Passion-Driven Statistics*

Rob Easterling
occasional itinerant visiting professor
Cedar Crest, New Mexico
IASE, August, 2011

*Ref: "Passion-Driven Statistics," The American Statistician, Feb.,2010

Overview

- The need for outreach: "We don't get no respect"
- The earliest/best opportunity for outreach: Stat101
- But, Stat101 is part of the problem

Actual Problem from Actual Text

- **Shoe Size Data**
  - recent sales

- *Test the hypothesis:*

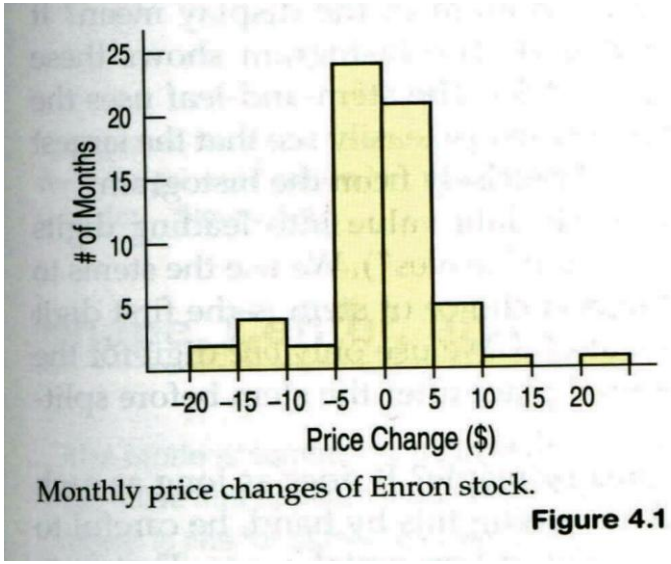
  The "population median" = 7.5

- **?#?*&*@*#? …. Say what!?**

Another Actual Example
Data: Enron Stock Prices

The message here:

Was turned into:



Monthly price changes of Enron stock.

Figure 4.1

Which destroyed the message!

What were they thinking?

Question

- *Do we want our future collaborators, customers, and fellow citizens to "learn" this sort of statistics?*

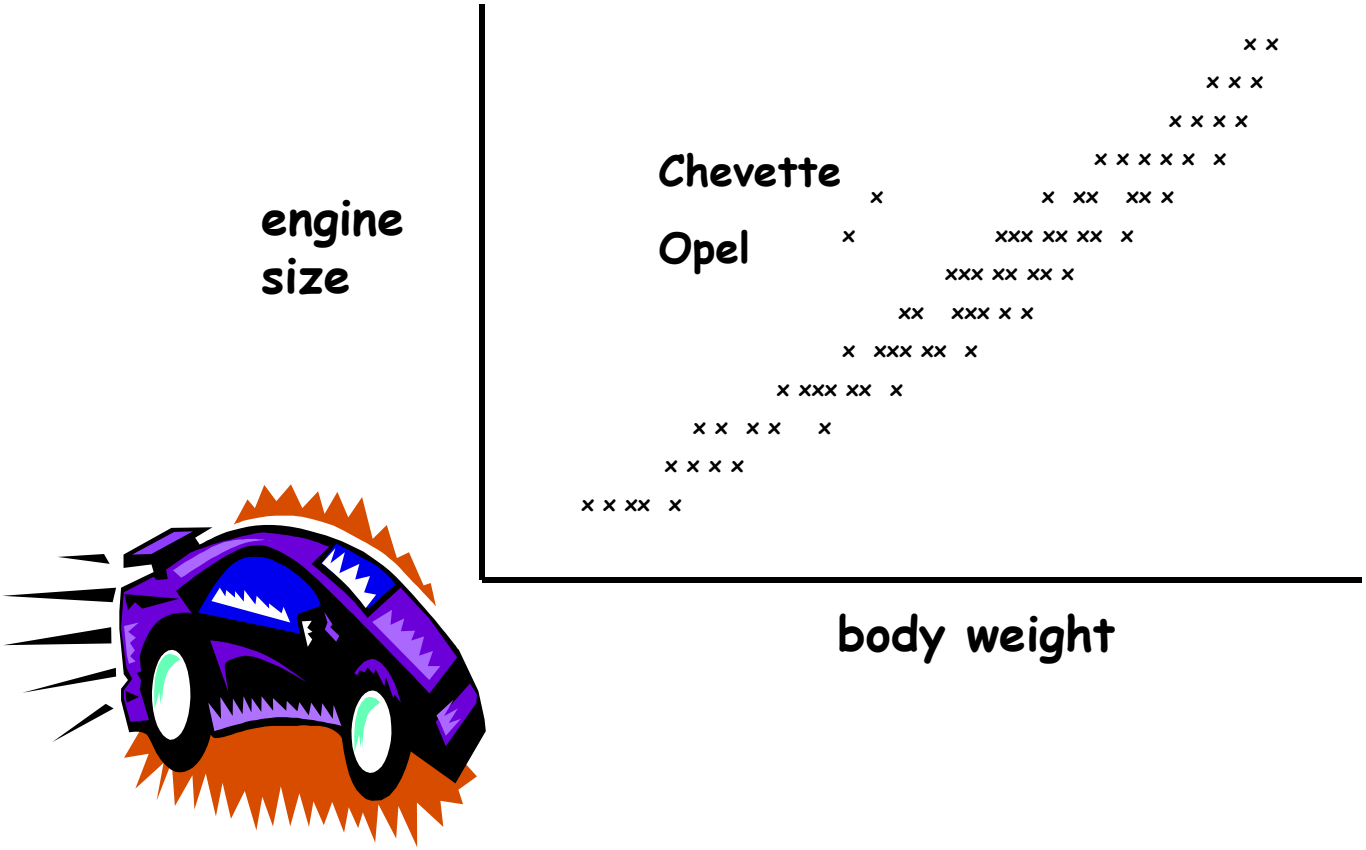
Chances are they do, now!

