mediana. 278. Z: Deixa eu ver aqui. Mediana, essa aqui tá... Deixa eu ver... Deixa eu ver como ficou o gráfico... Deu 2 mil e... 279. S: 2.909... 280. Z: Aqui, ó! A mediana tá no... Aqui... você tem que ajudar. 281. S: O desvio é... Apago isso, será? Essa bagunça... 282. Z: Cadê o branquinho?... Mediana tá aqui...
--	--

283. S: Apaga. 284. Z: Já vai achando a variância. 285. S: <SUSSURROS> 286. Z: Quietá! O desvio é... Pra cá, é negativo, pra lá, é positivo. Ó, desse lado, é negativo, daquele lado, é positivo. Lê a questão aí... 287. S: <LÊ O ENUNCIADO DA QUESTÃO C>	
<TERMINA O TEMPO DA AULA>	
TERCEIRA SESSÃO [00:22:28] – 30/11/2018	Produção escrita
Produção oral 288. S: <RELÊ O ITEM C> 289. Z: A moda dá pra achar, a moda é o 'ensino médio completo'. A média não dá pra achar, porque é qualitativa. A mediana dá pra achar também. 290. S: <SUSSURROS> 291. Z: É só colocar! 292. D: Fala baixo. 293. Z: Gente, é só colocar! 294. S: <SUSSURROS> 295. Z: Não interessa, cadê a mediana? Nem lembro como é que faz a mediana. 296. S: Nem lembro como é que faz a mediana. Vê no livro aí. 297. Z: Não tem resposta. 298. S: 13, 14 anos. 299. Z: 14, 15 anos, vê direito! 300. S: Eu não tô falando da média, tô falando da mediana. 301. Z: Eu tô falando da mediana também... A mediana é a média, tipo, aqui tem 40. A mediana é no '20'. Onde tá o '20'? É a vigésima pessoa, é isso? <SUSSURROS> Ensino médio completo e ensino superior incompleto. <PAUSA LONGA> <SUSSURROS> Não reclama, gente! <PAUSA> Eu vou ter que falar o que a gente vai fazer. <SUSSURROS> 302. D: A moda é 15. 303. Z: Cala a boca! 304. D: Foi você que falou que a moda é 15! 305. Z: Eu não disse isso, não! 306. D: A moda é o 'ensino médio completo'. A média e a	

variação não têm, porque é qualitativo. E a mediana fica entre o 'ensino médio completo' e o 'ensino superior incompleto'. 307. Z: E o desvio também não tem. 308. D: E o desvio também não tem. 309. Z: Agora a gente vai fazer o gráfico e... 310. S: E só! 311. Z: É dois gráficos que tem que fazer aqui: o do pai e da mãe. 312. S: Eu sei. 313. Z: Tem que fazer outro gráfico aqui. É o do pai e o da mãe. 314. S: Eu sei. 315. Z: Eu faço o gráfico do pai e você, o da mãe. 316. S: Eu não sei fazer. 317. Z: Minha filha, é só fazer igual ao meu... Cé tá fazendo errado, eu tô fazendo embaixo. Cé tem que fazer lá em cima. 318. S: Eu tô vendo quantos quadradinhos você pulou... 319. Z: Jesus Cristo! Faz o seu primeiro. 320. S: <SUSSURROS> <PAUSA LONGA> 321. Z: Agora fala tudo que a gente fez. Dezo! 322. D: <SUSSURROS> 323. Z: Isso, fala lá! 324. S: <SUSSURROS> 325. Z: Fala tudo que a gente fez! 326. S: Eu não sei explicar, não... 327. Z: Fala lá! 328. S: Explica você. 329. Z: Eu tô terminando de fazer o gráfico. Quando eu terminar, eu explico. 330. S: <SUSSURROS> 331. D: Acho que dá pra escutar ali. 332. Z: A média tá aqui, a moda tá aqui, a mediana tá ali. É só você explicar. 333. S: Então explica. 334. Z: Meu pai do céu! 335. S: Pra tirar a moda, a gente pega aqui... O quê... Como é que fala isso? 336. D: Tabela. 337. S: Nessa tabela... A frequência nessa tabela maior é 8 e a idade é...	
--	--

338. D: 16.
 339. S: 16. Então a moda é 16, porque...
 340. D: É o número maior.
 341. S: É o número maior. Aí, pra...
 342. Z: A moda é o que tem mais quantidade, não é o número maior!
 343. S: Então, é o que tem a frequência maior. Oito. Nesse caso, é o que tem a maior frequência. Esse aqui é o desvio, é o que você falou, né?
 344. Z: É.
 345. S: Pra tirar o desvio, você pega o número que você tirou da mediana, pra cá ou pra lá, como é que eu posso te explicar?
 346. D: Dá idade pra baixo.
 347. S: Sim, e da idade pra cima. Ou seja, se a frequência está entre 14 e 15, você vai pegar o de 14, 15 pra baixo. Que nem, 11 é 15 - 4, porque 4 é pra baixo.
 348. D: Quatro casas antes do 14.
 349. S: Não, quatro casas antes do 15. Que ficaria no meio.
 350. D: Do 14.
 351. S: Do 15.
 352. D: Do 14.
 353. S: 1, 2, 3, 4... do 15. Do 14, é 3. Ou você fez errado aqui, ou tá contando errado. Pra achar a variância entre a frequência e a idade, você pega a frequência, faz a frequência vezes, põe entre parênteses a idade e eleva ao quadrado, mais o número da frente, que é o 4. E aí você vai fazendo isso sucessivamente com todos os números. Aí você divide por 40.
 354. D: Que é o total.
 355. S: Que é o total. Aí, pra achar o desvio padrão, você faz a raiz quadrada da variância, que vai dar... No nosso deu mais ou menos um número quebrado, que vai ser... entre 15 e 16 o nosso. Aí, na segunda tabela, nós fizemos sobre renda familiar. A moda é 1.108,38, que é o salário mínimo. A média... Como é que faz a média mesmo? <SUSSURROS> Aí a gente fez a mediana...
 356. D: No caso, vai fazer a mesma coisa que fez na primeira tabela. Aí no caso é... <SUSSURROS> Aí no caso a mediana tá entre 2.501 e 3.000. Aí o desvio tá entre 2.001 e 2.500. Aí, no último, a moda tá em cima de... ensino médio completo.

357. S: Por quê?
358. D: Por quê? Não faz pergunta difícil.
359. Z: A moda? Porque é onde tem maior número de pessoas.
360. D: A média, o desvio e a variância não dá pra tirar, porque...
361. Z: São qualitativas.
362. D: São qualitativas.
363. S: Aí a mediana tá entre o 'ensino médio completo' e o 'ensino superior incompleto'. Entendeu?
364. Z: O gráfico já tá feito, aí já acabou.
365. D: Vai pôr isso também?
366. Z: É rascunho, ele pediu pra entregar separado.
367. S: Não tá errado...
368. Z: Tá errado, puxa aí, senão ele vai considerar.

Protocolo 2 – Grupo II

B: SEXO FEMININO, 17 ANOS
P: SEXO MASCULINO, 18 ANOS

SESSÃO ÚNICA [01:27:56] – 26/11/2018																																																																																																																	
Produção oral	Produção escrita																																																																																																																
APÓS A LEITURA DO EXERCÍCIO E DO ENUNCIADO DO ITEM A 1. B: Variável? 2. P: É. 3. B: É que tem variância... 4. P: É para você explicar o que é variação das idades utilizando isso aqui, construir uma das representações do gráfico usando medidas, moda, média, mediana, quartis, desvio padrão e amplitude... O que que é que é variação? 5. B: Variância e desvio padrão é pra ver, tipo, o quão errado estão os cálculos. Tipo, é como se fosse uma margem de erro. Uma variância é, tipo, um X, e a diferença entre a variância é o desvio padrão. <LEND O LIVRO DIDÁTICO> Há duas medidas desse tipo: a variância e o desvio padrão, que são usadas sempre... Variância é a média dos quadrados dos desvios... 6. P: Ai tem, tipo, que contar todas essas medidas. <RISOS> Putz! 7. B: <LEND O UM EXEMPLO DO LIVRO DIDÁTICO> Mas por que que está aqui o desvio -2? Tá vendo aqui? Desvio -2! 8. P: É que aqui... Sei lá! Tipo, bota aqui média, moda e mediana, o resto eu não sei fazer, não! 9. B: Eu também não sei fazer. <RISOS> 10. P: É o mais simples, véio! 11. B: No trabalho eu é que fiz o cartaz, e o que mais pegou foi essa parte de variância e desvio padrão. 12. P: Deve ser por aqui, no '16', tipo, esse é a média, a moda e a mediana. Pode ser que tenha a ver com isso. <RISOS> 13. B: A média é difícil, porque é o que menos aparece... 14. P: Como assim? <RISOS> 15. B: A moda é o que sempre aparece mais...	Atividade de Estatística Em uma escola pública de Santo André, um grupo de alunos do terceiro ano do Ensino Médio, realizando uma pesquisa escolar, resolveu investigar a opinião dos colegas sobre a política de cotas para o ingresso em universidades federais. Essa escola possui 938 alunos, sendo 417 estudantes do Ensino Fundamental, estudando no período vespertino (6º ao 9º ano) e 521 no Ensino Médio, estudando no período matutino (204 no 1º ano, 179 no 2º ano e 138 no 3º ano). Desse 938 alunos, 513 são mulheres. Esse grupo decidiu entrevistar 40 alunos, sendo 20 (dez homens e dez mulheres) do período matutino e 20 (dez homens e dez mulheres) do período vespertino. Elaborou um questionário com doze perguntas, cujas respostas podem ser observadas no banco de dados que segue. Analise os resultados utilizando as representações que julgar adequadas e as medidas que julgar adequadas. Lembre de sempre explicar ao máximo seu pensamento. a) Com relação à variável "idade", os alunos forneceram as seguintes respostas: <table><tr><th>Idade (anos)</th><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td><td>Total</td></tr><tr><th>Freq. simples</th><td>3</td><td>4</td><td>4</td><td>6</td><td>5</td><td>8</td><td>4</td><td>3</td><td>2</td><td>0</td><td>1</td><td>40</td></tr></table> Como você explica a variação das idades utilizando as ferramentas estatísticas que aprendeu: construa uma ou mais representações gráficas, explicando-as. Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de idades do grupo de 40 alunos pesquisados. b) Em relação à renda familiar, os alunos obtiveram as seguintes respostas: <table><tr><th>Sujeito</th><th>Renda Familiar (R\$)</th><th>Sujeito</th><th>Renda Familiar (R\$)</th><th>Sujeito</th><th>Renda Familiar (R\$)</th></tr><tr><td>1</td><td>1108,38 /</td><td>11</td><td>4800</td><td>21</td><td>3000</td><td>31</td><td>Não sabe</td></tr><tr><td>2</td><td>1500 /</td><td>12</td><td>5100</td><td>22</td><td>1108,38 /</td><td>32</td><td>3328,14</td></tr><tr><td>3</td><td>2521</td><td>13</td><td>1108,38 /</td><td>23</td><td>Pré-rev não declarar</td><td>33</td><td>4623</td></tr><tr><td>4</td><td>Pré-rev não declarar</td><td>14</td><td>1108,38 /</td><td>24</td><td>3820</td><td>34</td><td>4100</td></tr><tr><td>5</td><td>2800</td><td>15</td><td>1500 /</td><td>25</td><td>2216,76</td><td>35</td><td>5000</td></tr><tr><td>6</td><td>1108,38 /</td><td>16</td><td>Não sabe</td><td>26</td><td>2600</td><td>36</td><td>2400</td></tr><tr><td>7</td><td>3500</td><td>17</td><td>2750</td><td>27</td><td>1200</td><td>37</td><td>1108,38 /</td></tr><tr><td>8</td><td>4714</td><td>18</td><td>3400</td><td>28</td><td>1108,38 /</td><td>38</td><td>1700</td></tr><tr><td>9</td><td>1108,38 /</td><td>19</td><td>Pré-rev não declarar</td><td>29</td><td>30000</td><td>39</td><td>6330</td></tr><tr><td>10</td><td>2200</td><td>20</td><td>1108,38 /</td><td>30</td><td>1300</td><td>40</td><td>Não sabe</td></tr></table>	Idade (anos)	11	12	13	14	15	16	17	18	19	20	21	Total	Freq. simples	3	4	4	6	5	8	4	3	2	0	1	40	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	1	1108,38 /	11	4800	21	3000	31	Não sabe	2	1500 /	12	5100	22	1108,38 /	32	3328,14	3	2521	13	1108,38 /	23	Pré-rev não declarar	33	4623	4	Pré-rev não declarar	14	1108,38 /	24	3820	34	4100	5	2800	15	1500 /	25	2216,76	35	5000	6	1108,38 /	16	Não sabe	26	2600	36	2400	7	3500	17	2750	27	1200	37	1108,38 /	8	4714	18	3400	28	1108,38 /	38	1700	9	1108,38 /	19	Pré-rev não declarar	29	30000	39	6330	10	2200	20	1108,38 /	30	1300	40	Não sabe
Idade (anos)	11	12	13	14	15	16	17	18	19	20	21	Total																																																																																																					
Freq. simples	3	4	4	6	5	8	4	3	2	0	1	40																																																																																																					
Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)																																																																																																												
1	1108,38 /	11	4800	21	3000	31	Não sabe																																																																																																										
2	1500 /	12	5100	22	1108,38 /	32	3328,14																																																																																																										
3	2521	13	1108,38 /	23	Pré-rev não declarar	33	4623																																																																																																										
4	Pré-rev não declarar	14	1108,38 /	24	3820	34	4100																																																																																																										
5	2800	15	1500 /	25	2216,76	35	5000																																																																																																										
6	1108,38 /	16	Não sabe	26	2600	36	2400																																																																																																										
7	3500	17	2750	27	1200	37	1108,38 /																																																																																																										
8	4714	18	3400	28	1108,38 /	38	1700																																																																																																										
9	1108,38 /	19	Pré-rev não declarar	29	30000	39	6330																																																																																																										
10	2200	20	1108,38 /	30	1300	40	Não sabe																																																																																																										

16. P: Repetido!
17. B: Moda é, tipo, o que mais aparece mais. Aqui tá as idades...
18. P: Aham...
19. B: E aqui tá as frequências. E qual que mais aparece aqui é a de 8.
20. P: Ah, tá!
21. B: Ah, mas a variável já tá aqui, as idades que os alunos forneceram... Isso aqui já é a variável!
22. P: Então tem que fazer só... É só fazer o gráfico aqui.
23. B: <RELÉ O ENUNCIADO DO PROBLEMA> Dados, tem que fazer os dados.
24. P: É, tem que passar os dados junto com as contas. <RISOS>
25. B: Cé tá rindo de nervoso? <RISOS>
26. P: Tipo isso! <RISOS>
27. B: Hum...
28. P: Aqui já tá falando, tipo, da variação, né? Então, é só passar pro gráfico.
29. B: Mas tem que explicar, né?
30. P: É.
31. B: <RISOS> E se a gente usar gráfico de barras, assim, e aí, no eixo X, vai tar as idades e, no eixo Y, vai tar as frequências.
32. P: É.
33. B: Aí, de 11 anos, vai tar uma frequência, de 12 anos, outra frequência.
34. P: De 12, outra frequência... <RISOS> É, eu acho que é isso mesmo.
35. B: Mas aí precisa explicar, né? Então... <RELÉ O PROBLEMA>
36. P: Não era pra eu ter faltado nesse dia. <RISOS> <PAUSA MUITO LONGA> Como a gente vai explicar essa coisa? Essa variação? Não, mas você sabe o que é variação?
37. B: É isso que eu falei, variação, variável e desvio padrão é a mesma coisa... São parecidos. <PAUSA MUITO LONGA> <TOSSE>
38. P: Mais ou menos, esse desvio padrão tinha que fazer antes, tipo, a moda e a mediana tinha alguma coisa a ver. Não, tinha que ter a ver com esses negócios, com a média, para depois entrar nesses desvios. Não precisa usar todas... a gente podia, tipo, usar...
39. B: Pode colocar o que tem mais frequência, tipo, a moda...
40. P: É, coloca tipo, todas essas medidas, a idade, a frequência e, depois, tipo, a moda.

Como você explica a variação das rendas declaradas utilizando as ferramentas estatísticas que aprendeu: construa uma ou mais representações gráficas, explicando-as.

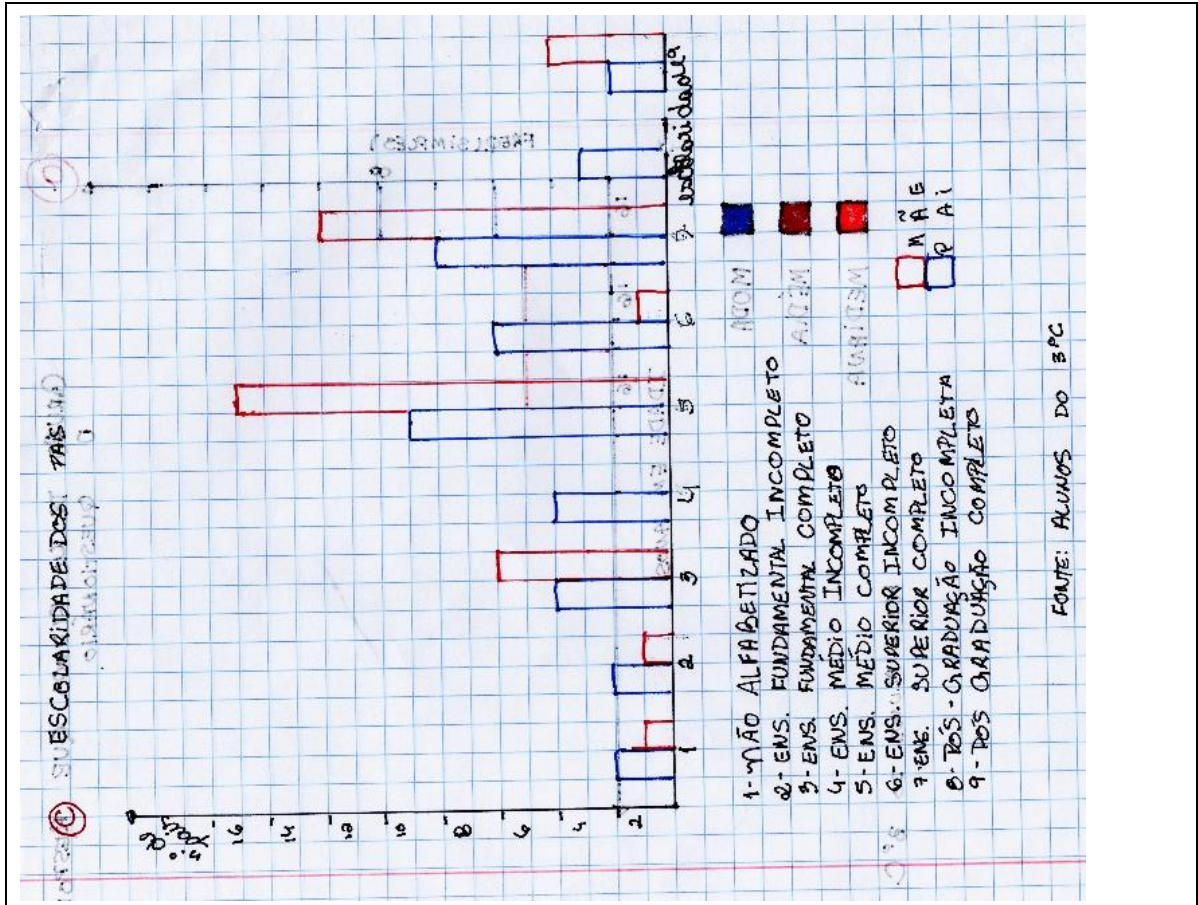
Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de dados

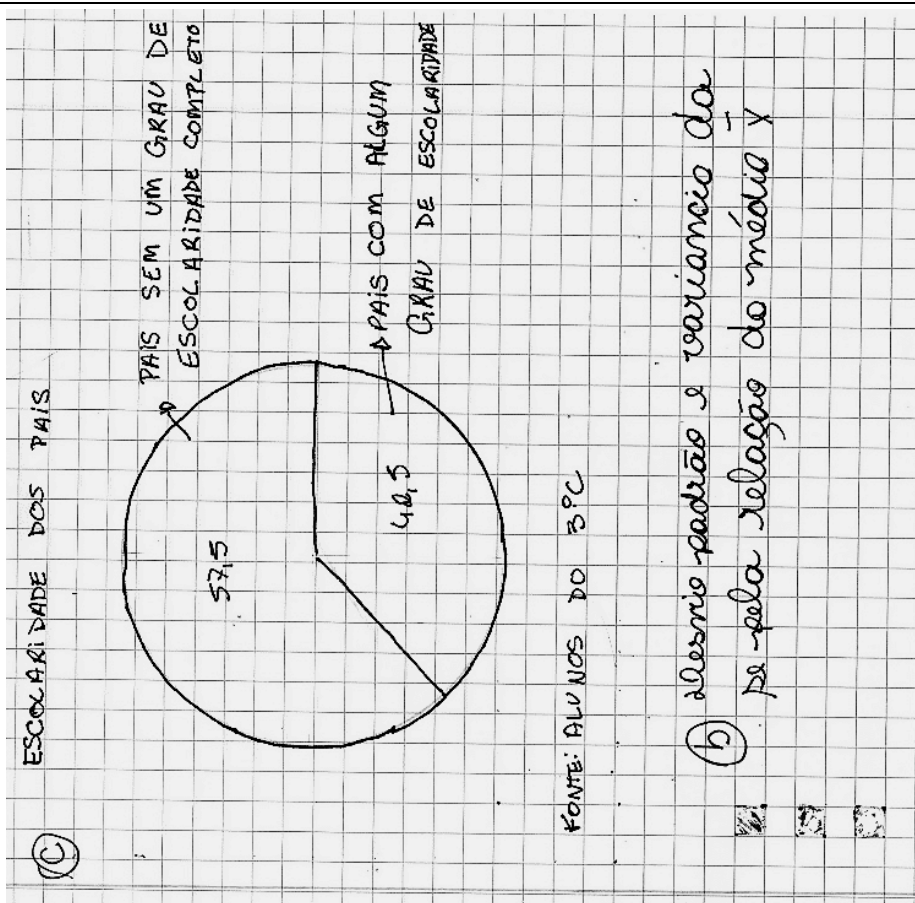
grupo de 40 alunos pesquisados.

c) Em relação à escolaridade dos pais, os alunos elaboraram a tabela a seguir. A tabela construída simplifica o banco de dados. Como você evidenciaria essa comparação se quisesse utilizar gráficos? Explique seu pensamento. É possível determinar uma medida-resumo que contribua com a análise desses dados? Qual? Como determinar? Explique seu pensamento.

