MTH/SDS 291 - 01: Multiple Regression

**Instructor:** Lu Lu

**Office:** McConnell 105

**Email:** llu@smith.edu

**Class meetings:** Tu.Thur 10:30am-11:50am, in Sabin Reed 301

**Other sections:**

*Andy Reagan (andy@andyreagan.com, Bass 218)*

Tu.Thur 10:30-11:50am, in McConnell 104

**Office hours:** Monday 1-2:30pm

Wednesday 1-2:30pm

Thursday 9:30-10:20am

or by appointment (in McConnell 105)

**Class materials:** Available at our Moodle course site

**Textbook:** *STAT 2: Building Models for a World of Data. New York: W.H.Freeman, 2013.*

by Cannon, Cobb, Hartlaub, Legler, Lock, Moore, Rossman, and Witmer.

It is available at the bookstore in the Campus Center. You may also acquire it from other places. Note that you do not need the online access code.

**Prerequisites**:

An introductory statistics course, e.g. MTH/PSY 201, GOV 190/203, MTH 219, MTH/SDS 220, ECO 220, SOC 203, or the equivalent or a score of 4 or 5 on the AP Statistics exam or permission of the instructor.

**Course description:**

In this course, students will learn how to use the four-step process (choosing, fitting, assessing, using) to build statistical models to analyze real world data. These models can be used for predictions, understanding relationships and assessing differences. Students will also learn how to communicate their findings effectively to a non-statistical audience.

Topics covered include simple and multiple regression, logistic regression, model selection, ANOVA, multicollinearity, analysis of residuals and outliers. We will use R-Studio, a statistical software package, to explore many real sets of data and to analyze them. Over the course of the semester, you will learn how to use this software effectively and how to interpret the output from it.

# Evaluation in MTH/SDS 291

**Grades:** Your course grade will be determined as follows (although this is subject to change):

Writing components of course grade:

Homework 25.0%

Quizzes 5.0%

Project 25.0%

Exams 40.0%

Participation 5.0%

Total 100.0%

# Homework: Homework represents an important aspect of the course. Readings and selected problems will be assigned regularly. The assigned problems and due dates will appear in the Course Schedule on Moodle. The assignments are designed to thoroughly test your knowledge of topics we discuss in class and are definitely not meant to be one-night jobs. Check Moodle for between-class announcements, hints, or modifications in assignments.

Homework should be completed in **RMarkdown.** All assignments will be due by 5pm of the due date. A hard copy of the knitted file should be submitted. The RMarkdown file should also be submitted via Moodle. Late homework turned in within 48 hours will be accepted with a 25% penalty. No credit will be given after 48 hours unless you are in extenuating situations with verifiable documents.

# Although you are encouraged to discuss problems with each other, I expect each person to submit his or her own work. You may choose to work with a partner on data analyses – that’s okay, but I want written interpretations and responses to be in your own words. You should always write up your own solutions. It is not acceptable to copy down your classmates’ work and in fact, doing this constitutes academic dishonesty.

# And remember, an answer to a statistics question is almost never just a number – I want thoughtful explanations and conclusions! In order to receive full credits on a problem, you will need to provide clear explanations and necessary steps to justify your answers. If your work is mostly correct, you will receive most credits. On the other hand, if you do not provide enough work, or follow wrong steps but happen to get a correct answer in the end, then it is likely that you receive no credit or very few partial credits.

**Quizzes:** There be weekly quizzes on Moodle. The purpose of the quizzes is to give you instant feedback on your learning. The quizzes will cover the materials discussed in the previous week and in some cases, the assigned readings for the following week. They will be due by midnight(11:59pm) of the due date. *There are no make-up quizzes.* At the end of the semester, the lowest quiz score will be dropped.

**Exams:** There will be two in-class exams. They are closed-book. You may bring a scientific calculator and a cheat sheet to the exam. More details will be forthcoming close to the exam dates.

Exam 1: October 18, Thursday

Exam 2: December 6, Thursday

**Project & Presentation:** An important part of this course is a group project to be completed in a group of 3 students. You will pick a question that is of interest to you and use the skills that you’ve learned throughout the semester to answer the question. This is a *team*-oriented task that will require extensive writing, data analysis, critical thinking, and an oral presentation. More detailed information about the project will be distributed at a later date.

**Software:** We will be using the software package, R-Studio, to analyze data. Today, virtually all data analysis is done with the aid of at least one computer software package – R is one of the most powerful and most flexible ones – it’s also probably the most widely used of all statistical packages. You can access R-Studio via the web at:

<http://rstudio.smith.edu>

**Class Preparation and Participation:**

Attendance to all of the classes and scheduled course events are required. You are responsible for any work missed. If the absence is unavoidable, please communicate with me in advance concerning such matters.

Your active engagement in class is an essential part to successful learning. In classes, I will do a mixture of lecturing and discussion. I might also ask you questions. If you are called upon but do not know the solution, try to give a partial answer. Your answer doesn’t have to be correct, but I appreciate that you make the effort to participate in the discussion.

Try to bring a laptop to the class as we might need it for class activities. However, it should not be used for activities not related to the class such as web browsing, chatting, etc. Please respect your classmates and help create a productive learning environment. In particular, please keep your cell phone in the silent mode during class.

**Other Resources:**

TA help sessions will be held Sunday-Thursday 7-9pm in McConnell 301.

There are statistics counselors and data assistants in the Spinelli Center for Quantitative Learning (Seelye 207).

In addition, the Jacobson Center (Seelye 307) provides writing services and resources for students. You may find it helpful when you draft the final paper for the project.

**Academic Honor Code Statement:**

“Smith College expects all students to be honest and committed to the principles of academic and intellectual integrity in their preparation and submission of course work and examinations. Students and faculty at Smith are part of an academic community defined by its commitment to scholarship, which depends on scrupulous and attentive acknowledgement of all sources of information, and honest and respectful use of college resources”.

**Acknowledgement:** Many materials used in this course have been developed by previous course instructors, including Ben Baumer, Amelia McNamara, Miles Ott and Ben Capistrant. I would like to thank them for sharing these materials with me.