Comment

- If all we want to teach is formula (or software) plug-in, these problems are as good as any
- If we want students to care passionately about statistics, *especially in their after-university life,* then we need good and sensible examples that a person could really care about

Passion-Driven Statistics:
Charlie Clark and the Car Charts

Chart: car data from Consumer's Reports



Moral of the Car Chart Story

- Charlie found the message in the data, not the authors
  - not because Charlie was a better statistician, but
  - *because Charlie cared more about the data!*
- Subject-matter knowledge and enthusiasm (*passion*) are indispensable to good statistics

Further Comment

- Of course, not all students are interested in car performance, but they all can take a lesson from this example:
  - graphical methods, done right, can help greatly in discovering and communicating interesting facts
  - subject matter knowledge counts!

Nouns and Verbs

- "The nice thing about statistics is that the nouns may change but the verbs stay the same." Carl Marshall, OK State U
- We (statisticians) get turned on by the verbs
- Our clients and collaborators make their living on the nouns
- *We gotta connect 'em!*
  - *in texts, in class, from Day One*

Verbs and Nouns: Archie Bunker
(*All in the Family TV show**)

"Don't give me no stastistics, Meathead. I want facts!"

* based on British show: *Till Death Do Us Part*

Connecting Nouns and Verbs, Facts and Stastistics

- Embed textbook problems in realistic business, political, scientific, … contexts
- Objective: show why someone might really care about the results
- Example: Box, Hunter, Hunter, *Statistics for Experimenters*

Example: Boys' Shoes

Exp. Objective: Compare wear of sole materials A (standard) and B (cheaper substitute).

Experiment: 10 boys; each wears one shoe of each mat'l. L/R random assignment

Response: amount of wear (% of sole worn away after specified time)