Escolaridade	do pai	da mãe
não alfabetizado	2	1
ens. fundamental incompleto	2	1
ens. fundamental completo	4	6
ens. médio incompleto	4	0
ens. médio completo	9	15
ens. superior incompleto	6	1
ens. superior completo	8	12
pós-graduação incompleta	3	0
pós-graduação completa	2	4
Total	40	40

67. P: AI, CALMA AÍ QUE EU FIZ ERRADO! Tem que colocar em ordem crescente. Que bos*! <PALAVRÃO>
68. B: <RISOS> Mas já tá em ordem crescente. Não, perai... Quando tá em ímpar, pega o do meio...
69. P: É, pega o central. Ou então, pega os dois do meio e divide por 2.
70. B: É.
71. P: 16, vai dar tudo 16. A média é 16, a moda é 16 e a mediana é 16. <RISOS> Na frequência.
72. B: Caramba!
73. P: A média é 3.
74. B: Hã? <PAUSA MUITO LONGA>
75. P: Então, aqui... <PAUSA MUITO LONGA> Mano, eu tô fazendo alguma coisa errada aqui, tá dando 10!
76. B: Por quê?
77. P: Tá faltando uns valores...
78. B: Tem certeza?
79. P: Aqui tem que ter 11 intervalos... mas tem 10!
80. B: Calma, vamos contar...
81. P: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. Já tá em ordem crescente. Não sei se está faltando alguma coisa.
82. B: 3 tem, 4 tem, 6 tem...
83. P: Calma, já achei, tem três '4's'. É que eu só coloquei dois! <PAUSA LONGA> A mediana vai ser 4. E a moda também, porque é a que mais se repete. O '4' é o que mais se repete. São três...
84. B: Não, você é muito inteligente! <RISOS> Quando eu crescer, eu quero ser como você!
85. P: Sonho de muitos. <RISOS>
86. B: Falta o desvio padrão, quartis...
87. P: Não, aqui ele fala, ó: "determine uma ou mais medidas". <RELÊ O ENUNCIADO DO PROBLEMA>
88. B: Pode ser, mas a gente precisa fazer um gráfico, um outro tipo de gráfico. E a gente precisa colocar fonte, título e fonte. <PAUSA LONGA> Acho que é isso.
89. P: Mano, eu tô muito fud* <PALAVRÃO> em Matemática!
90. B: Por quê?
91. P: Porque eu não fiz o trabalho do orçamento, da renda familiar...
92. B: Mas você apresentou?





93. P: Sim.
94. B: Socorro, Deus! <RISOS> Eu tenho que ir lá... no banheiro.
95. P: Vai lá, enquanto eu vou resolvendo... <PAUSA MUITÍSSIMO LONGA ATÉ O RETORNO DE B>
96. B: E aí?
97. P: Esse é bem mais complicado...
- LEITURA DO EXERCÍCIO E DO ENUNCIADO DO ITEM B
98. P: Então, agora tem que explicar o que é variação. Como a gente explicaria essa no gráfico? <PAUSA MUITO LONGA>
99. B: Pra achar a variância, você precisa achar primeiro o desvio...
100. P: Repete!
101. B: Pra achar a variância, você precisa, tipo, achar, primeiro, o desvio.
102. P: E como é que acha o desvio? <RISOS> <PAUSA MUITO LONGA> <CONVERSA PARALELA SOBRE NOTAS E SITUAÇÃO DOS ALUNOS QUANTO À APROVAÇÃO FINAL, FORMATURA>
- LEITURA DO EXERCÍCIO E DO ENUNCIADO DO ITEM C
103. B: <CANTAROLA ENQUANTO FAZ OS GRÁFICOS>
104. P: E agora, o que você está fazendo?
105. B: Estou contando quantos tem, para achar a moda... Enfim, essa é a moda
106. P: Então, você não disse que sabia calcular a variação?
107. B: Variância e desvio é mais complicado, vou fazer esse outro, que é mais fácil.
108. P: Esse é mais simples, eu acho. Então, nós poderíamos fazer um gráfico.
109. B: É.
110. P: Nós poderíamos fazer dois gráficos. Colocaria a escolaridade... do jeito que tá aqui, não alfabetizado, as outras... Daí o outro, poderia colocar a média, a moda e a mediana
111. B: É. <PAUSA MUITÍSSIMO LONGA, B ESBOÇA OS GRÁFICOS, ENQUANTO P TENTA DETERMINAR AS MEDIDAS RESUMO NA CALCULADORA> Acho que tá feito...

112. P: Tá nada! Aí, eu tava pensando em fazer assim o outro gráfico: comparar a escolaridade do pai e da mãe em cada... como chama? Tipo... anaf... analfabetização do pai e da mãe, comparar os dois, porque, tipo, quero fazer de cada, de cada coisa.
113. B: Acho que entendi... Estou incontente com esse gráfico... E se a gente fizer um gráfico de setor?
114. P: Pode ser! Mas, como a gente vai fazer, tipo, a analfabetização, tirar a média desse? Porque não tem outro jeito, esse é o único jeito.
115. B: Não tem como fazer. Tipo, são só dois, como é que você vai fazer? Eu pensei em fazer uma comparação...
116. P: Então, eu também!
117. B: Não sei...
118. P: Porque, por exemplo, tipo, quantos são a média, de pai e mãe que, tipo, não é alfabetizado. Aí, eu ia fazer isso, entendeu? Mas como é que vai fazer? 2? A média de 2?
119. B: 2 por 2. <RISOS>
120. P: Ou poderia fazer, tipo, os pais que completaram só o ensino médio completo até o ensino superior.
121. B: E se a gente pegar só os que tem completo?
122. P: Pode ser, mais fácil.
123. B: Entendeu o que eu quero fazer?
124. P: Pode ser. De cada, né? Ou o total?
125. B: De cada, né? Porque o total não vai fazer sentido, porque os valores são muito pequenos pra trabalhar em cima deles.
126. P: Número de pais que completaram o ensino completo.
127. B: Completaram o ensino completo... <RISOS>
128. P: De... <PAUSA MUITO LONGA>
129. B: Vai na sala do Cassio <PROFESSOR DE MATEMÁTICA DELES> e pega um compasso, por favor. <PAUSA MUITO LONGA>
130. P: Táí.
131. B: <CANTAROLA ENQUANTO FAZ OS GRÁFICOS>
<CONVERSA PARALELA SOBRE PROVAS E NOTAS> <PAUSA MUITO LONGA>
132. P: <BOCEJA>
133. B: Pronto, agora só falta o outro.
134. P: Como assim? Não sei nem por onde começar... Porque o desvio... Sei lá. Se tivesse como explicar de outra forma...

135. B: Pode explicar de outra forma, não pode?
136. P: <RELÊ O PROBLEMA>
137. B: A gente pode explicar com nossas palavras, não pode?
<CANTAROLA ENQUANTO FAZ OS GRÁFICOS> <PAUSA MUITO LONGA>
138. P: Sei lá, a gente pode deixar como tá... <PAUSA MUITO LONGA>
Tá com fome? <DESVIO> Nesse desvio, o Cassio ferrou. <RISOS>
139. B: <LENDO O LIVRO DIDÁTICO> "Desvio médio é uma variável definida em relação à média" ...
140. P: É, mas isso já... sabemos.
141. B: <LENDO O LIVRO DIDÁTICO> "Desvio padrão e variância dá-se em relação à média" ...
142. P: Tá bom, né?
143. B: Eu vou entregar assim para ele...

Protocolo 3 – Grupo III

N: SEXO FEMININO, 17 ANOS
L: SEXO FEMININO, 18 ANOS

PRIMEIRA SESSÃO [02:21:01] – 28/11/2018 (Nesta sessão, N estava sozinha, pois L havia faltado, mas, por solicitação do professor, comentou seus procedimentos em voz alta.)																																																																																																																																																											
Produção oral	Produção escrita																																																																																																																																																										
Atividade de Estatística																																																																																																																																																											
1. N: <LÊ O ENUNCIADO NOVAMENTE> 2. N: Na questão 'a', primeiro, vou fazer a média, aí eu vou somar todos os valores e dividir pela quantidade... Média... <PAUSA MUITO LONGA> 3. N: Primeiro, eu fiz a média das idades. Daí, eu somei todas as idades e, aí, eu dividi pela quantidade. Total deu 176 dividido por 11, que dá 16. Então, a média, com relação à variável de idade dos alunos, é 16. 4. N: A moda... é... 16 também, porque é o número que mais aparece. 5. N: E a mediana também é 16. 6. N: <RELÊ O ENUNCIADO> Deixa eu fazer a variância, nesse caso... A variância eu vou montar uma tabelinha com todas as idades e vou fazer as idades menos a média. Então, eu pego o valor das idades e subtraio da média. 7. N: Com esse valor, eu vou colocar ele ao quadrado, e os resultados ao quadrado eu vou somar e dividir pela quantidade pra ter a variância. Então, vamo lá: 11, 12, 13, 14. <SUSSURROS> Aí meu Deus, é complicado... Idade... <PAUSA MUITO LONGA> Então, vamo lá. 8. N: Então, pra calcular a variância, eu peguei o valor das idades e tirei com o valor da média. Eu não posso somar todos os resultados, porque vai dar zero. Então, eu preciso colocar, fazer todos os resultados ao quadrado, somar e, com o resultado ao quadrado, eu vou fazer a média deles, somar todos e dividir por 11, e essa vai ser a variância. <PAUSA MUITO LONGA> <SUSSURROS> 9. N: Então, na questão 'a', eu tenho um pouco de dificuldade de construir gráfico, mas... eu fiz a média, que eu já expliquei, a moda também, a mediana, que é só organizar, e a variância, que foi a última coisa que eu fiz. E agora, depois que eu somei tudo e dividi por 11, o resultado é 10 e o desvio padrão é 3, porque o desvio padrão é o resultado da	Em uma escola pública de Santo André, um grupo de alunos do terceiro ano do Ensino Médio, realizando uma pesquisa escolar, resolveu investigar a opinião dos colegas sobre a política de cotas para o ingresso em universidades federais. Essa escola possuía 938 alunos, sendo 417 estudantes do Ensino Fundamental, estudando no período vespertino (6º ao 9º ano) e 521 no Ensino Médio, estudando no período matutino (204 no 1º ano, 179 no 2º ano e 138 no 3º ano). Desses 938 alunos, 513 são mulheres. Esse grupo decidiu entrevistar 40 alunos, sendo 20 (dez homens e dez mulheres) do período matutino e 20 (dez homens e dez mulheres) do período vespertino. Elaborou um questionário com doze perguntas, cujas respostas podem ser observadas no banco de dados que segue. Analise os resultados utilizando as representações que julgar adequadas e as medidas que julgar adequadas. Lembre de sempre explicar ao máximo seu pensamento. a) Com relação à variável "idade", os alunos forneceram as seguintes respostas: <table><tr><th>Idade (anos)</th><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td><td>Total</td></tr><tr><th>Freq. simples</th><td>3</td><td>4</td><td>4</td><td>6</td><td>5</td><td>8</td><td>4</td><td>3</td><td>2</td><td>0</td><td>1</td><td>40</td></tr></table> Como você explica a variação das idades utilizando as ferramentas estatísticas que aprendeu: construa uma ou mais representações gráficas, explicando-as. Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de idades do grupo de 40 alunos pesquisados. b) Em relação à renda familiar, os alunos obtiveram as seguintes respostas <table><tr><th>Sujeito</th><th>Renda Familiar (R\$)</th><th>Sujeito</th><th>Renda Familiar (R\$)</th><th>Sujeito</th><th>Renda Familiar (R\$)</th><th>Sujeito</th><th>Renda Familiar (R\$)</th></tr><tr><td>1</td><td>1108,38</td><td>X</td><td>11</td><td>4800</td><td>X</td><td>21</td><td>3000</td><td>X</td><td>31</td><td>Não sabe</td><td>X</td></tr><tr><td>2</td><td>1500</td><td>X</td><td>12</td><td>5100</td><td>X</td><td>22</td><td>1108,38</td><td>X</td><td>32</td><td>3328,14</td><td>X</td></tr><tr><td>3</td><td>2521</td><td>X</td><td>13</td><td>1108,38</td><td>* X</td><td>23</td><td>Prefero não declarar</td><td>X</td><td>33</td><td>4623</td><td>X</td></tr><tr><td>4</td><td>Prefero não declarar</td><td>X</td><td>14</td><td>1108,38</td><td>X</td><td>24</td><td>3820</td><td>X</td><td>34</td><td>4100</td><td>X</td></tr><tr><td>5</td><td>2800</td><td>X</td><td>15</td><td>1500</td><td>X</td><td>25</td><td>2216,76</td><td>X</td><td>35</td><td>5000</td><td>X</td></tr><tr><td>6</td><td>1108,38</td><td>X</td><td>16</td><td>Não sabe</td><td>X</td><td>26</td><td>2600</td><td>X</td><td>36</td><td>2400</td><td>X</td></tr><tr><td>7</td><td>3500</td><td>X</td><td>17</td><td>2750</td><td>X</td><td>27</td><td>1200</td><td>X</td><td>37</td><td>1108,38</td><td>X</td></tr><tr><td>8</td><td>4714</td><td>X</td><td>18</td><td>3400</td><td>X</td><td>28</td><td>1108,38</td><td>* X</td><td>38</td><td>1700</td><td>X</td></tr><tr><td>9</td><td>1108,38</td><td>X</td><td>19</td><td>Prefero não declarar</td><td>X</td><td>29</td><td>30000</td><td>X</td><td>39</td><td>6330</td><td>X</td></tr><tr><td>10</td><td>2200</td><td>X</td><td>20</td><td>1108,38</td><td>X</td><td>30</td><td>1300</td><td>X</td><td>40</td><td>Não sabe</td><td>X</td></tr></table>	Idade (anos)	11	12	13	14	15	16	17	18	19	20	21	Total	Freq. simples	3	4	4	6	5	8	4	3	2	0	1	40	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	1	1108,38	X	11	4800	X	21	3000	X	31	Não sabe	X	2	1500	X	12	5100	X	22	1108,38	X	32	3328,14	X	3	2521	X	13	1108,38	* X	23	Prefero não declarar	X	33	4623	X	4	Prefero não declarar	X	14	1108,38	X	24	3820	X	34	4100	X	5	2800	X	15	1500	X	25	2216,76	X	35	5000	X	6	1108,38	X	16	Não sabe	X	26	2600	X	36	2400	X	7	3500	X	17	2750	X	27	1200	X	37	1108,38	X	8	4714	X	18	3400	X	28	1108,38	* X	38	1700	X	9	1108,38	X	19	Prefero não declarar	X	29	30000	X	39	6330	X	10	2200	X	20	1108,38	X	30	1300	X	40	Não sabe	X
Idade (anos)	11	12	13	14	15	16	17	18	19	20	21	Total																																																																																																																																															
Freq. simples	3	4	4	6	5	8	4	3	2	0	1	40																																																																																																																																															
Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)																																																																																																																																																				
1	1108,38	X	11	4800	X	21	3000	X	31	Não sabe	X																																																																																																																																																
2	1500	X	12	5100	X	22	1108,38	X	32	3328,14	X																																																																																																																																																
3	2521	X	13	1108,38	* X	23	Prefero não declarar	X	33	4623	X																																																																																																																																																
4	Prefero não declarar	X	14	1108,38	X	24	3820	X	34	4100	X																																																																																																																																																
5	2800	X	15	1500	X	25	2216,76	X	35	5000	X																																																																																																																																																
6	1108,38	X	16	Não sabe	X	26	2600	X	36	2400	X																																																																																																																																																
7	3500	X	17	2750	X	27	1200	X	37	1108,38	X																																																																																																																																																
8	4714	X	18	3400	X	28	1108,38	* X	38	1700	X																																																																																																																																																
9	1108,38	X	19	Prefero não declarar	X	29	30000	X	39	6330	X																																																																																																																																																
10	2200	X	20	1108,38	X	30	1300	X	40	Não sabe	X																																																																																																																																																

23. N: Então, o fato da idade dizer, é... 16 ser a variável com maior frequência é natural, já que a quantidade do ensino médio é maior.
24. N: Outro detalhe, também, é que os alunos do primeiro e do segundo ano também tem uma quant... Uma quantidade maior do que no terceiro ano, ent... então, também é natural de idade de 16 ser a idade maior, né?
25. N: Basicamente é isso.
26. N: Então, a questão 'a' foi isso. Agora, vou escrever... <PAUSA MUITÍSSIMO LONGA>
27. N: Então, minha explicação pra questão 'a' é: "É necessário entender que uma variação de dados pode ser explicada através do enunciado. Diante disso, resultados como a moda são nitidamente explicados com a análise dos dados transmitidos pelo texto." Então, é, basicamente, tudo que a gente fez no trabalho, porque antes da gente realizar o trabalho, a gente tinha que montar toda a amostra pra você... para o professor entender da onde a gente tirou esses resultados. Então é isso.
28. N: Então... 'b': <LÊ O ENUNCIADO DA QUESTÃO B>
29. <SUSSURROS>... Entrevistou 40 pessoas também... <PAUSA MUITO LONGA>
30. N: Então é 1.108,38 é a mediana, oh! É a moda, né?
- <SUSSURROS>... 'Não sabe'... Devo colocar 'Não sabe' no gráfico... 'Prefere não declarar', 1, 2, 3... tem um 'Não sabe', dois 'Não sabe'... Acho que só...
31. N: É, na questão 'b', primeiro eu vou me organizar, porqueeeee... pra conseguir fazer a média, eu vou... A média, a moda e mediana... Então, eu vou ver quantas pessoas não sabem, quantas pessoas prefere não declarar e vou ver os valores... É, desse jeito. É, tá. Prefiro, é... Renda familiar. <SUSSURROS> Renda familiar...
32. N: Acho que o negócio é fazer na mão... <SUSSURROS>... 'Prefere não declarar' é 1, 2, 3... 3, três pessoas. 'Não sabe' sua renda familiar... 1... Cadê?... 1, 2, duas pessoas. Agora, tem renda familiar... É... de 1.108,38. Vamo vê: é uma, duas... <SUSSURROS>... 3, 4, 5, 6, 7, 8, oito pessoas...
33. N: É... segundo passo, renda... de 1.500... Eu nem tô colocando 'reais', olha que doida! <SUSSURROS> 1.500. 1, 2, aqui já foi. 1, 2, 3, 4, 5, 6, 7, 8. Agora, 1.500, uma, duas, duas pessoas... duas pessoas? É, duas pessoas, uma, duas... é... renda de... <PAUSA

• Cálculo de Estatística

Questão 1 $Q = 26 + 21 + 14 + 10 + 8 + 7 + 6 + 5 + 4 + 3 + 2 + 1 = 140$

938 alunos

417 alunos (usuários do ensino fundamental)

521 alunos (ensino médio)

204 alunos 1º ano

179 " 2º ano

138 " 3º ano

513 mulheres

Investiga 40 alunos

20 de período matutino

20 de período vespertino

a) "Moda" variável

Média = $11 + 12 + 13 + 14 + 15 + 16 + 17 + 18 + 19 + 20 + 21$

Idade	Valor - média	Quadrado
11	-5	25
12	-4	16
13	-3	9
14	-2	4
15	-1	1
16	0	0
17	1	1
18	2	4
19	3	9
20	4	16
21	5	25

Média = 16

Desvio Padrão = 3,16

Variação = 10

Desvio Padrão = 3,16

Média = 16

Desvio Padrão = 3,16

Variação = 10

Desvio Padrão = 3,16

LONGA>

34. N: É, eu ia fazer... A primeira pra montar o gráfico, organizando...

Então, eu ia ver quantas pessoas preferem não declarar e, assim, porém, não vai funcionar, porque os únicos que se repetem muito é o 'Prefere não declarar', o 'Não sabe' e o de renda 1.108,38. Os outros valores não. Então, eu ia acabar perdendo tempo. Então, nesse caso, eu vou começar fazendo a média. Eu vou somar todos os valores dos 40 sujeitos, a renda familiar deles, e... o resultado eu vou... Continuo...<SUSSUROS>

35. N: É, no caso do 'Prefere não declarar' e 'Não sabe', eu vou colocar valor 0, mas eu vou colocar, porque é o que tem na amostra, então, eu tenho que usar...

