

# Introduction to Statistical Modeling in Psychology

PSYC G6006.001  
Spring 2017

*Lab Meeting:* TBD  
*Lecture Meeting:* Friday 12:10PM – 2:00PM  
*Location:* 405 Schermerhorn

## **Instructor:**

Becca Franks, PhD  
beccafranks@gmail.com

## **Teaching Assistant:**

Matti Vuorre  
mv2521@columbia.edu

## **Textbook and Readings:**

- ❖ Fox, John. (2008). *Applied regression analysis and generalized linear models (3<sup>rd</sup> ed.)*. Los Angeles, CA: Sage Publications.
- ❖ Supplemental reading to be provided throughout the term.
- ❖ The Internet (using it well is central to learning R)

## **Course Description:**

Using the freely available and acclaimed computing language, R, this course will cover the necessary basics for statistical modeling in the applied sciences, from plotting, to model specification, to write-up. In particular, be a practical introduction to linear regression, ANOVA, generalized linear models, and model diagnostics.

We will introduce topics in Friday lectures and discuss them in the following lab section. Students are expected to attend and participate in both.

Throughout the semester, students will conduct analyses on their own data. At the end of the semester, each student will organize his/her analyses into a final presentation and paper.

## **Assignments and Grading:**

After each lecture, students will have the remainder of the week to prepare the related lab, which will be due *the night before* the following lab section. For each lab assignment, students will (a) generate two discussion points from the reading, (b) submit one interesting blog post from R-bloggers and brief explanation of why it is of interest, (c) answer any assigned problem sets, and (d) complete an analysis of their own data.

The lab assignments are due the night before lab (email to Matti). There are 12 assignments worth 5 points each making up 60% of the grade. Students will receive 5 points for complete on-time assignments, 4.5 points for assignments handed in between lab and lecture, and 4 points for complete late assignments (handed in after lecture) and 0 points for incomplete assignments.

At the end of the semester, students will summarize the analyses of their own data in a 10-15 minute class presentation and written paper. The paper should follow the format of a specific journal of their choosing, providing a brief introduction to the topic, a detailed methods section, a thorough results section, and a concise discussion.

Presentations will be given in the last class and will be worth 20 % of the grade (all students are expected to attend all presentations).

Papers are due a week after the last class and will be worth 20% of the grade.

### **Details of the Lab Assignments:**

#### *(A) Response to the reading*

Students will complete the assigned reading and generate two responses. Each response should be a short paragraph in length and could be points that the student found interesting, a discussion of a section that was unclear and why, a connection to a previous lecture, the definition of a new vocabulary word and how it connects to the course material—anything that demonstrates active engagement with the text.

#### *(B) R-bloggers*

Students will sign up for daily emails from R-bloggers (<http://www.r-bloggers.com/>), a website that collects daily news and tutorials about R from various blogs. It will give students a sense of the R-community and will occasionally provide useful information. Every day of the week, students should scan the ‘headlines’ of R-bloggers to see if there is anything of interest (e.g. a cool plot, models that might be of use, subject matter that is interesting, etc.). They will pick one blog posting each week and write a brief comment on it (2-3 sentences). Only one posting is due each week, but students should read the headlines every day to make sure they don’t miss something of interest.

#### *(C) Problem sets*

For most labs, students will be required to answer one-two textbook questions.

#### *(D) Data analysis*

Following each lecture, students will submit an analysis of their own data in *R Markdown* that illustrates the analysis method presented in the lecture (if students do not have their own data, example datasets will be made available). Answers should conform to the following format:

- Explain the goal of the analysis for the particular dataset.
- Write several sentences interpreting the analyses/results/graphs

#### *(E) What confused you this week?*

Please write about at least one thing that confused you in the reading, lecture, R, or a combination. The more specific, the better.

Week	Lecture Dates	Events	Lecture Topic
1	Jan 20	<i>Classes start Jan 17</i>	Introduction to R and Stats
2	Jan 27		Introduction to Regression
3	Feb 3		Examining, Transforming, and Graphing
4	Feb 10		Cleaning Data
5	Feb 17		Linear Least Squares Regression
6	Feb 24		Statistical Inference for Regression
7	Mar 3		Dichotomous Predictors and Coding
8	Mar 10		Interactions
9	Mar 17	<i>Spring Recess</i>	
10	Mar 24		Diagnostics
11	Mar 31		Logistic Regression
12	Apr 7	<i>Project lab workshop</i>	The Generalized Linear Model
13	Apr 14	<i>Project Meetings</i>	Mediation/Power/Alpha Multivariate
14	Apr 21	<i>Becca Away</i>	Bayesian Statistics
15	Apr 28		Presentations/Flex
16	May 5	<i>Last day of classes May 1</i>	Presentations/Papers due

### **Additional Resources:**

- Quick-R: <http://www.statmethods.net/>
- Some free R-tutorials from <https://www.datacamp.com/>
- UCLA has a helpful statistics website: <http://www.ats.ucla.edu/stat/>
- John Fox's R Companion website: <http://socserv.mcmaster.ca/jfox/Books/Companion/>
- R-bloggers on learning R: <https://www.r-bloggers.com/how-to-learn-r-2/>
- A good matrix/linear algebra class on YouTube: Strang's 18.06 Linear Algebra