

An introduction to S-PLUS/R and SAS. These two very distinct statistical software packages are commonly used in the modern professional practice of statistics and are complementary in their strengths. S-PLUS is a computing environment for graphics, programming, statistical analysis and mathematical computing. SAS is a production tool for high-volume data management and statistical analysis. The first half of the term will cover S-PLUS/R and the second half will cover SAS. Examples from the practice of applied statistics will be used throughout the course.

Textbook: *STAT 462/862 Course Notes*

Several online manuals will be available and referenced throughout the course.

Instructor: A. Day

Prerequisite: CISC-101* *or* equivalent.

Corequisite: STAT-361* *or* permission of the instructor.

Evaluation:

| | |
|--------------|-----|
| Homework | 40% |
| Midterm test | 30% |
| Final test | 30% |

Topics:

1. S-PLUS

- Creating and working with data objects (vectors, matrices, lists, data frames, factors and time series)
- Introduction to statistical functions and data analysis
- Writing functions
- Programming techniques and computational efficiency
- Random numbers and simulation methods

2. SAS

- Introduction to SAS: the SAS programming environment, data sets, the output delivery system, and global options
- The DATA step: reading in data, creating data sets, manipulating data sets, deriving variables and generating text files
- Commonly used SAS procedures for: sorting and transposing data, printing, tabulating, formatting and obtaining summary statistics
- Introduction to statistical procedures and data analysis
- The SAS macro language
- Structured Query Language (SQL)
- Advanced topics (if time permits): call execute, IML (interactive matrix language), and the analyst application