36. N: Porque, como eles foram entrevistados, eu não posso ignorar o que foi falado por eles, porque senão, a minha conta não vai ter sentido, porque... eu não vou ter entrevistado as 40 pessoas.

37. N: Então, eu comecei fazendo a média... Média. <SUSSURROS> E como são muitos valores, eu vou tentar montar, fazendo a conta com o resultado... o gráfico... Média <SUSSURROS> <PAUSA MUITÍSSIMO LONGA>

38. N: ... 4, 7, 14, mais... Aprendi!... 1.108,38 + ... <SUSSURROS>

antes... Peraí, pra eu não me confundir, vou colocar mais esse valor: 1.108,38 +... mil... <SUSSURROS> <PAUSA MUITÍSSIMO LONGA>

39. N: Até aqui... 91.116,94... É... Tá... Em vez de eu coisa.... Ai, meu Deus... 4.623... Então, a média deu...

40. N: Então, a média eu vou conferir, porque são muitos valores, eu posso ter errado. Então, a média é 2.909... Não, é 2 milhões, 909.458, essa é a média. Agora eu vou conferir... <SUSSUROS> <PAUSA MUITÍSSIMO LONGA> Até aqui, foi 30.508,62, 31.568,08. <SUSSUROS> <PAUSA MUITÍSSIMO LONGA>

41. N: Parabéns, não bateu! Ai, meu Deus. <PAUSA> Mas eu acho que esse tá certo, porque esse eu fiz parando. O outro eu acho que eu somei duas vezes, daí eu chequei nesse resultado.

42. N: Não mudou tanto, mas mudou. Mas acho que eu vou conferir de novo... Mudou pouca coisa, mas mudou... Vamos ver.

<SUSSURROS> <PAUSA MUITÍSSIMO LONGA>

43. N: Se eu errar esse, eu erro todos. <SUSSUROS> <PAUSA MUITÍSSIMO LONGA>

အိတ်ကုသိုလ်

$$\rightarrow \text{Media} = 25 + 16 + 9 + 4 + 1 + 0 + 1 + 4 + 9 + 16 + 25 = 10$$

4

$$\sqrt{10} = 3.16$$

കുറിപ്പ്

é necessária interação que uma variável de dados pode ser explorada, através dos enunciados. Diante disso, julga-se, como a mesa, não necessariamente explorados sem a criação dos dados. Usam-se, neste

Question 2

Renda keluarga

$$\text{Média aritmética} = 1.108,38 + 1500 + 2521 + 0 + 2800 + 1.108,38 + 3500$$

$$+ 4714 + 1,108.38 + 22200 + 4800 + 5100 + 1,108.38 + 1,108.38 + 1500 + 0 + 2150$$

$$+ 3400 + 0 + 1,108.38 + 3000 + 1,108.38 + 0 + 3820 + 2,316.46 + 2,600$$

$$+ 1200 + 1108.38 + 30000 + 1300 + 0 + 3.332814 + 46.23 + 4100 + 5000 + 3400$$

$$0 + 0 \leq 6 + 1 \text{ for } 0 \leq 1 + 1$$

$$1100 + 30 + 100 + 330 + 40$$

W-116 298 33.

U₂
R1 = 196.34

W 3 909.450

$$M = \infty, 158$$

radius = 2,904,458

$\tau_{\text{obs}} = 1108.38$

[illegible]

$-1,108,38 - 1,108,38 - 1,108,38 - 1,108,38 - 1,108,38 - 1,200 - 1,300$

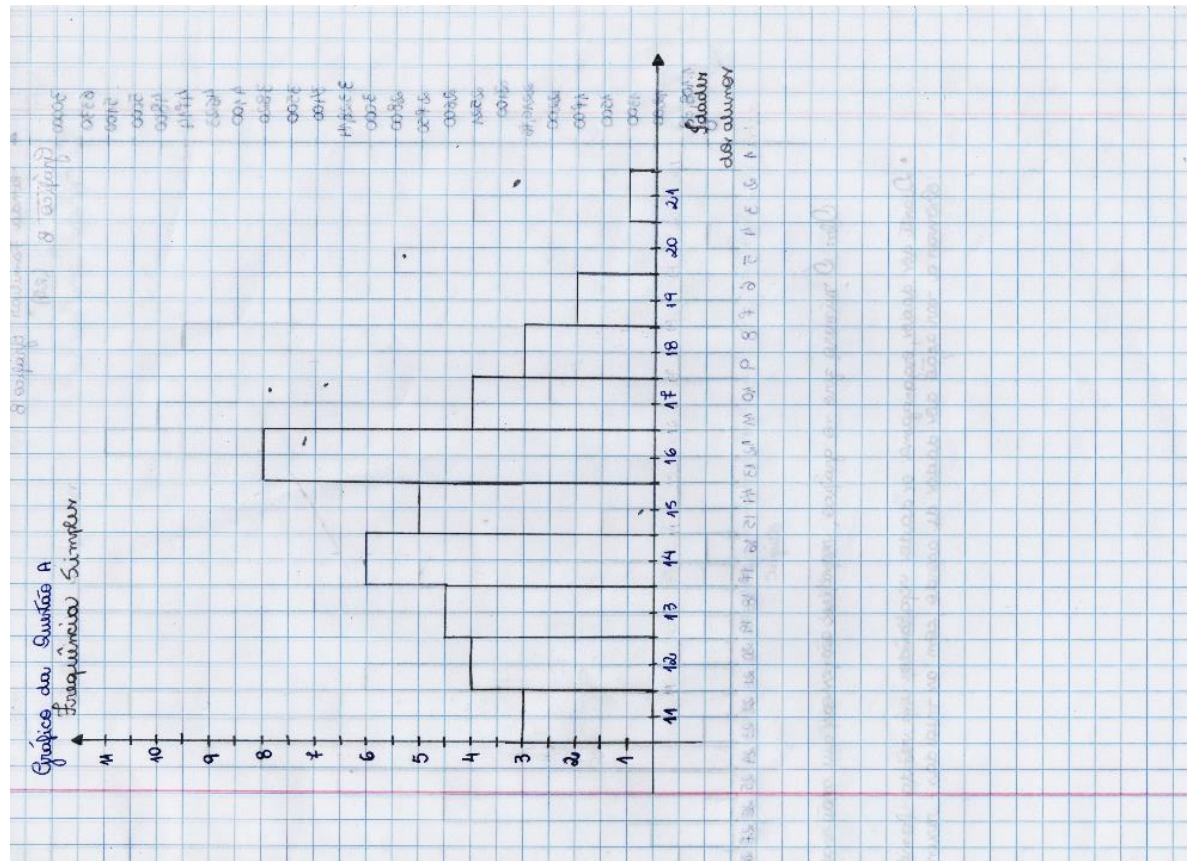
-1500 - 1500 - 1700 - 2200 - 2216.76 - 2400 - 2600 - 2750

$$= 2800 - 3000 - 33.28, 14 - 3400 - 3500$$

- 4800 - 5000 - 5100 - 6330 - 30000

Median = 1700 + 2

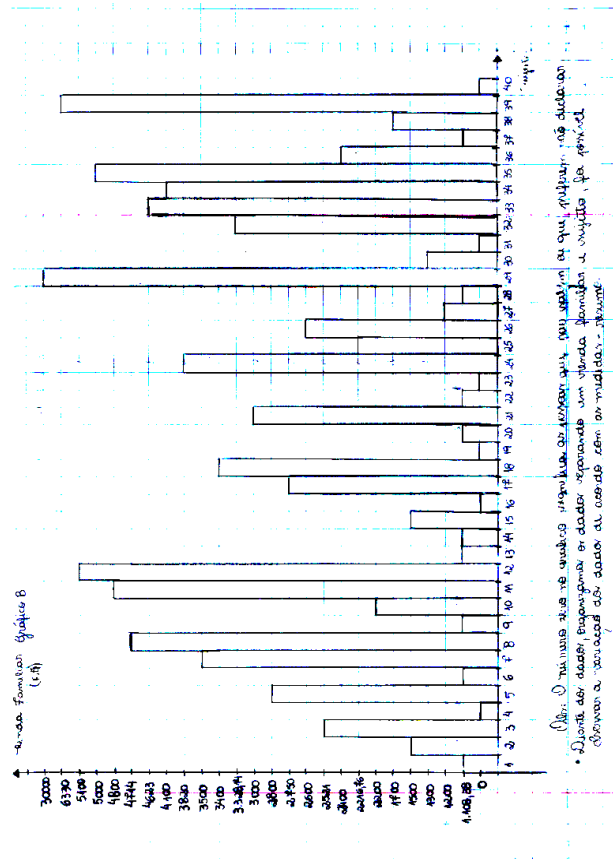
2 2



44. N: Deu o primeiro valor, então eu vou deixar esse. <PAUSA> Tá...
45. N: Aí, desculpe a demora... Na questão 'b', eu demorei, porque são muitos valores e eu conferi duas vezes. Eu já achei a média. Eu somei todos os valores e dividi pela quantidade, que é 40. A moda é 1.138, porque é o número que aparece com maior frequência... Nove, ele aparece nove vezes, então a moda é 1.108,38. <PAUSA>
46. N: Tá... a moda... A mediana... A mediana eu vou ter que organizar do menor pro maior e pegar os dois números do meio, somar e dividir por 2. Nesse caso, eu vou ter que pegar os dois, não vai ser que nem na questão 'a', que eu tinha um número ímpar. Então, foi mais simples de encontrar o meio da... da... da variável... dos resultados. E, nessa, eu vou ter que organizar e somar, porque é 40, então, eu vou precisar pegar dois valores, somar e dividir por 2. Então, eu vou ter a mediana. <SUSSURROS> <PAUSA LONGA>
47. N: Certo, então, primeiro vai ser os '0's'... Mas, na verdade, eu não vou somar, é só pra ir contando, 0, 0, 0... 1, 2, 3, 4, 5... Não. 1, 2, 3, 4, 5, 6... 1, 2, 3, 4, 5, 6... Tá, então, agora vamos ver o menor número. Vamos ver os menores números. Eu vou marcar o primeiro número da fileira... 1.108, nesse caso... Aí, 1.108 aparece oito vezes. Na verdade, acho que o meio '20 eu nem vou precisar dividir... 1, 2, 3, 4, 5, 6, 7, 8, 9. <SUSSURROS> <PAUSA LONGA>
48. N: 1, 2, 3, 4, 5, 6, 7, 8, 9... Tá, então, agora 1.108, a que é menor, 1.500... 1.200, 1.300... Então, aqui vai ser 1.200, agora vai ser 1.500. Então, calma, eu vou riscar todos que eu já fiz... Aí, aí tem 1.500 de novo, eu já fiz. <PAUSA LONGA>
49. N: Já foi, já foi, já foi, já foi... Agora é 1.700... 1.500 duas vezes. <SUSSURROS> 2.200 é o menor. Daí, depois, 2.400, corta. Deixa eu ver se tem mais algum '2.000'. <SUSSURROS> 2.600 não, é 2.521... <SUSSURROS> Daí, 3.328,14, risca. <SUSSURROS> <PAUSA MUITÍSSIMO LONGA>
50. N: A mediana vai ser 1.700. Vamos só ter certeza. <PAUSA LONGA>
51. N: Eu vou pegar os dois valores do meio e vou dividir por 2... 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19... Eu vou pegar os dois valores do meio... 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19... 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19...

52. N: Agora, eu vou pegar os dois valores do meio e dividir por 2...
<SUSSUROS> Então, a mediana desses valores vai ser 1.950.
53. N: Agora, eu vou fazer média, moda, mediana... variância. <PAUSA MUITÍSSIMO LONGA>
54. N: É, eu tava fazendo a variância, mas ela é um pouco mais demorada, então, agora eu vou montar o gráfico, porque, no enunciado, pede pra usar uma ou mais medidas-resumo, e eu determinei três. Então, eu acho que não é necessário a variância nesse caso, já que os números são bem grandes e vai ser um pouco mais demorado.
55. N: Então as três... Eu determinei a média, a moda e a mediana. Agora eu vou montar o gráfico...
56. N: Então, no gráfico da questão 'b', eu tô pensando em usar a média, a moda e a mediana, porque são muitos valores pra mostrar no gráfico. Então... <PAUSA LONGA>
57. N: Eu demorei um pouco na questão 'b'. Então, basicamente, o que eu fiz foi tentar organizar, mas seria perda de tempo... E eu perdi muito tempo, de qualquer jeito. <RISOS>
58. N: Na questão '2', quero dizer, 'b', depois eu fiz a média, a média aritmética, somei todos os valores e dividi por 40, então eu demorei um pouquinho. Depois, eu fiz a moda. A moda é o número que mais se repete, então, a moda é 1.138... 1.108,38. Depois, eu fiz a mediana e, agora, eu vou montar o gráfico.
59. N: O gráfico eu vou tentar mostrar isso, mas... Não, aí... Então ficou isso.
60. N: A questão 'b'. Então, como é que eu posso fazer?... Eu acho que eu vou fazer... <PAUSA LONGA> Eu acho que eu vou fazer a renda familiar e o número de pessoas... É que tem, tipo, poucos que...
61. N: É, nesse caso, no gráfico 'b' eu vou fazer de setores: eu vou colocar a parte que não precisa declarar e a moda, que vai ser o maior valor, porque é complicado fazer de setores com esse valor pequeno...
62. N: Não, não é que os valores sejam pequenos, é que pouca coisa se repete... Aí, já sei! A renda familiar eu posso fazer entre... hum... Acho que eu vou colocar aqui os valores entre 1.000 e 1.200... Renda familiar... E são 40 pessoas.

<TERMINA O TEMPO DA AULA



SEGUNDA SESSÃO [01:11:57] – 29/11/2018 (Nesta sessão, L participou.)	
Produção oral	Produção escrita
63. N: Vou te explicar o que eu fiz, daí você vê se concorda. Na primeira... <RELÊ O ENUNCIADO DA PRIMEIRA QUESTÃO> 64. L: Aham. 65. N: Aí o que que eu fiz: eu peguei todas as idades, eu somei e dividi por 11... 66. L: Que é a média? 67. N: Aham... Daí, deu esse valor: 16. 68. L: Aham. 69. N: Aí, a moda é o número que mais se repete, da frequência simples. É o '16' também. 70. L: Aham. 71. N: A mediana, se você organizar, porque, tipo, já tá em ordem, é 16, porque 16 é o meio. 72. L: Hum hum. 73. N: E a variância. Daí eu fiz... 74. L: Sim... 75. N: Daí, eu peguei as idades, fiz menos a média, fiz 16 – 11... mas ficou negativo... O que que eu fiz? 76. L: Não, é, deixa eu ver. 77. N: Ah... Eu fiz, 11 – 66, 12 – 16, 13 – 16. Aí, deu negativo os valores, mas eu não somei, porque, se somar, dá 0. O que que eu faço? Eu faço isso ao quadrado. 78. L: Todos? 79. N: Todos. Aí, eu tiro a média deles... 80. L: Que é a variância. 81. N: É, a variância, que é 10. Daí, o desvio padrão é 3,16. Aí, eu fiz isso. 82. L: E o gráfico? 83. N: Aí, eu só montei o das idades, que, aí, dá pra ver a moda, né? 84. L: Aham. E é praticamente a mesma coisa na 'c', né? 85. N: Cê acha que tá certo? 86. L: Aham, eu acho que tá certo. É isso mesmo. É o que eu vou fazer também. 87. N: Aí, a questão '2' eu achei mais complicada, porque eu tive que somar tudo isso <RISOS> pra tirar a média. Daí, eu somei, peguei o	<u>Questão 3</u> <u>PAIS</u> MODA → Enximo número completo com 9 alunos MEDIANA → 2 2 2 3 <u>4</u> 4 6 8 9 = 4 <u>MAES</u> MODA → Enximo número completo com 15 alunos MEDIANA → 0 0 1 1 <u>1</u> 4 6 12 15 = 1 Diante das informações obtidas no enunciado foi possível determinar a moda e a mediana sobre a escolaridade dos pais e das mães. Não foi possível calcular a média pois não foi fornecida para isso. Também não foi possível calcular o desvio padrão também não foi fornecido.

- valor. Daí, eu dividi por 40. Aí, tá! Essa é a média. Aí, a moda eu contei e o '1.108,38' se repete nove vezes. Então é a moda.
88. L: É.
89. N: A mediana, eu peguei os números do meio, que é o '1.700' e o '2.200', somei e dividi por 2.
90. L: Hum hum.
91. N: Aí, deu 1.850. Aí, essa é a mediana.
92. L: Aham.
93. N: Aí, eu ia fazer a variância, só que ia demorar muito, porque são muitos números, e como ele falou aqui no enunciado que era... "determine uma ou mais", se eu já determinei a moda, a mediana e a média...
94. L: É. Exatamente. Vamos pra última.
95. N: Só que aqui ele tá pedindo pra explicar. O que você acha que a gente pode fazer? Essa tem que montar gráfico?
96. L: Aham.
97. N: Aí eu explico e você...
98. L: Ele disse "se quiser utilizar gráfico"... "como você evidenciaria essa comparação se quisesse utilizar gráfico? Explique o seu pensamento. É possível determinar uma medida-resumo que contribua para a análise desses dados? Qual? Como determinar? Explique o seu pensamento". Medida-resumo... O que seria? Uma média?
99. N: É. Então, porque... eu não sei como fazer o gráfico da 'b'. Eu tô muito na dúvida, porque se for olhar, ó, vai pegar o que para comparar? Porque, se tivesse alguma idade, mas só tá numerando o sujeito...
100. L: Coloca de 1 a 40, aí coloca as rendas.
101. N: De cada um?
102. L: É, coloca a renda, o sujeito, faz bem pequenininho.
103. N: É, aí eu faço o gráfico, explico, e daí você faz a 'c'.
104. L: É, vamos logo, que eu quero pegar a revisão dele. Vai fazendo aí...
105. N: Vou fazer desse lado... É, mas eu tô em dúvida. Eu coloco a renda... Nesse gráfico aqui dá pra colocar a moda.
106. L: Mas não dá pra identificar.
107. N: Você acha que dá pra fazer? Pra identificar?

Medidas de Dispersão (Variabilidade)

Variação
 Sejam $x_1, x_2, \dots, x_N$ as medidas de valores observados para uma variável qualitativa de X e $\bar{x}$ a média aritmética desses valores.
 N é a quantidade de dados

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}$$

$$\sigma^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_N - \bar{x})^2}{N}$$

Desvio Padrão
 $\sigma = \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}$

$\sum_{i=1}^n$ = variáveis de n termos
 x_i = valor individual da amostra
 x = valores médios aritméticos
 n = número de vezes da repetição

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_N - \bar{x})^2}{N}}$$

- Quanto maior o desvio padrão, maior a variabilidade dos dados, ou seja, MENOR é a regularidade dos dados.
- Quanto menor o desvio padrão, menor a variabilidade dos dados, ou seja, MAIOR é a regularidade dos dados.