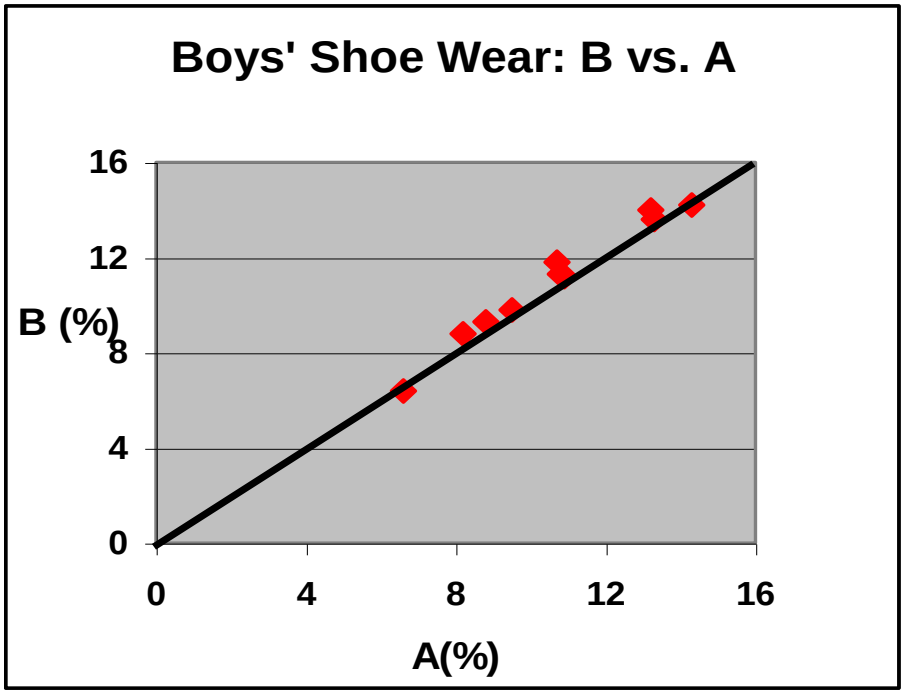
Management wants to know:

Can we get away with the cheap stuff?

Analysis 1. Plot the Data

Results:

A wins most of the time (less wear, 8 of 10 boys) and by a larger amount.



BHH: Paired t-test and rand'n. test give similar results (sig. test: low P) - stastistics

My Extended Story

Suppose shoe-life is defined as 50% wear

B's life would be 2% less than A's life

- e.g., wear out in 51 weeks, not 52 weeks
- upper 95% confidence limit: two weeks diff.

- Surely customers wouldn't notice

But, suppose the company slogan is 'Nothing But the Best'

Switching to B could be first step on slippery slope from loss in quality, loss of reputation… , to ruin

What are you going to recommend to mgt?

Ethics Alert!

- Sponsors, manufacturers, government agencies, managers, thesis advisors, … can all have agendas
  - I'm not being sinister, just realistic
  - You want people to care about experiments
- Statisticians and their collaborators must have the strength of character to assure honest experiments and analyses

Design Issues

- Other Shoe Sole Characteristics
  - comfort, flexibility, squeakiness, …
- Covariables needed
  - boys: age, weight, days-worn, …

Summary – Boys' Shoes Example

- Life doesn't end with a significance test
- The business, scientific, or other contexts (passions!) within which one works are critical to an analysis's meaningfulness, success, and effect

Impressions: Stat 101 Texts?

- **Questions:**
  - context capable of generating passion?
  - illustrate good statistical thinking and practice?

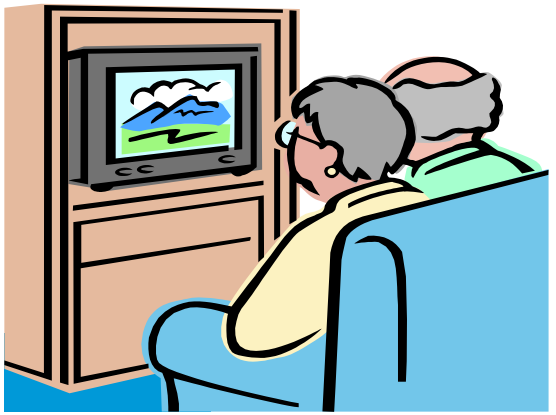
Observation

- Textbooks tend to focus on technique (*verbs - stastistics*), not on the story (*nouns - facts*)

- Analyses seem more driven by the section title than by what you might learn from appropriate data.

Actual Chapter 1 Example from Popular Text

- Histogram of state %s of persons age 65+
  - Passion?
    - *Story: where a young person would not like to live? Boring*
  - Analysis?
    - *loses state identification;*
    - *I would do geographical plot;*
    - *story: gov't. $ distribution*



Bottom Line:
Outreach Begins in the Classroom

- To students who are our future colleagues, customers, and bosses:
- To future (and present) statisticians:
- To the general public:

It's the FACTS!

(subject-matter context)

- That's where enthusiasm for and appreciation of our studies and data-analyses must come from

Responsibilities

- Instructors: We have to *teach* that message. We can't just rely on osmosis or a gradual dawning of awareness
- Authors, Editors, Reviewers: Make sure the message is included and communicated

The Goal: Statistical Heroics

"With time running out, he took an impossibly large amount of data and made something incredibly beautiful!"

(MicroSoft Office commercial)