THE NIGHTINGALE PROGRAMS

Chris Olsen
Iowa Community Schools
Cedar Rapids, Iowa

Introduction

The set of computer programs described in the present paper have been developed since the summer of 1984 in support of a teacher outreach program administered by the Woodrow Wilson Foundation in Princeton, New Jersey, and the Quantitative Literacy project sponsored by the American Statistical Association/National Council of Teachers of Mathematics Joint Committee on the Curriculum in Statistics and Probability. In general, the statistical techniques supported by the Nightingale programs are the Exploratory Data Analysis techniques which have appeared in Exploring Data (Landwehr & Watkins, 1986).

From inception, the Nightingale Programs have been designed to support teachers who would bring statistics and EDA techniques into the classroom. The present writer has used the programs in his own statistics class at the high school level in the United States and has supported other teachers' use in science and social studies classes. They have been field tested over the past two years by students and teachers, and myriad "perfecting amendments" have been offered and taken advantage of. The programs are now released for general use in Applesoft and Microsoft (IBM) BASIC source code, and they are absolutely public domain.

Description of the programs

The current version (4.0) of the programs will:

1. plot up to four horizontal or eight vertical box and whisker graphs;

2. do Stem and Leaf plots, including back-to-back, with one, two, or five lines per stem;

3. draw scatterplots, and median fit and simple regression lines, allowing analysis of residuals;

4. perform median smoothing on time series data;

5. graph histograms; and

6. calculate common statistics.

In addition, programs for file creation and simple editing capabilities allow the user to use interesting data as it appears, rather then having to depend on canned data sets.

Design of the programs – from the Teacher's View

A set of computer programs intended for use in a classroom or in classroom applications has only one client – the teacher. A set of programs which does not allow the teacher to do something heretofore impossible or at least make some task easier has no rightful claim to a teacher's attention. The Nightingale Programs have several features which have been built in with teachers in mind. Teachers are assumed to have no particular computer experience, and not a great deal of knowledge or experience with EDA techniques. Some typical teacher concerns will now be addressed.

The first concern is the wide variety of potential use. A teacher may want to get graphs to show the distribution of test scores, distribute data in graphic form for class discussion, demonstrate graphs for the class, have students input data, or in a statistics class actually teach students to use the programs. Thus the software must be simple to use for the occasional user, as well as powerful enough to do relatively advanced techniques without a great deal of fuss. In addition, a teacher or student should be able to progress from simple data analysis to more advanced data analysis without learning new or different kinds of commands. A screen dump option should be available for teachers who would like hard copy, and if different graphs are desired in real time (that is, during class while the students are easily bored by a blank screen) the graphs should be easily summoned from the disk in small amounts of time. Also, the same set of data should be capable of display in different formats without re-entry of data. For instance, a stem and leaf display and a box and whisker plot should not require two entries of the same data.

The Nightingale Programs have been written and modified with all this in mind. Prompts are intended to be very similar, and they appear in frames which are as homogeneous as possible. The programs are menu driven for ease of use. The menu items and the prompts are written in plain English, devoid of computer or statistical cant as much as possible. In addition a tutorial written in a conversational tone is supplied with the documentation, so that a teacher could almost teach herself without a computer guru's help.

A second concern is the range of acceptable data. If a teacher is going to bring in "real data" to a course of study he will sooner or later encounter data that is fractional, negative, and/or ill-behaved in other ways. The Nightingale programs have been designed and tested on increasingly pathological data in the past two years, and as a set of data pointed to deficiencies in the generality of the data handling capability, the programs were rewritten. That is not to say that any or all fractious sets of data will be digested without a hiccup in the programs – the potential user should regard the programs as reliably housebroken, but not reliable enough to bet the family farm.