108. L: Não sei se dá, não sei se ele pede isso, mas acho que é só pra colocar os dados do gráfico.
109. N: Então, ele pediu para construir um gráfico que mostra pelo menos uma dessas.
110. L: Ele disse "evidenciar a comparação". A comparação poderia... Porque eu acho que não tem...
111. N: Tá certo. Porque a moda fica nítida.
112. L: É, a moda a gente faz embaixo. A moda e todos os outros.
113. N: Eu vou ter que pegar outra folha. <RISOS> <CONVERSA PARALELA SOBRE AS PROVAS> Não sei se eu coloco as medidas na ordem... <PAUSA> Sabe o que eu vou fazer? Vou colocar do menor para o maior, porque senão fica muito sem sentido. 'Zero' tem que colocar... O ruim é que de 0 para 1.108,38...
114. L: Nossa, vai entender, né?
115. N: Por isso que eu tava... né, porque ontem eu buguei muito. Nossa, acho que eu vou colocar NS aqui, daí, eu coloco os números.
116. L: Aham.
117. N: Senão fica estranho. Você quer construir o gráfico aí?
118. L: Não, vou fazer aqui, depois eu passo a limpo.
119. N: Ele disse pra deixar os rascunhos junto com ele.
120. L: Ah, é?
121. N: É, que nem esse aqui que eu fiz ontem. <PAUSA LONGA> <CONVERSA PARALELA SOBRE ALUNOS DA TURMA> Eu conferi umas dez vezes o resultado.
122. L: É, mas tem que conferir mesmo, já pensou ter esse trabalho e perder tudo?
123. N: Aí, teve uma vez que eu conferi e não bateu... Aí, eu falei "Ai meu Deus!". E eu queria terminar ontem, porque eu quero ver a revisão dele. <PAUSA LONGA> Nossa, eu não acredito, não cabe mais na folha.
124. L: Nossa!
125. N: Cé tem branquinho aí?
126. L: Tenho.
127. N: Ninguém merece ter que refazer. O bom desse gráfico é que... Éita, e agora? E 'Não sabe'? 'Não sabe' é 0!
128. L: É.
129. N: <CONVERSA PARALELA SOBRE ALUNOS DA TURMA>



Problema:

Condição (Notas)	Nota - Média	Quadrado
7	0	0
10	3	9
6	-1	1
8	1	1
4	-3	9
7		
Média Sistemática	Média	Média
$\frac{7+10+6+8+4}{5} = 7$		

5

Média

$\frac{0+9+1+1+9}{5} = 4$

5

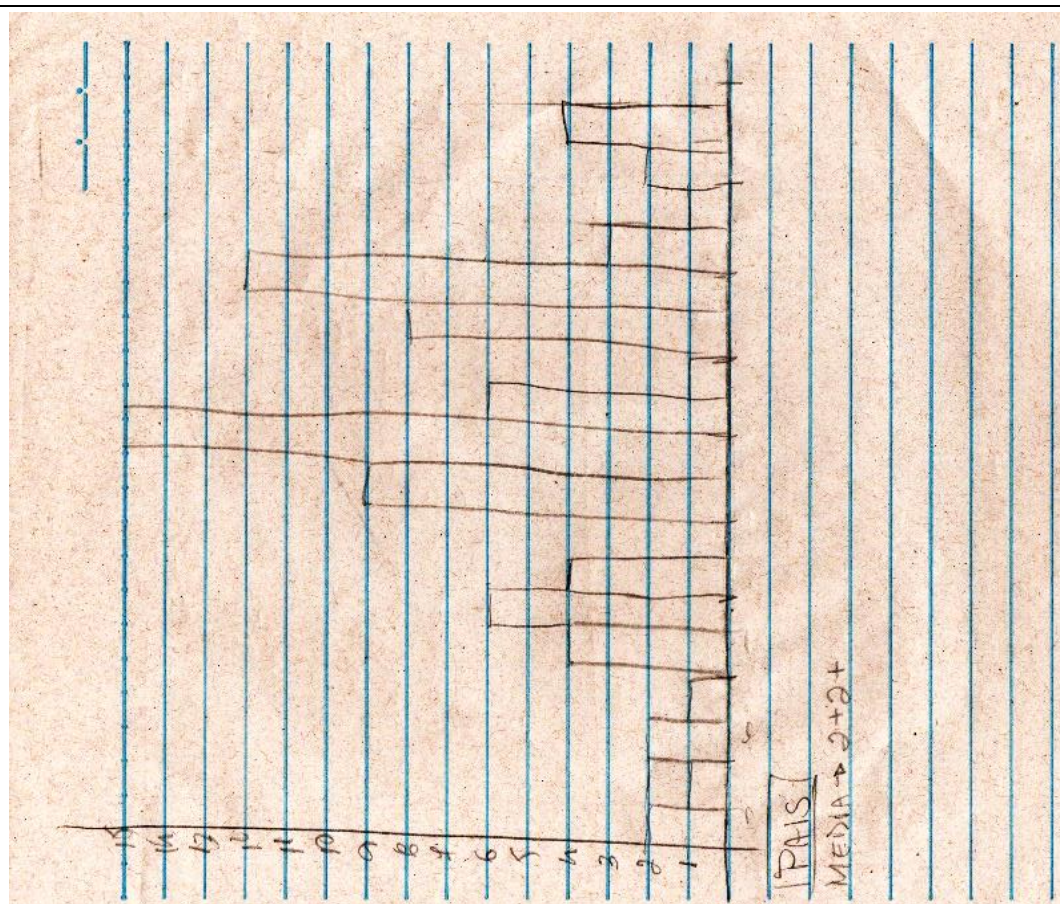
Média

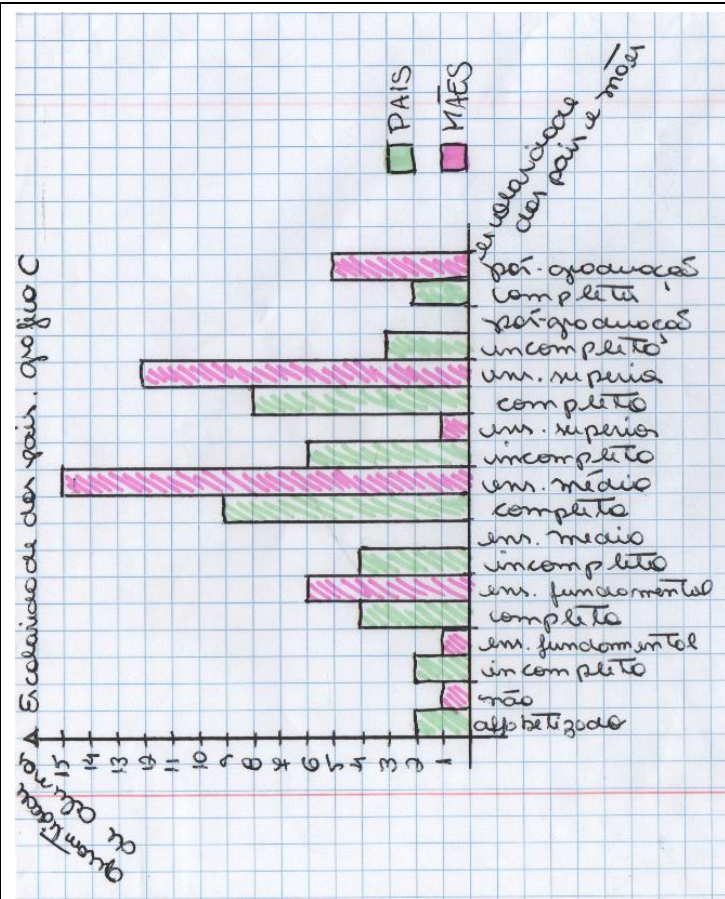
$\frac{0+9+1+1+9}{5} = 4$

Média

$\frac{0+9+1+1+9}{5} = 4$

130. L: E onde eu coloco? É na quadriculada? Eu coloquei questão '3'.
<CONVERSA PARALELA> Que mais que é? Média, mediana...
131. N: Moda, variância, desvio padrão. Tem que somar, você somou? Tem que tirar a do pai...
132. L: Claro.
133. N: Tem que somar separado, do pai e da mãe, você somou separado?
134. L: É.
135. N: Oh, ele entrevistou 40 pessoas. Dessas 40 pessoas, 15 a mãe tem o ensino médio completo.
136. L: É.
137. N: Então, a moda é ter o ensino médio completo.
138. L: É.
139. N: Agora, a média, você vai fazer...
140. L: A moda é o que mais aparece?
141. N: Isso.
142. L: Ó, ficou bom o gráfico assim...
143. N: É.
144. L: O sujeito '6' é 1.108?
145. N: É. <SUSSURROS> <PAUSA LONGA>
146. L: Mediana é sempre o do meio?
147. N: Isso, você coloca eles na ordem, daí cê acha. Daí, se o número for, tipo, 40, você vai ter que somar os dois e dividir por 2.
148. L: Entendi.
149. N: Ai, é melhor fazer em dupla do que sozinha.
150. L: <RISOS>
151. N: Ai, ontem eu fiz sozinha. <RISOS>
152. L: Por exemplo, aqui tem mais de um '2', eu coloco todos ou só um?
153. N: Coloca todos. Eu fiquei com essa dúvida ontem, mas se não colocar, na hora da conta, não bate. Ai que boba, olha o que eu fiz, acho que eu errei... Tem 414? Tem 416... 416, não, tem 4 mil?
154. L: Cadê?
155. N: Só tem 14... Eu errei, 4014, 4714...
156. L: Sabe por que eu acho que não dá pra tirar? Porque eu somo tudo e dá 40... Daí, eu divido e dá 20. A moda dá, a mediana dá.
157. 157. N: Ele falou que é normal, que tem umas coisas que não dá pra calcular. <SUSSURROS>.





158. L: Será que pode ir lá perguntar para ele?
159. N: Acho que pode, não pode sair no áudio, eu acho.
160. L: É isso mesmo que eu pensei, eu consigo tirar a moda e a mediana, a média não, ele disse que se tivesse o número de anos que as pessoas estudaram, teríamos como calcular a média disso.
161. N: A gente fica com medo de não colocar, né, depois ele pede, né? <CONVERSA PARALELA SOBRE A TURMA> Vai dar trabalho...
162. L: Eu posso fazer aqui na margem?
163. N: É muito grande, eu usei. <CONVERSA PARALELA SOBRE AVALIAÇÃO> Eu acho que eu vou falar que não fiz a variância porque dá muito trabalho.
164. L: É, realmente, ele disse que é opcional.
165. N: <CONVERSA PARALELA>
166. L: Será que pode pintar, usar marca-texto, fazer uns risquinhos?...
167. N: Acho que pode. No meu caso, não precisa... Onde eu tô, meu Deus? É muito fácil de se perder... 1.200, 1.138...
168. L: Mano, é muito grande o negócio, onde é que eu tô? Deixa eu ver.
169. N: Na minha explicação é melhor falar que 'Não sabe' é 0... Falar que... O que é que ele falou que tinha que ter nos gráficos, mesmo?
170. L: Não.
171. N: Título, fonte, esses negócios.
172. L: Pode deixar a lápis isso aqui?
173. N: Eu deixei o meu e ele não falou nada ontem. A última é o que? 8.630?
174. L: Sei lá, vou esperar você achar aí...
175. N: Meu Deus do céu... Será que ele vai entender? Será que ele vai entender que esse é do '1' e esse é do '2'? É melhor colocar no meio. Eu acho que eu não vou colocar título no gráfico, porque eu vou colocar que título?
176. L: Eu coloquei aqui...
177. N: Acho que ele entende, né? Já tá todo borrado... <PAUSA LONGA> Eu coloquei aqui que o '0' são as pessoas que não sabem ou não querem declarar, porque é melhor que não colocar nada, né?
178. L: É porque senão ele vai pensar "por que zero?".
179. N: Ai, tomara que ele entenda.
180. L: O Cassio que me desculpe...

181. N: É a variância aí? 182. L: Não, é o desvio padrão... Tem que ler as questões? Pra gravar? 183. N: Você tem que ler a '3', porque essas eu li... 184. L: Não, porque é pra explicar... Eu leio o gráfico? 185. N: Não, você fala só o que você fez... <SUSSUROS> 186. L: Será que se eu deixar a lápis aqui, tudo bem? 187. N: Acho que sim. <TOCA O SINAL> Aí, meu Deus, vai começar a aula... 188. L: Eu tirei a média, eu somei tudo, daí, eu dividi por 40. Aí, eu fiz a moda, que é o que tem mais vezes, que, aí, no caso, é o '1.108,38' e eu fiz a mediana, somando os dois números do meio, que é $1.700 + 2.200$, que é 1.950. Daí, eu fiz o gráfico, né? Que o sujeito, a renda, aí eu fiz o gráfico. Tá certo, né? O que você acha que eu explico? Lê aqui e você o que eu preciso explicar. 189. N: Fala o que você fez na '3'. 190. L: <LÊ O ENUNCIADO DA QUESTÃO 'C'> Então, eu calculei a moda e a mediana, que foi o que eu consegui. A média não deu, porque não tem número pra calcular a média e eu calculei separado dos pais e das mães. A dos pais, a moda foi o 'ensino médio completo', com nove alunos, os pais de alunos, na verdade, a mediana foi 4, porque eu coloquei na ordem dos valores dos pais de alunos. As mães, também a moda foi o 'ensino médio completo', só que com 15 pais... mães de alunos, mer* <PALAVRÃO> Tá tudo errado o que eu falei. <RISOS> 191. N: Por quê? <RISOS> 192. L: Peraí, volta tudo, eu calculei separado, a moda foi o 'ensino médio completo', com nove pais de alunos, e a mediana foi igual a 4. Das mães, o 'ensino médio' foi 15 mães de alunos e a mediana eu também coloquei mais ou menos em ordem, e o resultado foi 1. A média não foi possível calcular, porque não tem número. Assim, não foi possível calcular a variância e o... e o desvio padrão eu também não consegui... Não consegui, não! Não tem como, porque não tem valor pra isso. O gráfico eu coloquei na vertical a quantidade de alunos, na horizontal eu coloquei a escolaridade dos pais, em verde, e, em rosa, a das mães, informando a quantidade de alunos nas colunas. E só. 193. N: Ele falou que tudo que tiver de rascunho é pra deixar... Será que tem alguma coisa pra explicar nós dois?	
---	--

194. L: <SUSSURROS>
195. N: Eu não sei o que eu coloco...
196. L: É medidas de resumo, né?
197. N: É, acho que tá bom, né?
198. L: É.
199. N: Esse é o dele, deixa o gravador aí que ele que vai desligar, a gente avisa ele.
200. L: Ai, acabou, que bom, né?
201. N: Não precisava falar, né?

Protocolo 4 – GRUPO IV

W: SEXO MASCULINO, 17 ANOS
V: SEXO MASCULINO, 17 ANOS
A: SEXO MASCULINO, 17 ANOS

PRIMEIRA SESSÃO [01:26:58 + 00:29:18 = 01:56:16 (intervalo de 20 minutos entre a primeira e a segunda parte)] – 29/11/2018																																																																																																																			
Produção oral	Produção escrita																																																																																																																		
A: <LÊ O ENUNCIADO GERAL> W: <LÊ O ENUNCIADO DO ITEM A> 1. A: Tem que tirar isso aqui: média, moda, mediana, quartis, desvio padrão... 2. V: Desvio padrão a gente tem no caderno. 3. A: Ó, aqui tem média, moda, mediana, quartis, variância, desvio padrão e gráfico. 4. V: Aqui também. 5. A: Tem no livro? 6. V: Tem no livro! Tô falando! <PAUSA> Qual que é pra eu fazer aí? 7. A: Eu vou fazer a 'a'. 8. V: Não, qual que é pr' eu fazer aí? 9. W: É pra fazer média, moda, mediana... 10. A: Tem tudo aqui no livro. Ó, tem que medir tudo. Não, média... 11. V: Média tem só no caderno. 12. W: Preciso de um livro. 13. A: Não, ó, média... Calma aí, aqui tem moda, tem média e moda. No caderno, tá muito simples... A moda é, tipo, a idade que mais aparece. 14. W: A que mais aparece? 15. A: É, 16 anos. Mas, primeiro, tem que tirar a média. 16. W: A média. 17. A: É, primeiro, precisa calcular a média. 18. W: Tem que colocar alguma coisa aqui para multiplicar, com esse aqui, com esse aqui, com esse aqui, aí divide por 40 no total. 19. A: Peraí, tem aqui, ó. Tem no caderno. <PAUSA> Esse com esse, esse com esse, esse com esse, esse com esse... 20. W: Aí soma tudo. E divide por 40. <INTERRUPÇÃO DE OUTRO PROFESSOR E ALUNOS NA SALA DE LEITURA>	Atividade de Estatística Em uma escola pública de Santo André, um grupo de alunos do terceiro ano do Ensino Médio, realizando uma pesquisa escolar, resolveu investigar a opinião dos colegas sobre a política de cotas para o ingresso em universidades federais. Essa escola possuía 938 alunos, sendo 417 estudantes do Ensino Fundamental, estudando no período vespertino (6º ao 9º ano) e 521 no Ensino Médio, estudando no período matutino (204 no 1º ano, 179 no 2º ano e 138 no 3º ano). Desses 938 alunos, 513 são mulheres. Esse grupo decidiu entrevistar 40 alunos, sendo 20 (dez homens e dez mulheres) do período matutino e 20 (dez homens e dez mulheres) do período vespertino. Elaborou um questionário com doze perguntas, cujas respostas podem ser observadas no banco de dados que segue. Analise os resultados utilizando as representações que julgar adequadas e as medidas que julgar adequadas. Lembre de sempre explicar ao máximo seu pensamento. a) Com relação à variável "idade", os alunos forneceram as seguintes respostas: <table><tr><th>Idade (anos)</th><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td><td>Total</td></tr><tr><th>Freq. simples</th><td>3</td><td>4</td><td>4</td><td>6</td><td>5</td><td>8</td><td>4</td><td>3</td><td>2</td><td>0</td><td>1</td><td>40</td></tr></table> Como você explica a variação das idades utilizando as ferramentas estatísticas que aprendeu: construa uma ou mais representações gráficas, explicando-as. Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de idades do grupo de 40 alunos pesquisados. b) Em relação à renda familiar, os alunos obtiveram as seguintes respostas <table><tr><th>Sujeito</th><th>Renda Familiar (R\$)</th><th>Sujeito</th><th>Renda Familiar (R\$)</th><th>Sujeito</th><th>Renda Familiar (R\$)</th><th>Sujeito</th><th>Renda Familiar (R\$)</th></tr><tr><td>1</td><td>1108,38</td><td>11</td><td>4800</td><td>21</td><td>3000</td><td>31</td><td>Não sabe</td></tr><tr><td>2</td><td>1500</td><td>12</td><td>5100</td><td>22</td><td>1108,38</td><td>32</td><td>3328,14</td></tr><tr><td>3</td><td>2521</td><td>13</td><td>1108,38</td><td>23</td><td>Prefere não declarar</td><td>33</td><td>4673</td></tr><tr><td>4</td><td>Prefere não declarar</td><td>14</td><td>1108,38</td><td>24</td><td>3820</td><td>34</td><td>4100</td></tr><tr><td>5</td><td>2800</td><td>15</td><td>1500</td><td>25</td><td>2216,76</td><td>35</td><td>5000</td></tr><tr><td>6</td><td>1108,38</td><td>16</td><td>Não sabe</td><td>26</td><td>2600</td><td>36</td><td>2400</td></tr><tr><td>7</td><td>3500</td><td>17</td><td>2750</td><td>27</td><td>1200</td><td>37</td><td>1108,38</td></tr><tr><td>8</td><td>4714</td><td>18</td><td>3400</td><td>28</td><td>1108,38</td><td>38</td><td>1700</td></tr><tr><td>9</td><td>1108,38</td><td>19</td><td>Prefere não declarar</td><td>29</td><td>30000</td><td>39</td><td>6330</td></tr><tr><td>10</td><td>2200</td><td>20</td><td>1108,38</td><td>30</td><td>1300</td><td>40</td><td>Não sabe</td></tr></table>	Idade (anos)	11	12	13	14	15	16	17	18	19	20	21	Total	Freq. simples	3	4	4	6	5	8	4	3	2	0	1	40	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	1	1108,38	11	4800	21	3000	31	Não sabe	2	1500	12	5100	22	1108,38	32	3328,14	3	2521	13	1108,38	23	Prefere não declarar	33	4673	4	Prefere não declarar	14	1108,38	24	3820	34	4100	5	2800	15	1500	25	2216,76	35	5000	6	1108,38	16	Não sabe	26	2600	36	2400	7	3500	17	2750	27	1200	37	1108,38	8	4714	18	3400	28	1108,38	38	1700	9	1108,38	19	Prefere não declarar	29	30000	39	6330	10	2200	20	1108,38	30	1300	40	Não sabe
Idade (anos)	11	12	13	14	15	16	17	18	19	20	21	Total																																																																																																							
Freq. simples	3	4	4	6	5	8	4	3	2	0	1	40																																																																																																							
Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)																																																																																																												
1	1108,38	11	4800	21	3000	31	Não sabe																																																																																																												
2	1500	12	5100	22	1108,38	32	3328,14																																																																																																												
3	2521	13	1108,38	23	Prefere não declarar	33	4673																																																																																																												
4	Prefere não declarar	14	1108,38	24	3820	34	4100																																																																																																												
5	2800	15	1500	25	2216,76	35	5000																																																																																																												
6	1108,38	16	Não sabe	26	2600	36	2400																																																																																																												
7	3500	17	2750	27	1200	37	1108,38																																																																																																												
8	4714	18	3400	28	1108,38	38	1700																																																																																																												
9	1108,38	19	Prefere não declarar	29	30000	39	6330																																																																																																												
10	2200	20	1108,38	30	1300	40	Não sabe																																																																																																												

