

**Syllabus**  
**Statistics 627: Professional Skills in Data Science**  
**Fall 2015 (1 cr)**

Instructor: Brian Yandell with Doug Bates

Phone: 263-3304

e-mail: [brian.yandell@wisc.edu](mailto:brian.yandell@wisc.edu)

Office Hours: Immediately after class or by appointment.

Based on course taught in Fall 2014 as Stat 692, Network Skills for Statisticians, and on open course notes of Jenny Bryan, University of British Columbia.

**Course Description:** This course is aimed at providing statistics graduate students with an understanding of and experience with important aspects of professional development in statistics, including skills with internet tools, sophisticated use of statistical languages (such as R) and other emerging topics.

**Learning Objectives:** After completing the course, a student will be able to

- use RStudio as platform for statistical computing
  - install and use RStudio application on personal computer
  - use RStudio on departmental computers
- use R Markdown language for literate programming
- understand data representation in R, including reading data, working with factors and data frames, and manipulation and display of data
- visualize data with the ggplot2 package for R
- organize work into well documented collections such as R packages
- make sound choices about algorithm efficiency and graphical presentation
- understand aspects of underlying statistical computing algorithms, including
  - the R formula language and its use for fitting linear models and generalized linear models
  - contrast specifications for factors and their impact on interpretation of coefficient estimates
- appreciate benefits and challenges of other statistical language choices

Computer software: R, RStudio and related applications.

Text: Reference materials will be provided online.

Assignments: There will be periodic assignments.

Exams: There will be no exams.

Grading: Based on homework (80%) and class participation (20%). There will be homework every other week.

Optional final project will replace one homework.

**Reading material** can be found at:

<https://ay15-16.moodle.wisc.edu/prod/course/view.php?id=280> (Fall 2015 Moodle)

<http://stat545-ubc.github.io> (Jenny Bryan's course)

<http://www.stat.wisc.edu/network-skills> (Web pages)

<http://www.stat.wisc.edu/network-skills/stat692-notes> (Yandell notes from Fall 2014)

<https://github.com/dmbates/stat692> (Bates GitHub site with his notes)