

1 Introduction to Statistical Packages

When I first taught this class (30 years ago!) we used punch cards on the IBM mainframe computer. That was slow, and we could not cover a great deal of material. We evolved into remote submission of card images (that was faster) and then into actual interactive work on remote unix servers. That mode stuck with us for a long time – and was how we did it when I last taught this course about 20 years ago.

Most real commercial statistics packages now run on MS Windows-based computers, some of them solely on Windows (although the real “buffs” probably use linux (unix) based platforms), and that is how we will operate. On my web page I list the packages I commonly use, and all run on Windows (some on unix, but that’s not where I usually run them). The modern platform allows us to be overwhelmingly more productive than in the “old days.” Learning enough computing to use the packages used to be a real issue, but “anybody” can use computers these days.

We will use **Minitab** for this class. This package is available on the computers at the CIRT pod (in the main CIRT building) – although it was mis-installed so not useable as of 8/22/05. It also is available in the ESC-South (Engineering and Science Computing) pod, also run by CIRT, and does seem correctly installed there. Next semester we will use SAS.

If you prefer to not use CIRT facilities, you can **rent** Minitab for \$29.99 a semester and put it on your PC at home or work. To find out more about this, go to the course web page and click on the link for Inexpensive Rentals. I would do it this way if you can.

You can not permanently store files on the CIRT PCs, so any files that you create while doing statistical analysis or word processing must be saved to the floppy drive, to a USB drive, or on your briefcase@myunm.edu. Ask a consultant how to use the last method.

Introduction to Minitab

MINITAB is a complete package for data summary and analysis. Version 14 (the most recent version) is available on the CIRT machines. There are some fairly substantial changes between Versions 13 and 14, but if you have the earlier version you should be alright. I still use it in my work at HSC.

1. Strengths - easy to use and learn, interactive.
2. Weakness - limited data management facilities. Not appropriate for very large data sets (census tapes, etc.). It’s getting better, though.

We used to have to scrounge around on the web or write our own introductions to Minitab (or assign an expensive book!), but now there is a good introduction that installs with it, and also is available for you to download (they don’t let you save the one that installs). We are going over that tutorial during the first lecture. To bring up the tutorial, follow the path

Start > All Programs > Minitab 14 > Meet Minitab.

After the first lecture, I hope you have seen

1. How to start up Minitab
2. The various windows, icons, and bars that pop up with MINITAB: Data Window, Session Window, Information Icon, History Icon, Menu Bar, and On-Line Help.

3. Entering data in the Worksheet
4. Generating Simple Descriptive Statistics
5. Generating Graphical Summaries
6. Transforming Data
7. The Command Line Editor
8. Basics on saving data, output, graphs

As a first assignment you should go through the tutorials , then go to the CIRT pod (or wherever you are using it) and try things out. Minitab is very easy to use, so I will not spend much time teaching you about the program. The subsequent notes describe how many of the data analytic procedures are implemented in Minitab.

You should make extensive use of the tutorials and the on-line help! There are lots of other resources on the Minitab site and off links on the course web page.