21. A: Aqui, ó, cê tem que pegar $11 \times 3, 12 \times 4 \dots$ Daí, somar tudo. <INTERRUPÇÃO DE OUTRO PROFESSOR E ALUNOS NA SALA DE LEITURA> 22. W: Daí, vai tirar a média. 23. A: A moda é 16... Peraí, como é que calcula a moda? 24. W: Mas primeiro não é a média? 25. V: A gente vai tirar agora a média. <PAUSA> Nossa! 26. W: O que você tá fazendo? 27. V: Vou somar tudo. 28. A: Vai falando pra mim. 29. W: $11 \times 3 + 12 \times 4 + 13 \times 4 + 14 \times 6 + 15 \times 5 + 16 \times 8 + 17 \times 4 + 18 \times 3 + 19 \times 2 + 20 \times 0 + 21 \times 1$. 30. A: Igual... 31. W: Agora tem que somar. 32. A: Tem que pegar a calculadora. <PAUSA> 33. W: Agora a gente vai usar a calculadora. <PAUSA LONGA> 34. A: Agora tem que somar tudo. $33 + 48 + 52 + 84 + 75 + 128 +$ <PAUSA> $+ 68 + 54 + 38 + 21 \dots$ dividido por 40. 35. V: $33 + 48 + 52 + 84 + 75 + 128 + \dots$ peraí... $+ 68 + 54 + 38 + 21 \dots$ 601. 36. A: Deu 601 só? Mas tinha que ter dividido. 37. V: Por 40. 38. A: É. 39. V: 15,25. 40. W: É a média de idade. A média é 15... 15 anos. 41. A: É 15,25. 42. W: É 15, vamos arredondar... Tem que fazer gráfico. 43. A: Primeiro é a moda... A moda é 16. 44. W: Não tem nenhuma conta? 45. V: Não, é só 16. 46. W: É 16... Ah, por causa desse 8. 47. V: Agora é? 48. A: Quartis... Não, mediana, mediana, mediana... 49. W: É mediana. 50. V: Mediana ou média aritmética. 51. A: <LÊ A DEFINIÇÃO DE MÉDIA ARITMÉTICA DO LIVRO DIDÁTICO> <PAUSA> Ah... Tem que pegar a frequência absoluta. O ponto médio de cada classe. E a frequência simples. Frequência	Como você explica a variação das rendas declaradas utilizando as ferramentas estatísticas que aprendeu: construa uma ou mais representações gráficas, explicando-as. Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de ~~dados~~ <i>rendas familiares</i> do grupo de 40 alunos pesquisados. c) Em relação à ~~escolaridade~~ <i>escolaridade</i> dos pais, os alunos elaboraram a tabela a seguir. A tabela construída simplifica o banco de dados. Como você evidenciaria essa comparação se quisesse utilizar gráficos? Explique seu pensamento. É possível determinar uma medida-resumo que contribua com a análise desses dados? Qual? Como determinar? Explique seu pensamento. <table><tr><th>Escolaridade</th><th>do pai</th><th>da mãe</th></tr><tr><td>não alfabetizado</td><td>2</td><td>1</td></tr><tr><td>ens. fundamental incompleto</td><td>2</td><td>1</td></tr><tr><td>ens. fundamental completo</td><td>4</td><td>6</td></tr><tr><td>ens. médio incompleto</td><td>4</td><td>0</td></tr><tr><td>ens. médio completo</td><td>9</td><td>15</td></tr><tr><td>ens. superior incompleto</td><td>6</td><td>1</td></tr><tr><td>ens. superior completo</td><td>8</td><td>12</td></tr><tr><td>pós-graduação incompleta</td><td>3</td><td>0</td></tr><tr><td>pós-graduação completa</td><td>2</td><td>4</td></tr><tr><td>Total</td><td>40</td><td>40</td></tr></table>	Escolaridade	do pai	da mãe	não alfabetizado	2	1	ens. fundamental incompleto	2	1	ens. fundamental completo	4	6	ens. médio incompleto	4	0	ens. médio completo	9	15	ens. superior incompleto	6	1	ens. superior completo	8	12	pós-graduação incompleta	3	0	pós-graduação completa	2	4	Total	40	40
Escolaridade	do pai	da mãe																																
não alfabetizado	2	1																																
ens. fundamental incompleto	2	1																																
ens. fundamental completo	4	6																																
ens. médio incompleto	4	0																																
ens. médio completo	9	15																																
ens. superior incompleto	6	1																																
ens. superior completo	8	12																																
pós-graduação incompleta	3	0																																
pós-graduação completa	2	4																																
Total	40	40																																

	109. A: Por quê? 110. W: Desvio padrão tem que fazer a raiz. Aqui, parça, isso aí é a variância. 111. A: Tem que fazer a variância? 112. W: Depois de fazer a variância, é que a gente saca... Tem que fazer a variância. Depois que a gente tira a raiz dele. Aqui, ó, raiz de 10,95. Aí, dá o desvio padrão. Não tem que apagar, não. A gente fez certo, dá... 3,30. 113. A: 3,3? Aproximadamente 3,3. Sei lá, vou colocar aqui. 114. W: Vamos pra amplitude. Ou quer voltar pro quartis? 115. A: Mas a gente não achou os quartis. <PAUSA LONGA> 116. W: Amplitude, eu não tô achando amplitude. <PAUSA LONGA> 117. A: Vamos pros gráficos. 118. V: Pega a régua aqui. <RISOS> <CONVERSA PARALELA> 119. W: Vamos fazer um bagulho bonito. 120. A: Tem que ter 21 quadradinhos. 121. V: Vamos fazer meio centímetro cada. 122. W: Vamos começar a chutar. 123. V: 1, 2, 3, 4, 5, 6, 7, 8... <SUSSURROS> 124. A: Mas vai ter que explicar o gráfico. 125. W: Explica a média. 126. A: A média é mais fácil... A média, a moda e a variância... e o desvio padrão... 11, pode colocar '11' aí... Não, pera, 11 mulheres no período matutino. 127. W: Aqui tem homens e mulheres. 128. A: Tem 20 homens e 20 mulheres. 129. V: Aqui embaixo... Vamos fazer por idade. 130. A: Aqui não tem quantas mulher tem 18 anos, quantas mulher tem 19, quantos homens... 131. W: Faz assim, coloca o número de homens. Vai, quantas crianças tem 11 anos. Coloca o '11' e o '3'. 132. W: Vamos chamar o Cassio. Não dá pra fazer antes de confirmar. 133. A: Vamos fazer o gráfico. Você já fez? Quantos quadradinhos você já fez? Nossa, você fez 36 quadradinhos! Tem 11. <RISOS> Pra que isso tudo? 134. W: Ele é doido. <RISOS> 135. A: Olha aqui: você vai começar do '11'... As meninas têm lápis de cor... <PAUSA MUTÍSSIMO LONGA> <CONVERSA PARALELA>
--	--

<RISOS> <SUSSURROS> 136. A: Agora, tem de fazer a 'b'. Vai lendo aí a 'b'. Vê aí o que precisa fazer. 137. W: Mesma coisa: média, mediana, moda, variância, quartis, desvio padrão. Escolhe uma. Qual você quer fazer? Vamos fazer a média? 138. A: Não tem como pular, tem que fazer tudo isso. Vamos fazer o gráfico. 139. W: Gráfico e mediana. Não tem como fazer gráfico e mediana. 140. V: É só um gráfico que tem que fazer. 141. A: Eu sei, mas, se tiver que fazer o desvio padrão, ó só o tanto de coisa que vai ter que fazer. Olha aqui, vamos ver o primeiro. '1.108,38' tem 1, 2, 3, 4, 5, 6, 7, 8, 9. Nove tem o mesmo salário. Depois tem '1.500', tem 1, 2, 3. 142. W: O Cassio devia ter explicado. 143. A: Mas foi... Vamo lá, 1, 2, 3, 4, 5, 6, 7, 8, 9, tá certo, é isso aí. <PAUSA MUTÍSSIMO LONGA> <CONVERSA PARALELA> 'Prefere não declarar!' Como assim, 'Prefere não declarar'? Vai colocar o quê? 0. 144. W: É o '0'... Ah, entendi o que você tá fazendo. Aí divide por 40? 145. A: É, divide por 40... 1.200 <SUSSURROS>... 4.714. 146. W: '4.714' uma vez só? 147. A: É, '4.714', '2.200', vamos ver... só um? 148. W: Só um. '4.800' também é só um. 149. A: '5.100'? 150. V: Só um. 151. A: '1.500' já foi. 152. W: '2.750'. <SUSSURROS> 153. A: Só um também. 154. V: '1.108' já foi... <SUSSURROS> <SEGUEM LENDO A TABELA> 155. A: E isso aqui é só a média... Aí, tem a mediana, a mediana é fácil, a moda... 156. W: A moda também é fácil. 157. A: A moda é o que repete mais? 158. W: A moda é o que parece mais. E a amplitude. 159. A: Lê aí. 160. W: <LÊ A TABELA DO ITEM A> 161. A: Acabou? 162. W: Dividido por 40.	
--	--

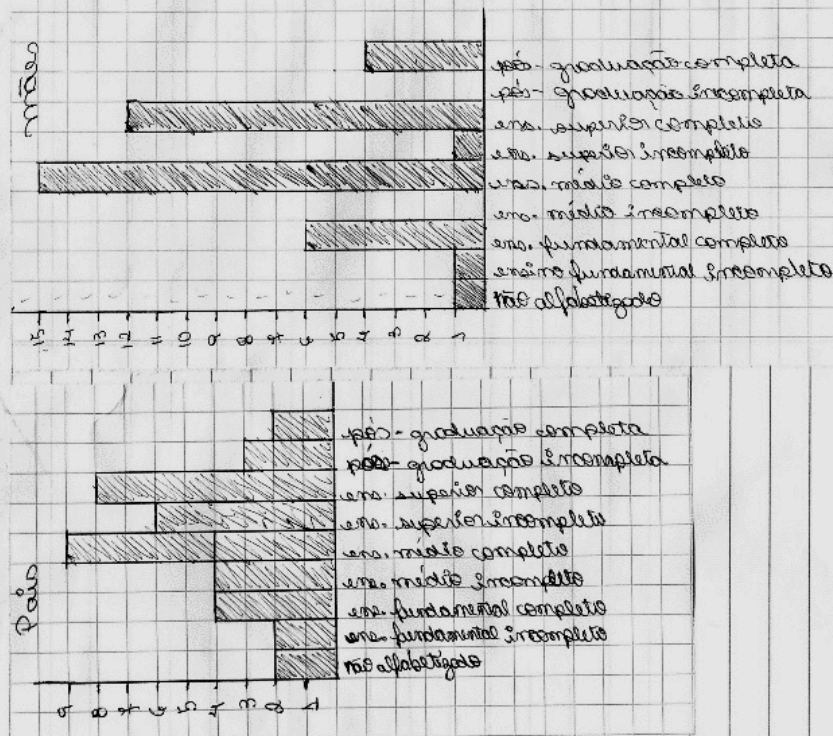
	163. A: Dá 2.909... 2.909,45 aproximadamente, é 2.909,4545. 164. V: A moda? 165. A: Moda acabou. Média. 166. W: A moda é 1.108,38. Agora, a mediana. 167. A: Tem que ver o menor e o maior. 168. W: O menor é '1.108', o maior é '30.000'. 169. A: Tem que colocar em ordem, parça. Se não colocar em ordem, não dá pra achar a mediana. 170. W: <RELÊ A TABELA DO ITEM A> 171. A: Agora, tem que fazer a média. 1, 2, 3, 4, 5, 6, 7, 8... peraí, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18,19, 20, 21, 22, 23, 24, 25 é o número '13' ..., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, é o '3.000'. 172. W: A mediana não é o primeiro e o último... 173. A: Não. A média é o '3.000'. 174. V: É o meio. 175. A: A mediana é o '3.000'. 3.000. 176. V: Não tem que fazer o gráfico da mediana? 177. A: Tem. 178. V: Vai fazendo que eu vou traçando... 179. W: Vou beber água. <PAUSA LONGA> <SUSSURROS> 180. A: Quantos salários tem mesmo? 181. V: Salário? 182. A: 25. Tem 25.... <PAUSA LONGA> Aqui vai até 30.000. Marca lá!.... O 'mil' é aqui, o parça! <RISOS> 183. V: Tem menos que mil ou não? 184. A: Cé tá chapando! 185. V: Tem menos que mil ou não? 186. A: Peraí, peraí, que tem que ver qual é o salário e quantas pessoas tem. Tem igual de pessoas que tem o salário. 187. W: Aqui vai os salários... 188. A: Parça, peraí, vamos fazer o gráfico de pizza... Não, não dá, gráfico de pizza também é muito ruim, mano. Não, tem que especificar quantas pessoas tem. 189. V: Acho que é de barras. 190. W: E o que vai escrever na barra? 191. A: Não pode ser... Não dá... 192. W: Aqui, o salário e, aqui, as pessoas.
--	---

193. A: Tem que ser quarenta quadradinhos. Mas vai só até o '30'.
 194. W: Faz outro, parça.
 195. A: Tem que fazer outro.
 196. W: Por que 40? São 34 pessoas, porque seis não declararam.
 197. A: Calma, deixa eu pensar um baguio... Vou falar com o Cassio.
 Peraí. <PAUSA LONGA> <CONVERSA PARALELA> <RETORNA>
 198. A: Vamos fazer de pizza mesmo.
 199. W: Isso, a gente faz de tanto a tanto, de tanto a tanto e pinta.
 Gráfico tem que ser pintado.
 200. A: Calma aí, eu vou colocar de 1.000 a 3.000, de 3.000 a 6.000 e de 6.000 a 7.000.
 201. W: de 6.000 a 30.000.
 202. A: Não, calma aí, tem um salário que... que tá 6.330. Então, o que eu vou fazer? Eu vou colocar de 3.500 a 6.500 e de 6.500 e... de 1.000 a 3.499.
 203. W: E como é que calcula isso na porcentagem do gráfico?
 204. V: Tem que pôr mais ou menos 30.000.
 205. A: Nós tamos ferrados. Tem que pôr porcentagem. Deixa eu pensar. <SUSSURROS>
 206. V: Isso não vai acabar hoje.
 207. W: Isso tem que acabar hoje.
 208. A: Ah, já sei. Isso é mais fácil... <SUSSURROS> Ah, eu sei... 1.200. Quantas pessoas que pegam 1.200? Olha aí para mim.
 209. W: 1.200? Você não marcou.
 210. A: Olha aí, tá marcado! <PAUSA>
 211. W: Eu nunca gostei de fazer trabalho.
 212. A: Nem eu.
 213. W: Você vai fazer que faculdade? <CONVERSA PARALELA>
 214. V: Todos ganham uma.
 215. A: Olha aí, 1.500 ganha uma pessoa também?
 216. V: Uma, duas.
 217. A: Duas? 1.500 recebem duas... <SUSSURROS>
 218. W: Porra, anota os valores aí, depois você vai fazendo as barrinhas. <PAUSA LONGA> Todos é um, menos '1.500'.
 219. A: E '2.750'?
 220. V: Uma.
 221. A: '2.800'?
 222. V: Uma.

223. W: É tudo uma, parceiro. Todo mundo ganha uma. 224. A: Calmaí, deixa eu contar. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34. 225. V: Falta seis. 226. W: Tá certo, seis pessoas não ganham nada. 227. A: Quatro pessoas não ganham nada. 228. W: Seis, seis pessoas não ganham nada. 229. A: Deixa eu ver, tem certeza? 1, 2, 3, 4, 5. 230. V: Olha aqui, 1, 2, 3, 4, 5, 6, seis pessoas não ganham nada. 231. A: Deixa eu ver, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34. Tá certo, mas calma aí, acho que tem coisa errada, eu dividi por 36. Não, tá certo, eu dividi por 34. <RISOS> <CONVERSAS PARALELAS> Nossa, esqueci de colocar o 'Não precisa declarar'. Quantas pessoas colocaram 'Não deseja declarar'? 232. W: Três. 233. A: E 'Não sabe'? 234. V: Três, também. 235. W: Olha, já vai dar o sinal, vamos chamar ele, a gente faz a 'c' amanhã. 236. A: Ele vai vir aqui. 237. V: Eu vou falar pra ele que a gente vai terminar amanhã. 238. A: Fala pra ele que só falta a 'c'. <TERMINA O TEMPO DA AULA>	
SEGUNDA SESSÃO [00:34:06] – 30/11/2018	
Produção oral	Produção escrita
<REINICIA NO DIA SEGUINTE> 239. A: Nós parou em qual? 240. W: Na 'c'. 241. A: Pode crer, na 'c'. <LÊ O ENUNCIADO DO ITEM C> 242. W: Tem que fazer média, mediana e moda de novo? 243. A: Espera aí. <RELÊ O ENUNCIADO DO ITEM C> Se fosse usar gráfico, teria que usar o gráfico de barras. Se for o gráfico de pizza, ficaria ruim de fazer, tanto para dividir quanto para fazer.	

244. W: Vai usar o de barra mesmo.
 245. A: Tem que fazer uma medida-resumo, vamos fazer a média, a mediana e moda.
 246. V: Pega o livro aí.
 247. A: Não, mas no livro não tem como fazer. Aí não tem, porque a medida é qualitativa, não dá pra fazer média nem mediana... Não, mediana tem, só tem como fazer moda e mediana, média não tem, porque é qualitativa. Qual é a moda? Moda entre os pais... Tem que colocar os dois. <PAUSA> <CONVERSA PARALELA> A moda é 'ensino médio completo'. Quantas pessoas? 24 pessoas... 24 pessoas.
 248. W: Por que 24 pessoas? Porque aqui tá 40, 40 pais e 40 mães.
 249. A: 9, 9 + 15?
 250. W: Calma, se tá falando o quê? Ensino médio? <SUSSURROS>
 251. A: 'Ensino médio completo' tá aqui, são 9 pais e 15 mães. Eu coloquei 24, sendo 9 pais e 15 mães... Agora é mediana... Como é que se calcula mediana aqui? Não tem como calcular a mediana.
 252. W: Pergunta pro Cassio.
 253. A: Calma aí, não precisa, tem no caderno a coisa qualitativa. <PAUSA> Pai e mãe. Tem que calcular os dois, porque a mediana tá no meio dos dois. <PAUSA> Tem '0', se bem que 0 não entra... A mediana seria o mais... tem que estar no meio dos dois... seria o '4', a mediana é 4. <SUSSURROS> Agora, é o gráfico, tem que fazer o gráfico.
 254. W: Gráfico de barras, qual a altura dele?
 255. A: Tem que fazer dois gráficos... Deixa que eu faço, pode deixar... Eu faço do pai e você faz da mãe. Pode ser?
 256. W: Pode ser. Qual a altura da mãe?
 257. A: É 15.
 258. W: Não é melhor fazer 40 mãe? 40 de pai e 40 de mãe?
 259. A: Não, é 15 por 10 que você tem que fazer.
 260. W: 15 por 10?
 261. 261. A: É... <PAUSA LONGA> O meu vai ter só nove: 1, 2, 3, 4, 5, 6, 7, 8, 9...
 262. W: Nossa!
 263. A: Aqui você vai colocar aqui, 'Não alfabetizado'... 'Fundamental incompleto'... <PAUSA LONGA> Eu tô fazendo qual? Eu tô fazendo do pai, viu?

C) mediana: 24 pessoas sendo 9 pais e 15 mães
 mediana: 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40
 a mediana entre os pais é 4 e entre os pais é 4
 a mediana é 4



264. V: Nós também.
 265. A: Cara, é coisa... <PAUSA MUITO LONGA> Tá difícil aí.
 266. V: Não.
 267. W: Não.
 268. A: Cara, o '4' é aqui! Não devia ter feito isso... Aí depois vem o '6'?
 Aí depois vem o '8'?
 269. V: O '8' não, vem o '6'.
 270. A: Aí o '3' e o '2'?
 271. V: É.
 272. A: <LÊ A TABELA> <PAUSA LONGA> Tá difícil aí?
 273. W: Não.
 274. A: Cê tá confundindo tudo aí. Cê é burro! Cê tem que colocar aqui
 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, cê bota até 6, né?
 275. V: Não sei por que você colocou o 0..
 276. W: Porque tem 0! Vai colocar aonde?
 277. A: <RISOS> <PAUSA LONGA> <CONVERSAS PARALELAS>
 278. W: Você tem que explicar o que tá fazendo.
 279. A: Como assim?
 280. W: Você tem que explicar tudo que a gente fez. <CONVERSAS
 PARALELAS>
 281. A: Agora explica.
 282. W: A gente tirou a média, que o resultado deu 15,0,25, foi uma
 média de 15 anos pela sala de aula. Tiramos a moda, que deu 16
 anos <RISOS DOS OUTROS DOIS>, tiramos a mediana que foi 16
 anos, o quartil a gente não conseguiu fazer, o desvio padrão deu 3, a
 amplitude deu 10 e o mapa, a gente fez o mapa de barras
 simbolizando a idade e o número de pessoas que tem cada idade.
 Agora eu vou passar para ele que vai fazer a 'b'.

ANEXO 1 – Termo de Consentimento Livre e Esclarecido



PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO (PUC-SP)

Pesquisa: Processos de Ensino e Aprendizagem de Matemática em Ambientes Tecnológicos

Termo de Consentimento Livre e Esclarecido

Eu, _____, com ____ anos de idade, portador do RG _____, residente na _____, nº _____, bairro _____, cidade _____, estado _____, número de telefone _____, e-mail _____ abaixo assinado, dou meu consentimento livre e esclarecido para que meu filho(a) _____, participe como voluntário da pesquisa supra citada, sob responsabilidade do pesquisador, sob a responsabilidade do Prof. Dr. Saddo Ag Almouloud, vinculada ao Programa de Estudos Pós-graduados em Educação Matemática da PUC-SP.

Assinando este Termo de Consentimento, estou ciente de que:

- 1) O objetivo do questionário, parte da pesquisa desenvolvida Cassio Cristiano Giordano, é o de realizar um estudo a respeito das concepções mobilizadas por estudantes do Ensino Médio sobre Estatística Descritiva.
- 2) Assim que for terminada a pesquisa, o relatório com seus resultados será divulgado para a comunidade acadêmica.
- 3) Estou livre para interromper, a qualquer momento, minha participação nesta pesquisa.
- 4) Poderei entrar em contato com o pesquisador responsável, Saddo Ag Almouloud pelo e-mail saddoag@pucsp.br ou do pesquisador Cassio Cristiano Giordano pelo e-mail ccgiordano@gmail.com e telefone (11) 4425-3396.
- 5) Obtive todas as informações necessárias para poder decidir conscientemente sobre a minha participação na referida pesquisa;

São Paulo, _____ de _____ de 2018.