A third major concern is graceful recovery from error. The programs should be fault tolerant, especially since computer neophytes will frequently be in the driver's seat – a teacher may, for instance, have students do actual keystrokes during a lecture. Having raised the specter of fault tolerance, one is forced to admit only relative success. During this writer's statistics classes, and over a significant number of teacher in-service presentations around the United States, users were carefully monitored for problems. When common problems were detected steps were taken to minimize the effects of error; however, the programs will not be necessarily tolerant of wild deviations from expectation.

A disturbingly common view among those who have never had to actually write programs for others is that computer programs should have omnipresence (to detect user errors everywhere), omniscience (to read the user's mind) and, of course, omnipotence (to correct the user errors). Among those who have actually written programs for others there is a more reasonable view: user-friendliness must be balanced against available memory, language limitations, reasonable execution time desires, hardware differences, prompt complexity, screen formatting capability, transportability, etc. The Nightingale Programs represent an attempt to balance all these competing values in a manner that seems to work very well in a variety of situations with a wide variety of students and teachers.

Design of the Programs – The Programmer's View

The Nightingale programs have been written with an eye toward modification by others. The code is modular, internally documented, and transportation should be relatively easy. Peeks and pokes are at an absolute minimum, as well as any techniques which seemed specific to a particular disk operating system. The Applesoft version is admittedly less tractable to modification, due to a less capable DOS, a less capable BASIC dialect, less capable I/O, and frankly a less capable programmer. The present writer learned much of what he knows about Applesoft by hammering out the Nightingale programs, and the code will be more stilted due to this lack of experience with Applesoft.

Variable names are descriptive where possible (it is obviously more possible on the Microsoft version), and are consistent across programs. Algorithms are also consistent across programs. Suppose for example that two different scatterplots (say for regression and QQ plots) could have been programmed differently so that each was more efficient. Rather than have different methods for generating points on a screen, the same method would be used so that modification would be easier due to a smaller number of distinct algorithms.

The programs use common files, and raw data files are formatted using the Nightingale Data Protocol, discussed by Swift (1985) and formulated at the Woodrow Wilson National Fellowship Summer Institute in statistics at Princeton in 1984. Temporary files are created by the computer programs as needed and are given characteristic extensions for easy housekeeping should a data disk become cluttered through use. In both Applesoft and Microsoft versions a separate data disk is required. In the Applesoft version, use on a single drive system is supported, but the programs nearly fill one disk and disk shuffling is required. The Nightingale Programs are the product of approximately 2000 hours of coding, refining, testing and debugging. Data is regularly fed into the programs for the present writer's statistics and other mathematics classes; in addition, data is collected for teacher in-service presentations in social studies and science areas. All anomolous results are checked and every program is tested against textbook problems regularly, working from raw data. The rate at which the programs are modified has slowed significantly, and no changes in program format are now anticipated. Since code is not being altered, the entrance of new bugs is halted and thus the total number of undetected bugs should decrease monotonically.

Transportability problems have been minor and few. The programs function on Apple II e's and c's. Apple II +'s are problematical only in that lower case alpha characters are not supported by the machine; code for the Apple family does not differ. The Microsoft version has been successfully run on Leading Edge, Zenith, and Tandy 1000's. Transportability problems for the Microsoft version from IBM to the Tandy 1000 have been in the area of graphics screen parameters. Leading Edge and Zenith have had no problems to date.

Availability of the Programs

The Nightingale programs are available from:

Chris Olsen
1750 34th Street S. E.
Cedar Rapids, IA 52403
USA

The programs are public domain, and users are encouraged to share with others. Prospective users are asked to send $10.00 U.S. and specify whether the Applesoft or Microsoft (IBM) versions (or both) are desired. The ten dollars covers the cost of diskettes, documentation, postage and handling. If a photocopyable copy of the documentation is desired, please so specify. If an Apple II+ version is desired, be sure to specify.

Bibliography

Landwehr, J., & Watkins, A. (1986). *Exploring Data*. Palo Alto, CA. Dale Seymour Publications.

Swift, J. (1985). The Nightingale Network. *Teaching Statistics*, *7*(2), 57-61.