Assinatura do participante

Assinatura do pesquisador

ANEXO 2 – Folhas de respostas dos quatro grupos de alunos

Grupo I

Atividade de Estatística

Em uma escola pública de Santo André, um grupo de alunos do terceiro ano do Ensino Médio, realizando uma pesquisa escolar, resolveu investigar a opinião dos colegas sobre a política de cotas para o ingresso em universidades federais. Essa escola possuía 938 alunos, sendo 417 estudantes do Ensino Fundamental, estudando no período vespertino (6º ao 9º ano) e 521 no Ensino Médio, estudando no período matutino (204 no 1º ano, 179 no 2º ano e 138 no 3º ano). Desses 938 alunos, 513 são mulheres. Esse grupo decidiu entrevistar 40 alunos, sendo 20 (dez homens e dez mulheres) do período matutino e 20 (dez homens e dez mulheres) do período vespertino. Elaborou um questionário com doze perguntas, cujas respostas podem ser observadas no banco de dados que segue. Analise os resultados utilizando as representações que julgar adequadas e as medidas que julgar adequadas. Lembre de sempre explicar ao máximo seu pensamento.

a) Com relação à variável “idade”, os alunos forneceram as seguintes respostas:

Idade (anos)	11	12	13	14	15	16	17	18	19	20	21	Total
Freq. simples	3	4	4	6	5	8	4	3	2	0	1	40

Como você explica a variação das idades utilizando as ferramentas estatísticas que aprendeu: construa uma ou mais representações gráficas, explicando-as. Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de idades do grupo de 40 alunos pesquisados.

b) Em relação à renda familiar, os alunos obtiveram as seguintes respostas

Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)
1	1108,38	11	4800	21	3000	31	Não sabe
2	1500	12	5100	22	1108,38	32	3328,14
3	2521	13	1108,38	23	Prefere não declarar	33	4623
4	Prefere não declarar	14	1108,38	24	3820	34	4100
5	2800	15	1500	25	2216,76	35	5000
6	1108,38	16	Não sabe	26	2600	36	2400
7	3500	17	2750	27	1200	37	1108,38
8	4714	18	3400	28	1108,38	38	1700
9	1108,38	19	Prefere não declarar	29	30000	39	6330
10	2200	20	1108,38	30	1300	40	Não sabe

Como você explica a variação das rendas declaradas utilizando as ferramentas estatísticas que aprendeu: construa uma ou mais representações gráficas, explicando-as. Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de dados do grupo de 40 alunos pesquisados.

L> rendas
familiares

- c) Em relação à escolaridade dos pais, os alunos elaboraram a tabela a seguir. A tabela construída simplifica o banco de dados. Como você evidenciaria essa comparação se quisesse utilizar gráficos? Explique seu pensamento. É possível determinar uma medida-resumo que contribua com a análise desses dados? Qual? Como determinar? Explique seu pensamento.

Escolaridade	do pai	da mãe
não alfabetizado	2	1
ens. fundamental incompleto	2	1
ens. fundamental completo	4	6
ens. médio incompleto	4	0
ens. médio completo	9	15
ens. superior incompleto	6	1
ens. superior completo	8	12
pós-graduação incompleta	3	0
pós-graduação completa	2	4
Total	40	40



1- Meda = 16 anos

Media = 15 e 16 anos

Mediana = 14 e 15 anos

Desvio =

11	12	13	14	15	16	17	18	19	20	21	total
3	4	4	6	5	8	4	3	2	0	1	40

$$11 = -4^2 = 16$$

$$21 = 6^2 = 36$$

$$12 = -3^2 = 9$$

$$20 = 5^2 = 25$$

$$13 = -2^2 = 4$$

$$19 = 4^2 = 16$$

$$14 = -1^2 = 1$$

$$18 = 3^2 = 9$$

$$17 = 2^2 = 4$$

$$16 = 1^2 = 1$$

$$15 = 0^2 = 0$$

Variação =

$$V = \frac{3 \cdot (11)^2 + 4 \cdot (12)^2 + 4 \cdot (13)^2 + 6 \cdot (14)^2 + 5 \cdot (15)^2 + 8 \cdot (16)^2 + 4 \cdot (17)^2 + 3 \cdot (18)^2 + 2 \cdot (19)^2 + 0 \cdot (20)^2 + 1 \cdot (21)^2}{40}$$

$$V = \frac{9255}{40} \approx 231,3$$

$$\text{Desvio padrão} = \sqrt{231,3} \approx 15,2 \quad 15 \text{ e } 16 \text{ anos}$$

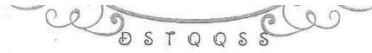
2- Meda = 1108,38

Media = 2909,45

Mediana = 2501 a 3000

Desvio =



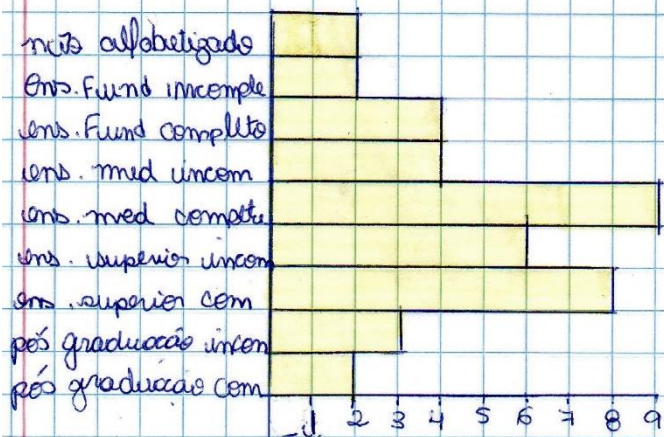
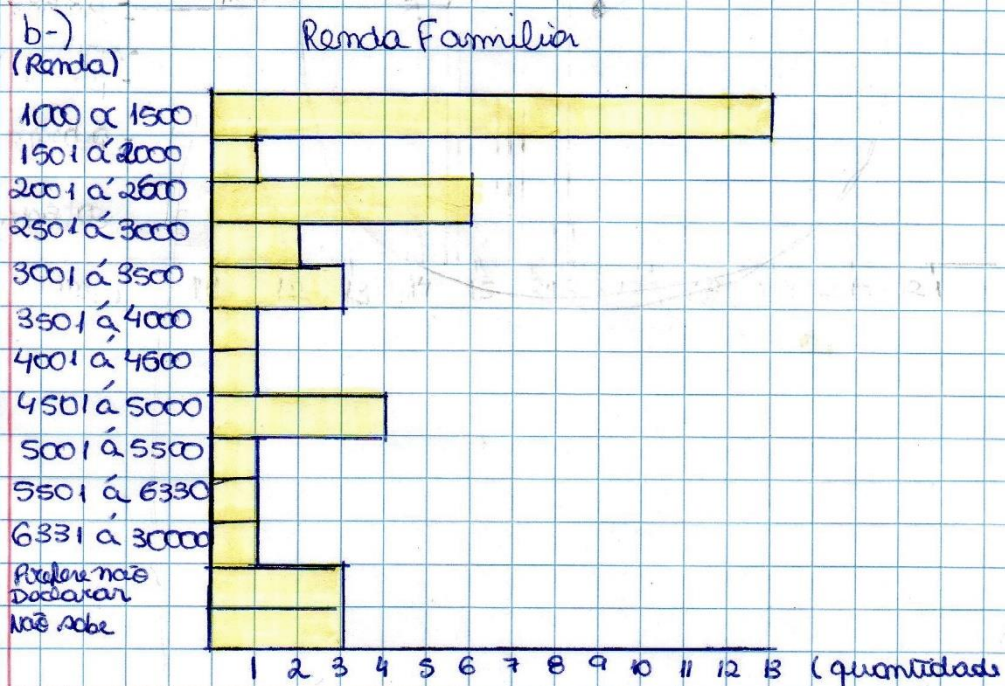
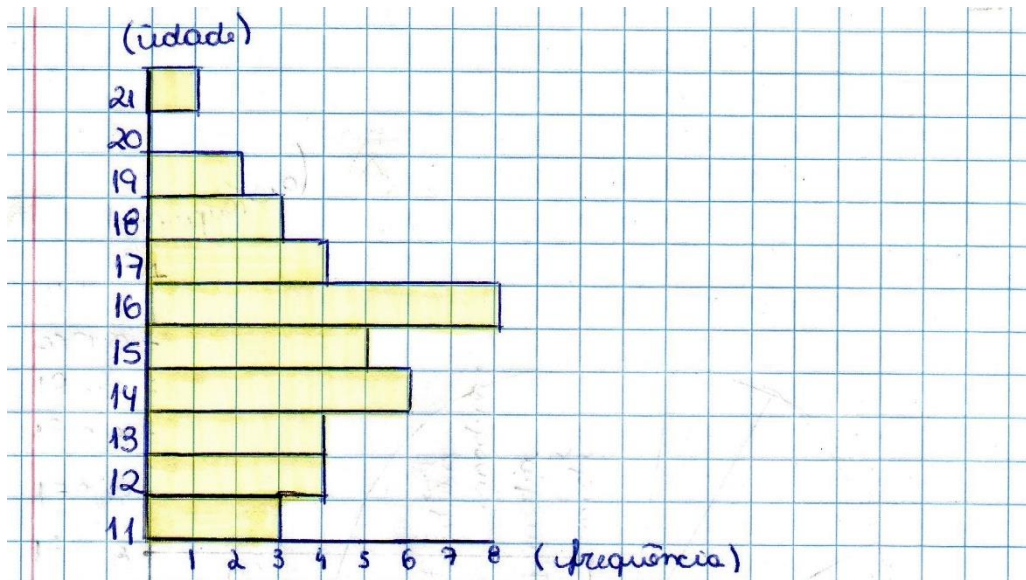


$$2001 \times 2500 = -$$

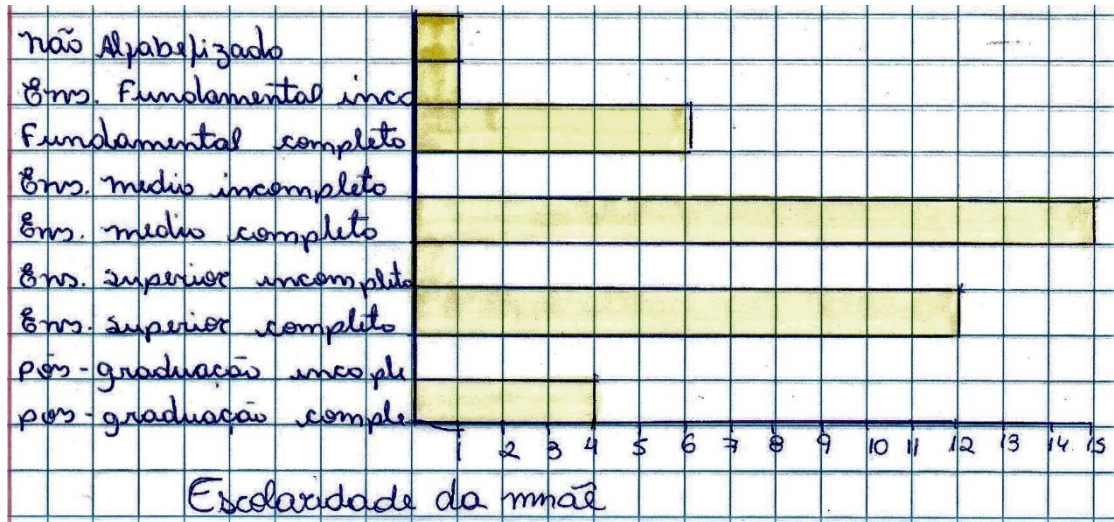
$$2001 \times 2500 = -$$

3) moda = Ensino médio completo
 mediana = Esta entre o Ensino médio completo e o Ensino Superior incompleto.





Educação do Pai



1108.38
3000

~~4,327,347.38~~
~~341,972.92~~
~~957,149.54~~
~~246870~~
162,674.6

4,022,910.54
40

254,918.08

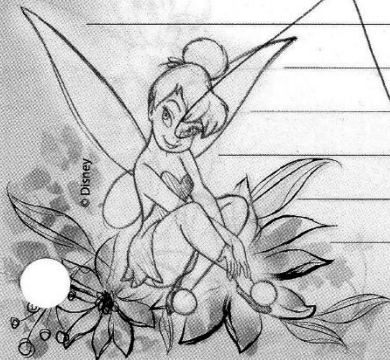
100,572.7635

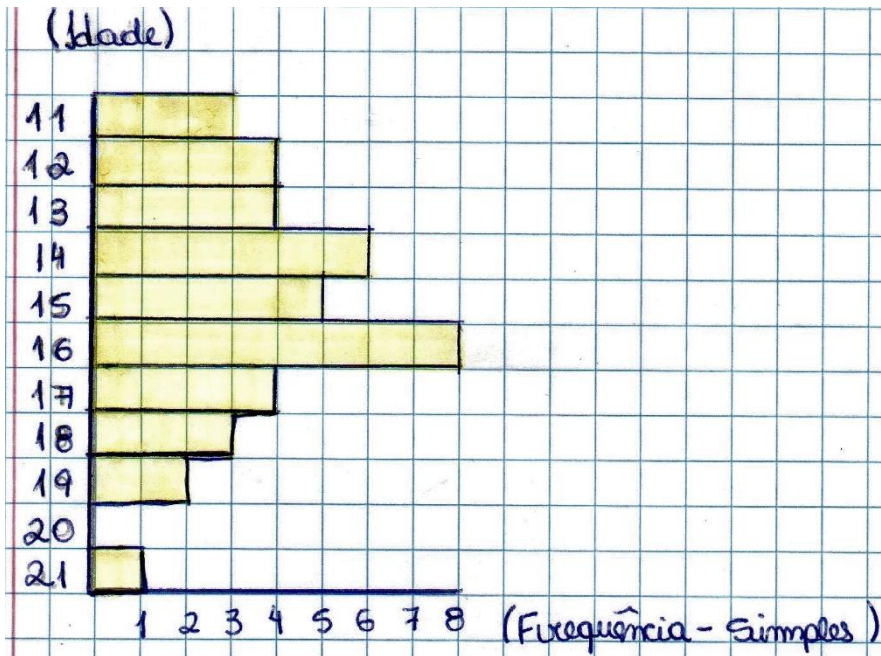
2,909.458

9,975.42
3,000
2,521
0
2,800
3,500
4,114
2,200
4,800
5,100

2,750
3

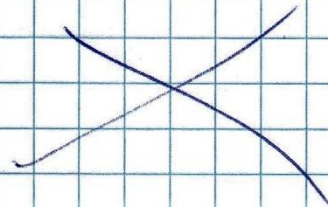
88,897.18





~~1000 a 1500~~
~~1501 a 2000~~
~~2001 a 2500~~

1000 a 1500
 1501 a 2000
 2001 a 2500
 2501 a 3000
 3001 a 3500
 3501 a 4000
 4001 a



Grupo II

Atividade de Estatística

Em uma escola pública de Santo André, um grupo de alunos do terceiro ano do Ensino Médio, realizando uma pesquisa escolar, resolveu investigar a opinião dos colegas sobre a política de cotas para o ingresso em universidades federais. Essa escola possuía 938 alunos, sendo 417 estudantes do Ensino Fundamental, estudando no período vespertino (6º ao 9º ano) e 521 no Ensino Médio, estudando no período matutino (204 no 1º ano, 179 no 2º ano e 138 no 3º ano). Desses 938 alunos, 513 são mulheres. Esse grupo decidiu entrevistar 40 alunos, sendo 20 (dez homens e dez mulheres) do período matutino e 20 (dez homens e dez mulheres) do período vespertino. Elaborou um questionário com doze perguntas, cujas respostas podem ser observadas no banco de dados que segue. Analise os resultados utilizando as representações que julgar adequadas e as medidas que julgar adequadas. Lembre de sempre explicar ao máximo seu pensamento.

a) Com relação à variável "idade", os alunos forneceram as seguintes respostas:

Idade (anos)	11	12	13	14	15	16	17	18	19	20	21	Total
Freq. simples	3	4	4	6	5	8	4	3	2	0	1	40

Como você explica a variação das idades utilizando as ferramentas estatísticas que aprendeu: construa uma ou mais representações gráficas, explicando-as. Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de idades do grupo de 40 alunos pesquisados.

b) Em relação à renda familiar, os alunos obtiveram as seguintes respostas

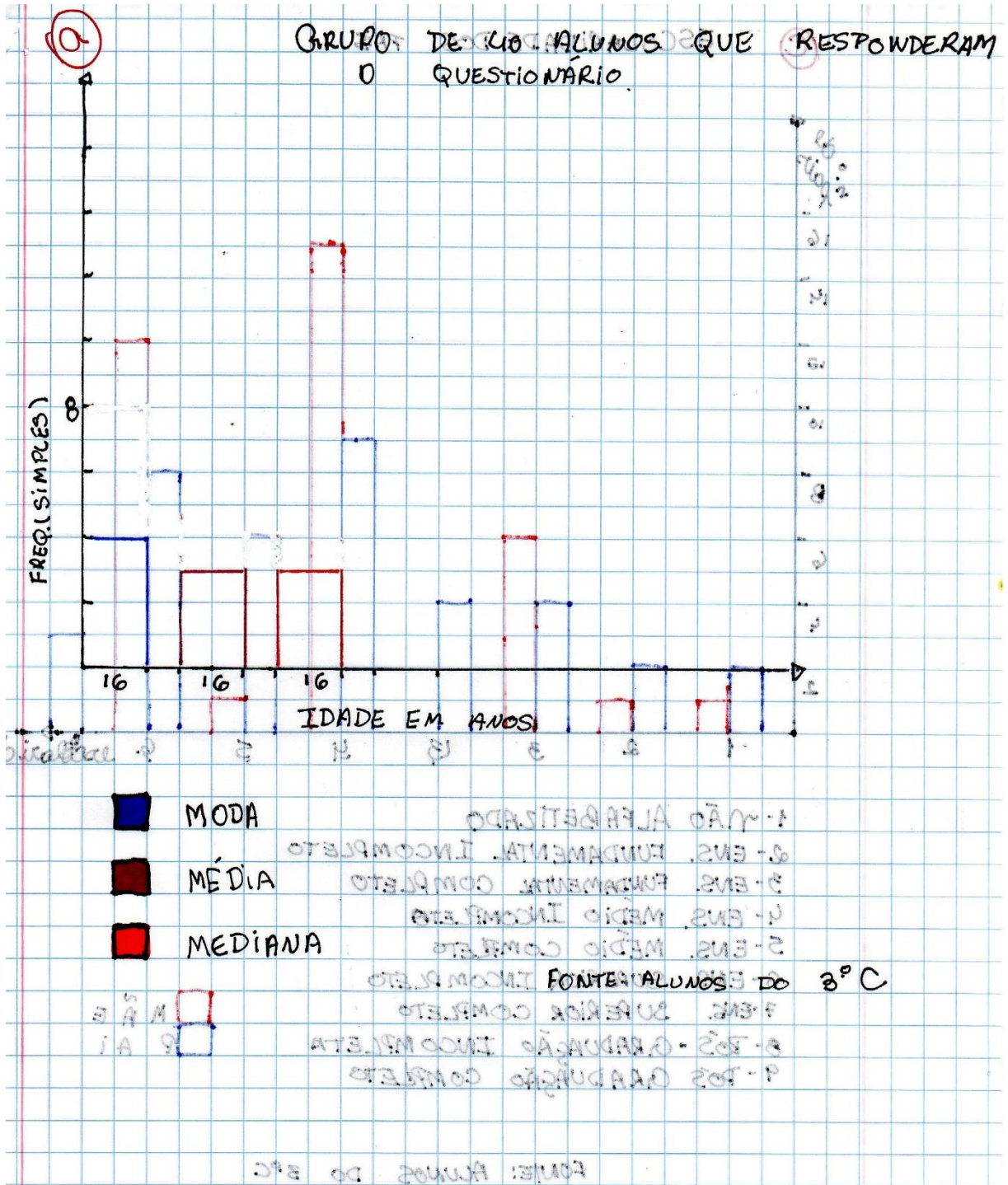
Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)
1	1108,38 /	11	4800	21	3000	31	Não sabe
2	1500 /	12	5100	22	1108,38 /	32	3328,14
3	2521	13	1108,38 /	23	Prefere não declarar	33	4623
4	Prefere não declarar	14	1108,38 /	24	3820	34	4100
5	2800	15	1500 /	25	2216,76	35	5000
6	1108,38 /	16	Não sabe	26	2600	36	2400
7	3500	17	2750	27	1200	37	1108,38 /
8	4714	18	3400	28	1108,38 /	38	1700
9	1108,38 /	19	Prefere não declarar	29	30000	39	6330
10	2200	20	1108,38 /	30	1300	40	Não sabe

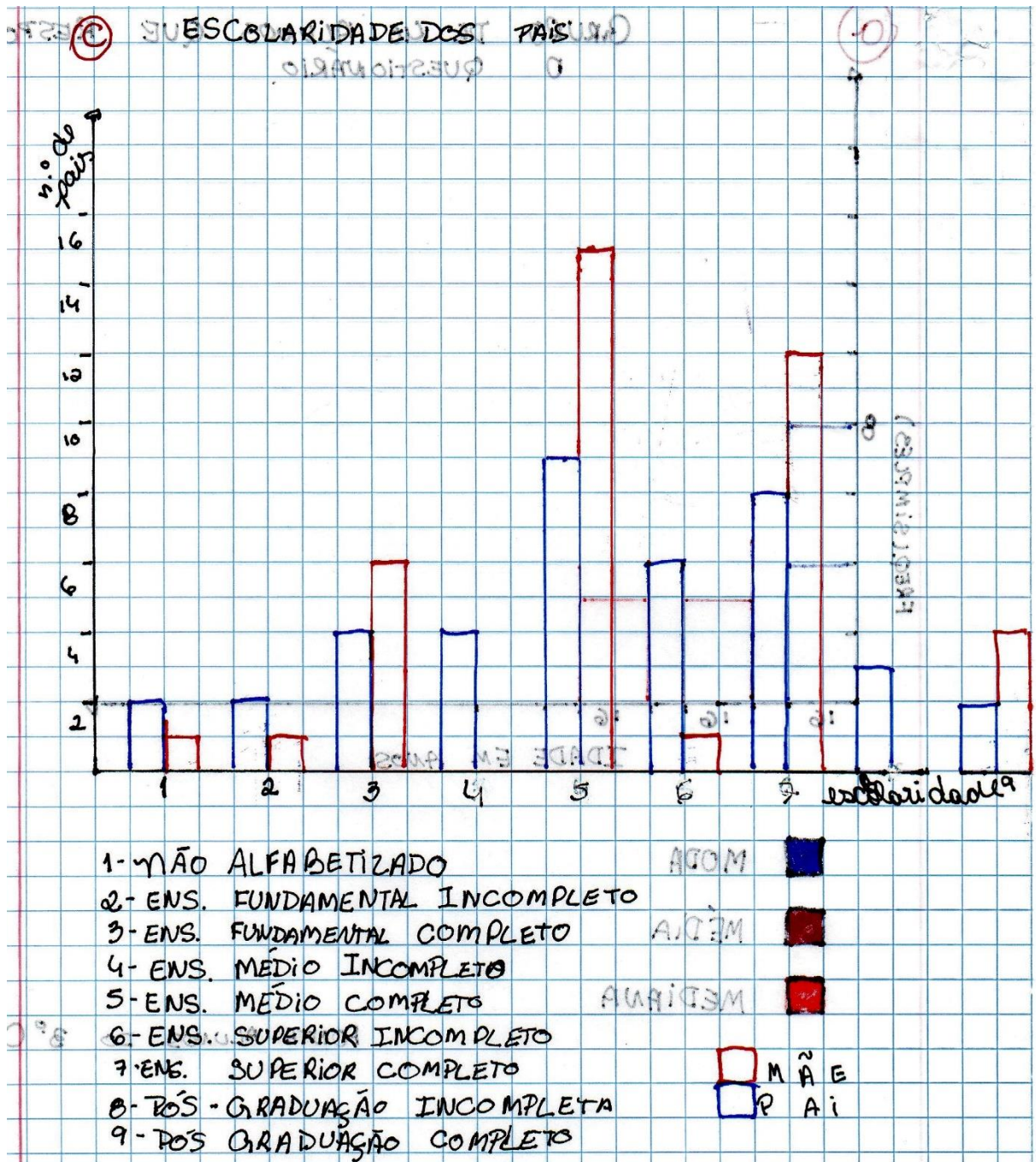
a.
b

Como você explica a variação das rendas declaradas utilizando as ferramentas estatísticas que aprendeu: construa uma ou mais representações gráficas, explicando-as. Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de ~~idades~~ ^{RENDAS FAMILIARES} do grupo de 40 alunos pesquisados.

c) Em relação à ~~escolaridade~~ ^{escolaridade} dos pais, os alunos elaboraram a tabela a seguir. A tabela construída simplifica o banco de dados. Como você evidenciaria essa comparação se quisesse utilizar gráficos? Explique seu pensamento. É possível determinar uma medida-resumo que contribua com a análise desses dados? Qual? Como determinar? Explique seu pensamento.

	Escolaridade	do pai	da mãe
1	não alfabetizado	2	1
2	ens. fundamental incompleto	2	1
3	ens. fundamental completo	4	6
4	ens. médio incompleto	4	0
5	ens. médio completo	9	15
6	ens. superior incompleto	6	1
7	ens. superior completo	8	12
8	pós-graduação incompleta	3	0
9	pós-graduação completa	2	4
	Total	40	40

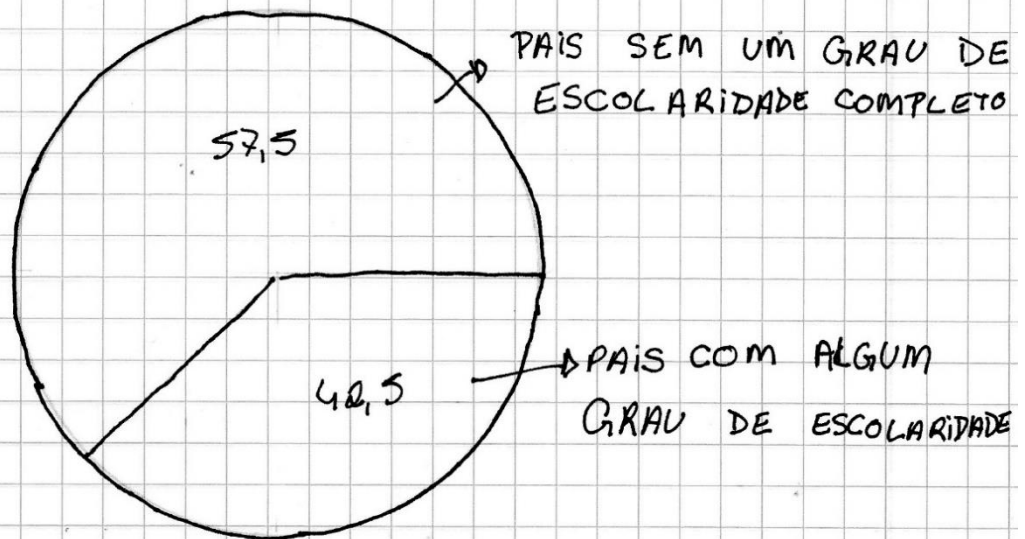




FONTE: ALUNOS DO 3º C.

(C)

ESCOLARIDADE DOS PAIS



FONTE: ALUNOS DO 3º C

(b) Desvio padrão e variância da
pe pela relação do médio $\bar{x}$

Grupo III

Atividade de Estatística

Em uma escola pública de Santo André, um grupo de alunos do terceiro ano do Ensino Médio, realizando uma pesquisa escolar, resolveu investigar a opinião dos colegas sobre a política de cotas para o ingresso em universidades federais. Essa escola possuía 938 alunos, sendo 417 estudantes do Ensino Fundamental, estudando no período vespertino (6º ao 9º ano) e 521 no Ensino Médio, estudando no período matutino (204 no 1º ano, 179 no 2º ano e 138 no 3º ano). Desses 938 alunos, 513 são mulheres. Esse grupo decidiu entrevistar 40 alunos, sendo 20 (dez homens e dez mulheres) do período matutino e 20 (dez homens e dez mulheres) do período vespertino. Elaborou um questionário com doze perguntas, cujas respostas podem ser observadas no banco de dados que segue. Analise os resultados utilizando as representações que julgar adequadas e as medidas que julgar adequadas. Lembre de sempre explicar ao máximo seu pensamento.

a) Com relação à variável “idade”, os alunos forneceram as seguintes respostas:

Idade (anos)	11	12	13	14	15	16	17	18	19	20	21	Total
Freq. simples	3	4	4	6	5	8	4	3	2	0	1	40

Como você explica a variação das idades utilizando as ferramentas estatísticas que aprendeu: construa uma ou mais representações gráficas, explicando-as. Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de idades do grupo de 40 alunos pesquisados.

b) Em relação à renda familiar, os alunos obtiveram as seguintes respostas

Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)
1	1108,38 X	11	4800 X	21	3000 X	31	Não sabe X
2	1500 X	12	5100 X	22	1108,38 X	32	3328,14 X
3	2521 X	13	1108,38 X	23	Prefere não declarar X	33	4623 X
4	Prefere não declarar X	14	1108,38 X	24	3820 X	34	4100 X
5	2800 X	15	1500 X	25	2216,76 X	35	5000 X
6	1108,38 X	16	Não sabe X	26	2600 X	36	2400 X
7	3500 X	17	2750 X	27	1200 X	37	1108,38 X
8	4714 X	18	3400 X	28	1108,38 X	38	1700 X
9	1108,38 X	19	Prefere não declarar X	29	30000	39	6330 X
10	2200 X	20	1108,38 X	30	1300 X	40	Não sabe X

Como você explica a variação das rendas declaradas utilizando as ferramentas estatísticas que aprendeu: construa uma ou mais representações gráficas, explicando-as. Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de ~~dados~~ *rendas familiares* do grupo de 40 alunos pesquisados.

- c) Em relação à ~~questão 2.93~~ escolaridade dos pais, os alunos elaboraram a tabela a seguir. A tabela construída simplifica o banco de dados. Como você evidenciaria essa comparação se quisesse utilizar gráficos? Explique seu pensamento. É possível determinar uma medida-resumo que contribua com a análise desses dados? Qual? Como determinar? Explique seu pensamento.

Escolaridade	do pai	da mãe
não alfabetizado	2	1
ens. fundamental incompleto	2	1
ens. fundamental completo	4	6
ens. médio incompleto	4	0
ens. médio completo	9	15
ens. superior incompleto	6	1
ens. superior completo	8	12
pós-graduação incompleta	3	0
pós-graduação completa	2	4
Total	40	40

• Atividade de Estatística

Questão 1 $Q1 = 26 + 31 + P + 1 + 1 + 0 + 1 + P + P + 31 + 26 = 116$

938 alunos

417 alunos (estudantes do Ensino Fundamental)

521 alunos (Ensino médio)

204 alunos 1º ano

179 " 2º ano

138 " 3º ano

513 mulheres

Investiga 40 alunos

20 do período matutino

20 do período vespertino

a) "Idade" variável

Média = $\frac{11 + 12 + 13 + 14 + 15 + 16 + 17 + 18 + 19 + 20 + 21}{11}$

Média = $\frac{176}{11}$	Idade	Valor - Média	Quadrado
	11	-5	25
Média = 16	12	-4	16
Moda = 16	13	-3	9
Mediana = 16	14	-2	4
Variância = 10	15	-1	1
Desvio Padrão = 3,16	16	0	0
	17	1	1
	18	2	4
	19	3	9
	20	4	16
	21	5	25
	Média = 16		

$$\rightarrow \text{Média} = \frac{25 + 16 + 9 + 4 + 1 + 0 + 1 + 4 + 9 + 16 + 25}{11} = 10$$

$$\sqrt{10} = 3,16$$

• Explicação

É necessário entender que uma variação de dados pode ser explicada através do enunciado. Diante disso, resultados como a meda são nitidamente explicados com a análise dos dados transmitidos pelo texto.

Questão 2

Renda Familiar

$$\begin{aligned} \text{Média aritmética} = & 1.108,38 + 1500 + 2521 + 0 + 2800 + 1.108,38 + 3500 \\ & + 4714 + 1.108,38 + 2200 + 4800 + 5100 + 1.108,38 + 1.108,38 + 1500 + 0 + 2750 \\ & + 3400 + 0 + 1.108,38 + 3000 + 1.108,38 + 0 + 3820 + 2.216,76 + 2600 \\ & + 1200 + 1.108,38 + 30000 + 1300 + 0 + 3.328,14 + 4623 + 4100 + 5000 + 2400 \\ & + 1.108,38 + 1700 + 6330 + 0 \end{aligned}$$

$$M = \frac{116.378,32}{40}$$

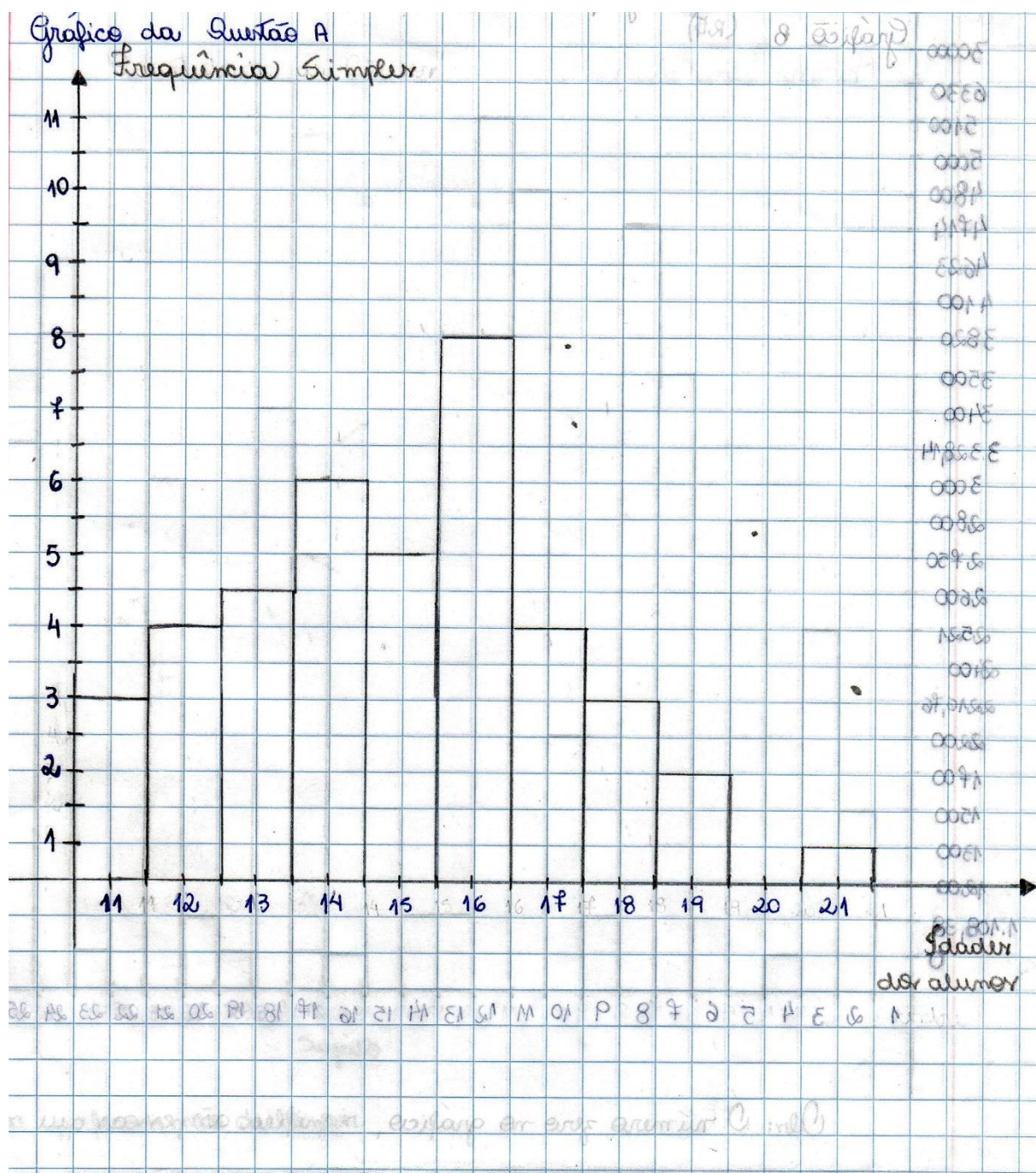
$$M = 2.909,458$$

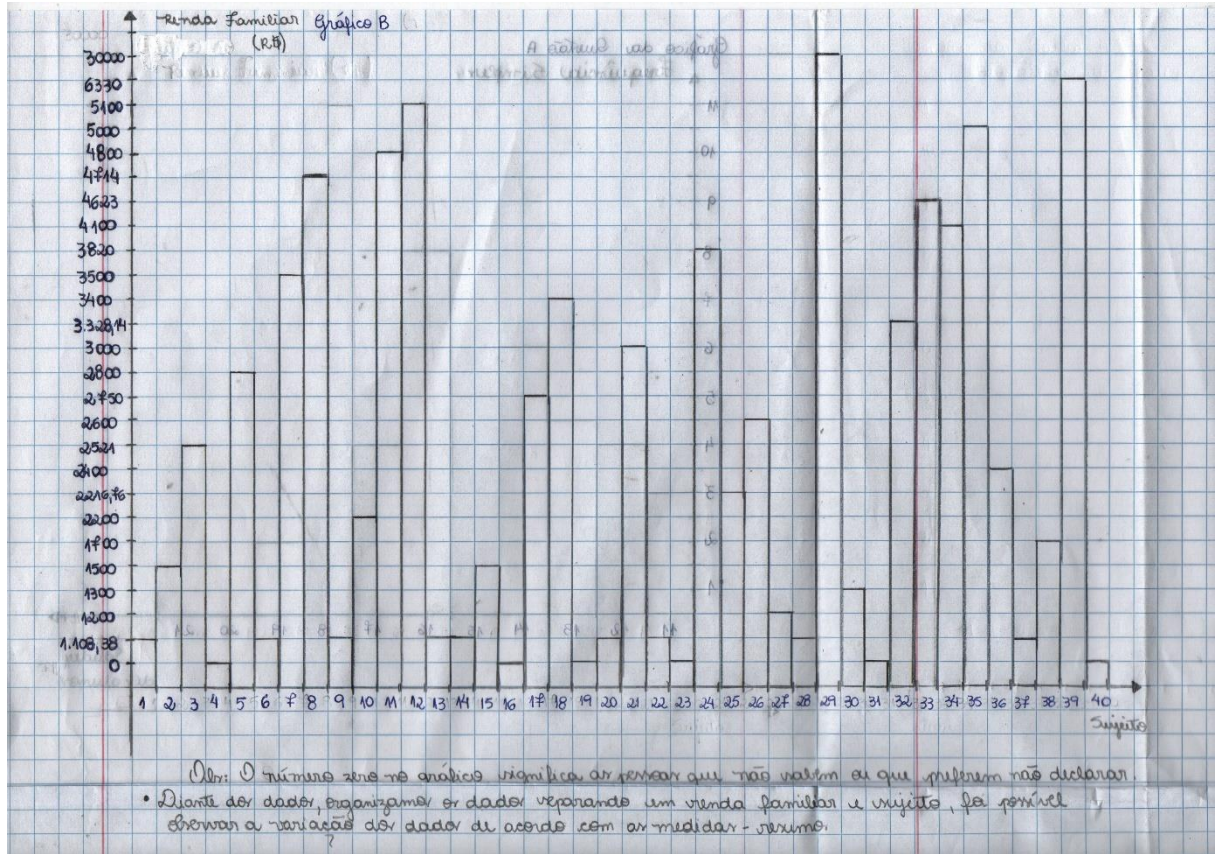
$$\text{Média} = 2.909,458$$

$$\text{Meda} = 1.108,38$$

$$\begin{aligned} \text{Mediana} = & 0 - 0 - 0 - 0 - 0 - 0 - 1.108,38 - 1.108,38 - 1.108,38 \\ & - 1.108,38 - 1.108,38 - 1.108,38 - 1.108,38 - 1.108,38 - 1.108,38 - 1.200 - 1300 \\ & - 1500 - 1500 - 1700 - 2200 - 2216,76 - 2400 - 2521 - 2600 - 2750 \\ & - 2800 - 3000 - 3328,14 - 3400 - 3500 - 3820 - 4100 - 4623 - 4714 \\ & - 4800 - 5000 - 5100 - 6330 - 30000 \end{aligned}$$

$$\text{Mediana} = \frac{1700 + 2200}{2} = \frac{3900}{2} = 1950$$





Questão 3

PAIS

MODA $\rightarrow$ Enxmo médio completo com 9 alunos

MEDIANA $\rightarrow 2 \ 2 \ 2 \ 3 \ (4) \ 4 \ 6 \ 8 \ 9 = 4$

MAES

MODA $\rightarrow$ Enxmo médio completo com 15 alunos

MEDIANA $\rightarrow 0 \ 0 \ 1 \ 1 \ (1) \ 4 \ 6 \ 12 \ 15 = 1$

Diante das informações obtidas no enunciado foi possível determinar a moda e a mediana sobre a escolaridade dos pais e dos mães. Não foi possível calcular a média pois não foi dada nenhuma informação quantitativa a respeito do ensino médio também não foram calculadas

• Medidas de Dispersão (Variabilidade)

Variância

→ Sejam $x_1, x_2, \dots, x_N$ a coleção de valores assumidos por uma variável quantitativa de X e $\bar{x}$ a média aritmética desses valores.

→ N é a quantidade de dados

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{N}$$

$$\sigma^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_N - \bar{x})^2}{N}$$

Desvio Padrão

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

$\sum$ = soma de n termos

x_i = valor individual da amostra

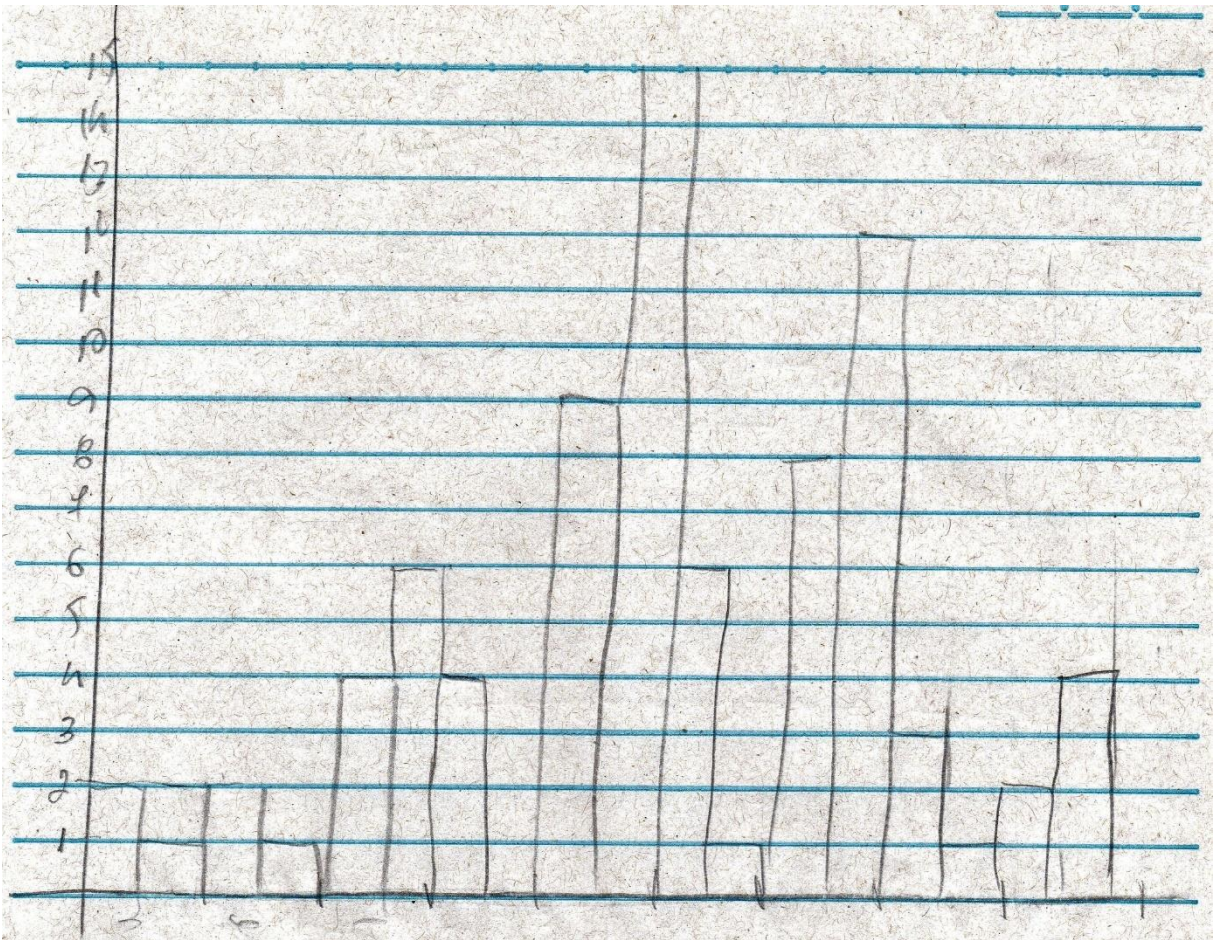
$\bar{x}$ = valor médio aritmético

n = número de vezes da repetição

$$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_N - \bar{x})^2}{N}}$$

• Quanto maior o desvio padrão, maior a variabilidade dos dados, ou seja, MENOR é a regularidade dos meus dados.

• Quanto menor o desvio padrão, menor a variabilidade dos dados, ou seja, MAIOR é a regularidade dos meus dados.



1º PAIS

MEDIA $\rightarrow 2+2+$

Grupo IV

Atividade de Estatística

Em uma escola pública de Santo André, um grupo de alunos do terceiro ano do Ensino Médio, realizando uma pesquisa escolar, resolveu investigar a opinião dos colegas sobre a política de cotas para o ingresso em universidades federais. Essa escola possuía 938 alunos, sendo 417 estudantes do Ensino Fundamental, estudando no período vespertino (6º ao 9º ano) e 521 no Ensino Médio, estudando no período matutino (204 no 1º ano, 179 no 2º ano e 138 no 3º ano). Desses 938 alunos, 513 são mulheres. Esse grupo decidiu entrevistar 40 alunos, sendo 20 (dez homens e dez mulheres) do período matutino e 20 (dez homens e dez mulheres) do período vespertino. Elaborou um questionário com doze perguntas, cujas respostas podem ser observadas no banco de dados que segue. Analise os resultados utilizando as representações que julgar adequadas e as medidas que julgar adequadas. Lembre de sempre explicar ao máximo seu pensamento.

a) Com relação à variável “idade”, os alunos forneceram as seguintes respostas:

Idade (anos)	11	12	13	14	15	16	17	18	19	20	21	Total
Freq. simples	3	4	4	6	5	8	4	3	2	0	1	40

Como você explica a variação das idades utilizando as ferramentas estatísticas que aprendeu: construa uma ou mais representações gráficas, explicando-as. Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de idades do grupo de 40 alunos pesquisados.

b) Em relação à renda familiar, os alunos obtiveram as seguintes respostas

Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)	Sujeito	Renda Familiar (R\$)
1	1108,38	11	4800	21	3000	31	Não sabe
2	1500	12	5100	22	1108,38	32	3328,14
3	2521	13	1108,38	23	Prefere não declarar	33	4623
4	Prefere não declarar	14	1108,38	24	3820	34	4100
5	2800	15	1500	25	2216,76	35	5000
6	1108,38	16	Não sabe	26	2600	36	2400
7	3500	17	2750	27	1200	37	1108,38
8	4714	18	3400	28	1108,38	38	1700
9	1108,38	19	Prefere não declarar	29	30000	39	6330
10	2200	20	1108,38	30	1300	40	Não sabe

Como você explica a variação das rendas declaradas utilizando as ferramentas estatísticas que aprendeu: construa uma ou mais representações gráficas, explicando-as. Determine uma ou mais medidas-resumo (média, moda, mediana, quartis, desvio-padrão, amplitude) e explique seu significado quando queremos analisar esse conjunto de ~~dados~~ *rendas familiares* do grupo de 40 alunos pesquisados.

- c) Em relação à ~~escolaridade~~ *escolaridade* dos pais, os alunos elaboraram a tabela a seguir. A tabela construída simplifica o banco de dados. Como você evidenciaria essa comparação se quisesse utilizar gráficos? Explique seu pensamento. É possível determinar uma medida-resumo que contribua com a análise desses dados? Qual? Como determinar? Explique seu pensamento.

Escolaridade	do pai	da mãe
não alfabetizado	2	1
ens. fundamental incompleto	2	1
ens. fundamental completo	4	6
ens. médio incompleto	4	0
ens. médio completo	9	15
ens. superior incompleto	6	1
ens. superior completo	8	12
pós-graduação incompleta	3	0
pós-graduação completa	2	4
Total	40	40

$$a) \text{ média: } 11 \cdot 3 + 12 \cdot 4 + 13 \cdot 4 + 14 \cdot 6 + 15 \cdot 5 + 16 \cdot 8 + 17 \cdot 4 + 18 \cdot 3 + 19 \cdot 2 + 20 \cdot 0 + 21 \cdot 1$$

$$40$$

$$= 33 + 48 + 52 + 84 + 75 + 128 + 68 + 54 + 38 + 0 + 21$$

$$40$$

média = 15,025, a média da idade é 15 anos
moda = 16 anos

mediana = 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21

a mediana é 16 anos

Quartis $Q_1 = 11, Q_2 = 16, Q_3 = 19$

35,3011 RA = 0,0001

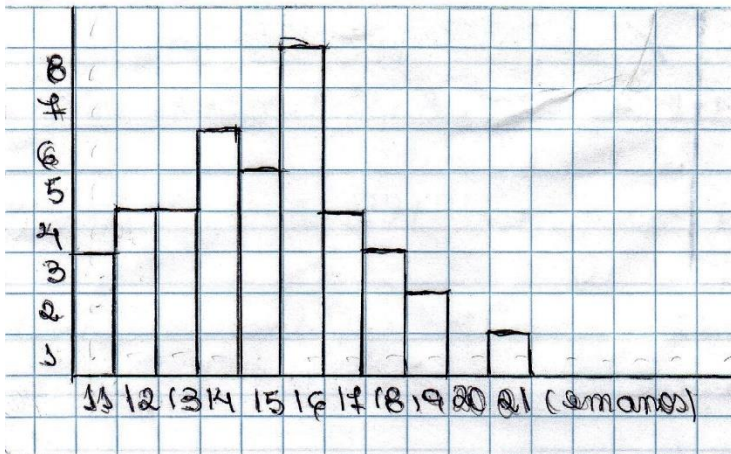
Desvio padrão $\approx 3,3$

$$\begin{aligned} & (11 - 15,025)^2 + (12 - 15,025)^2 + (13 - 15,025)^2 + (14 - 15,025)^2 + (15 - 15,025)^2 \\ & + (16 - 15,025)^2 + (17 - 15,025)^2 + (18 - 15,025)^2 + (19 - 15,025)^2 \\ & + (20 - 15,025)^2 + (21 - 15,025)^2 \div 11 = (-4,025)^2 + (-3,025)^2 + (-2,025)^2 + (-1,025)^2 + (-0,025)^2 + (0,975)^2 + (1,975)^2 \\ & + (2,975)^2 + (3,975)^2 + (4,975)^2 + (5,975)^2 \div 11 \approx 16,2 + 9,13 \\ & 4,1 + 1,05 + 0,000625 + 0,95 + 3,9 + 8,85 + 15,8 + 24,75 + 35,7 \div 11 \approx 10,95 = 3,3 \end{aligned}$$

Amplitude = 10

11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21

$$11 - 21 = 10$$



minim (o)

$60 + 8H + 50 = 190$

$60 + 8H + 50 =$

$21 = \text{minim}$

$21 = \text{maxim}$

média: ~~2500~~ R\$ 3,422,88

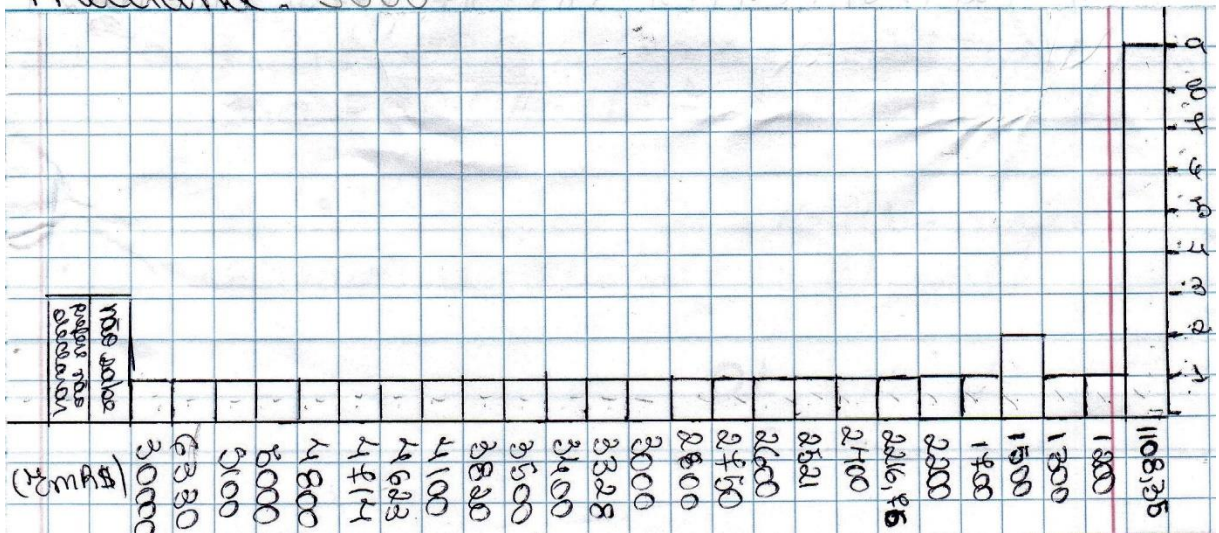
$$b) 1108,38 \cdot 9 + 1500 \cdot 2 + 2521 \cdot 1 + 0 \cdot 5 + 2800 \cdot 1 + 3500 \cdot 1 + 4714 \cdot 1 + 2200 \cdot 1 + 4800 \cdot 1 + 5100 \cdot 1 + 2750 \cdot 1 + 3400 \cdot 1 + 3000 \cdot 1 + 3820 \cdot 1 + 2216,76 \cdot 1 + 2600 \cdot 1 + 1200 \cdot 1 + 30000 \cdot 1 + 1300 \cdot 1 + 3328 \cdot 1 + 4623 \cdot 1 + 4100 \cdot 1 + 5000 \cdot 1 + 2400 \cdot 1 + 1400 \cdot 1 + 6330 \cdot 1 \div 34 = 2,909,45$$

meda = R\$ 1108,38

~~5,8~~ ~~27,01~~ = ~~20,19~~ ~~20,19~~

mediana: 1108,38; 1200; 1300; 1500; 1700; 2200; 2216,76; 2400; 2500; 2600; 2750; 2800; 3000; 3328; 3400; 3500; 3820; 4100; 4623; 4714; 4800; 5000; 5100; 6330; 30000 = 50,1) + 3(850,8) +

mediana: 3000 $(127P,1) + 1(17P,8) + 2(27P,8) +$

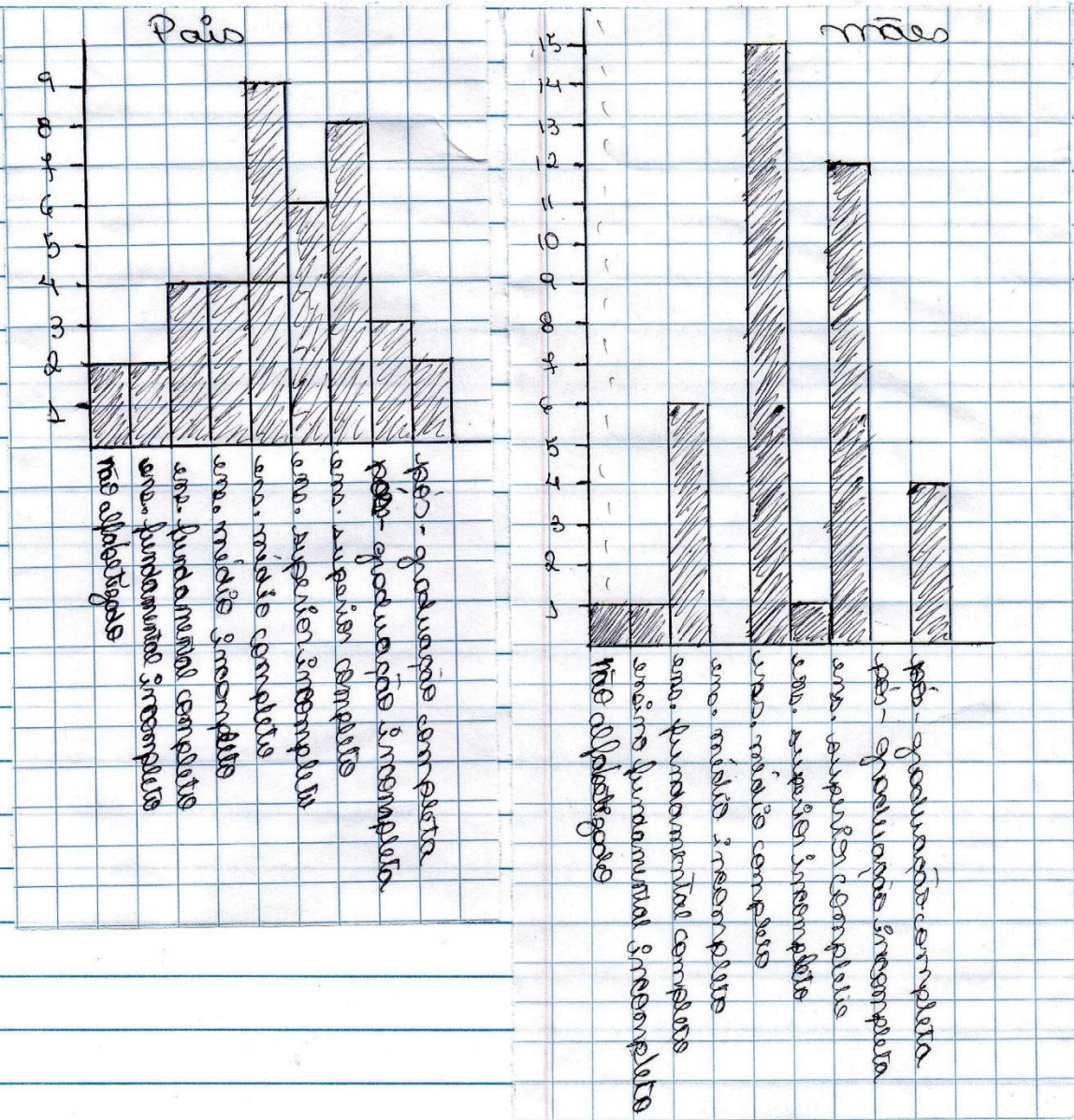


C) moda: 24 pessoas sendo 9 pais e 15 mães

mediana: 1, 1, 1, 4, 6, 12, 15 / 22, 2, 3, 4, 4, 6, 8, 9

a mediana entre as mães é 4 e entre os pais é 4

a mediana é 4



ANEXO 3 – Carta de esclarecimento sobre o projeto de pesquisa



PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO (PUC-SP)

Pesquisa: Processos de Ensino e Aprendizagem de Matemática em Ambientes Tecnológicos

Pesquisador: Saddo Ag Almouloud

Informações sobre o projeto e sobre a pesquisa:

A pesquisa a ser realizada faz parte de projeto em desenvolvimento no Programa de Estudos Pós Graduated em Educação Matemática da Pontifícia Universidade Católica de São Paulo (PUC-SP).

A pesquisa tem como objetivos:

- Dar subsídios teóricos ao trabalho de professores e pesquisadores interessados na integração de pesquisas em Educação Matemática e de tecnologias em aulas de Matemática.
- Produzir conhecimento na área de formação de professores de matemática, utilizando tecnologias como recurso de apoio (como instrumento, no sentido que explicitaremos mais adiante neste texto).
- Elaborar em conjunto uma sequência utilizando o Geogebra e/ou Cabri 3D para ensinar conteúdos matemáticos para ser aplicada com grupo de professores de Ensino Básico e Superior.
- Contribuir para os debates a respeito do papel das tecnologias na formação de professores

Para este fim, serão levantados dados a partir de observações, questionários, documentos produzidos pelos participantes da pesquisa.

A pesquisa intitulada CONCEPÇÕES SOBRE ESTATÍSTICA: UM ESTUDO COM ALUNOS DO ENSINO MÉDIO é subprojeto de Processos de Ensino e Aprendizagem e Matemática em Ambientes Tecnológicos, conjugando assim dos objetivos e procedimentos metodológicos aqui apresentados.

Destacamos que, a qualquer momento, os participantes são livres para deixar de participar da pesquisa, sem qualquer ônus ou penalização de qualquer parte. Caracteriza-se firmemente a participação voluntária.

Os dados coletados serão tratados de forma a manter o sigilo dos participantes, sem possibilidade de identificação dos mesmos.

ANEXO 4 – Atestado de aprovação do projeto

Saúde
Ministério da Saúde

Plataforma Brasil

[Informe o E-mail](#) [Informe a Senha](#) [LOGIN](#)

[Esqueceu a senha?](#) [Cadastre-se](#) v3.0

Você está em: Público > Buscar Pesquisas Aprovadas > Detalhar Projeto de Pesquisa

DETALHAR PROJETO DE PESQUISA

DADOS DO PROJETO DE PESQUISA

Título Público: PROCESSOS DE ENSINO E APRENDIZAGEM DE MATEMÁTICA EM AMBIENTES TECNOLÓGICOS
 Pesquisador Responsável: Saddy Ag Almouloud
 Contato Público: Saddy Ag Almouloud
 Condições de saúde ou problemas estudados:
 Descritores CID - Gerais:
 Descritores CID - Específicos:
 Descritores CID - da Intervenção:
 Data de Aprovação Ética do CEP/CONEP: 19/10/2015

DADOS DA INSTITUIÇÃO PROPONENTE

Nome da Instituição: Faculdade de Ciências Exatas e Tecnologia
 Cidade: SÃO PAULO

DADOS DO COMITÊ DE ÉTICA EM PESQUISA

Comitê de Ética Responsável: 5482 - Pontifícia Universidade Católica de São Paulo-PUC/SP
 Endereço: Rua Ministro Godói, 969 - sala 63 C
 Telefones: (11) 3670-6466
 E-mail: cometica@pucsp.br

CENTRO(S) PARTICIPANTE(S) DO PROJETO DE PESQUISA

CENTRO(S) COPARTICIPANTE(S) DO PROJETO DE PESQUISA

[Voltar](#)

DATASUS
Departamento de Informática de Saúde

Este sistema foi desenvolvido para os navegadores Internet Explorer (versão 7 ou superior),
ou Mozilla Firefox (versão 9 ou superior).

Conselho Nacional de Saúde **SUS** Ministério da Saúde **GOVERNO FEDERAL**
BRASIL
PATIA EDUCADORA