Instructors: Nicole Lazar    Lily Wang
Office: Statistics Building 224    Statistics Building 223
email: nlazar@stat.uga.edu    lilywang@uga.edu
Phone: (706) 542–0632    (706) 542–0455
Office Hours: By appointment

Teaching Assistant: Alex Lyford, Statistics Building 261
email: ajlyford@uga.edu
Office Hours: By appointment.

Class will meet Monday and Wednesday, 2:30–3:20 PM in Statistics Building Room 306.

INTRODUCTION

The capstone course has several goals: to introduce you to a variety of advanced statistical data analysis techniques, to give you experience working with a real data set (and with real scientists who care about the outcome of your analysis), and to help improve your communication skills – both written and oral.

It is increasingly important that graduating Statistics majors have experience beyond the textbook; that is, experience with analyzing a real data set. Real data sets are messy, they have missing or problematic observations, it isn’t always clear how to analyze them. None of these issues are faced by students who learn all of their data analysis from textbooks, where data sets are clean and analysis is obvious (in most instances), since problems are assigned to chapters that focus on particular methods. Furthermore, it is a useful skill to be able to interact with a scientist who has a real problem to solve. This skill can truly only be acquired through experience.

Our national statistical society, the American Statistical Association (ASA) provides
curriculum guidelines for undergraduate programs in the statistical sciences. Some of the recommended skills ASA advocates for a successful Statistics undergraduate are: effective technical writing and presentations; teamwork and collaboration; planning for data collection; data management. At the end of this year-long course, you should have gained the requisite experience and acquired these skills, thereby making you fully a part of the statistical community.

**COURSE OUTLINE**

This course will be more “free form” than you might be accustomed to, due to its nature. We will start with a review of exploratory data analysis, basic regression and analysis of variance (ANOVA), and progress from there, according to the interests of the class and the instructors. There is a large range of topics we can decide to cover, including: generalized linear models, time series, modern regression, bootstrap and other resampling methods, data mining, classification and regression trees, factor analysis, principal component analysis, and more! The idea is not to delve too deeply into any one topic, but rather to give you an overview of the techniques, how they can be used, and, finally, to give you a chance, via homework sets, to apply them to data.

In addition to the advanced topics modules, a major component of this course will be working on real data problems. There will be two aspects of this.

1. We have been contacted by the Campus Transit Authority to help them with evaluating use of municipal buses by campus riders. As you probably know, students, faculty, and staff can ride city buses for free with the swipe of their UGA ID card. The University pays Athens Transit for campus riders (that’s how you get to ride free). Trips that originate and end on campus presumably should not be charged to the University, since those rides could also have been taken on campus buses in many instances. Campus Transit Authority is interested in estimating the number of passengers per year on city buses whose rides start and end on campus. Everyone in the class will work on this project, helping to build the sampling frame
(i.e. working out a list of bus routes and times), design a sampling scheme, collect
data, analyze the results, and write up a final report of results. Most likely we will
split you into teams, with each team being responsible for certain aspects of the
work. We anticipate that work on the Transit Authority project will last all year,
although some groups will be doing more work early, while others will be doing
more later.

2. In addition, sometime around the middle of this semester we will be inviting several
researchers from around campus to present data problems to the class. Working in
groups of three or four, each of you will tackle one of these projects. **If you have
a data set that you think might make a suitable project, either for your
group or for another group in the class, please let us know soon!** We are
happy to have you analyze data that you might have collected for another class,
for an internship, *etc.*

So, roughly from the middle of this term until the end of Spring term (in 5020), we will
be working on two parallel tracks – learning new statistical methods, and performing
actual data analysis.

The Capstone Course is also a part of the University’s “Writing Intensive Program” (or
WIP). The goal of the program is to incorporate innovative techniques to help students
improve their written communication skills. One implication is that assignments and
project reports will be graded not only on the quality of statistical analysis, but also on
the quality of writing. We will also have special writing assignments in addition to the
more statistics-oriented homework. Another implication is that we have been awarded
a special TA, Alex Lyford (see above), who will be working full-time as a TA for this
course. He will be very involved in the planning of the course and in improving your
writing and communication skills. He will also help to supervise all project work.

**TEXTS**

There are no required texts for this course. Regular attendance is therefore all the more
critical, since you won’t be able to make up a missed class by “reading the book.”

GRADING

We plan to assign homework on each of the advanced topics that we will cover in the class, which means that there will be an assignment due roughly every two weeks.

At the end of this semester, each group will be required to make a short (15 minute) presentation of the project data set, covering the subject matter, some questions of interest and simple exploratory data analysis (basic plots and tables). The purpose of this presentation is two-fold: (1) to introduce the rest of the students to the details of the projects and (2) to guarantee that you are able to access and work with your data set. Also, you will be asked to turn in a short (3–5 page) written report. We will talk about both of these in more detail later in the semester. These will lead in to more in-depth presentations and reports to be made during the Spring semester.

Similarly, progress on the Transit Authority project will be presented orally and in writing at the end of the term. We will give you more information on the format later.

Grading will be based on participation in class, homework, the presentation of the data set for analysis, and the written report. In addition, you will be asked to evaluate the performance of your team members, including yourself. The breakdown of grading is Homework: 35%, Class Presentation: 20%, Athens Transit Project: 10%, Participation: 10% Written Report: 15%, Topic/Group Preference: 5%, Peer Evaluation: 5%.

GROUND RULES

1. Late option: Each of you has the option of moving the deadline by one week on a homework assignment. You can exercise that option only once this term. We suggest you save it up so that you can re-schedule an assignment deadline when you really have a crisis. This option applies only to individual homework, not to any group assignment!
2. No help will be given on homework in the 24 hours before it is due. We will always
give you plenty of time for assignments. Assignments in this class will often be
more “open ended” than you are used to from previous courses, so you should start
them ahead of time to do a good job.

3. Since your active attention and participation in class are required, there will be no
texting, web surfing, Facebooking, and so on during our time together. This should
go without saying, but we are saying it anyway!

The course syllabus is a general plan for the course; deviations announced to
the class by the instructors may be necessary.

Students with a disability or health-related issue who need a class accommodation should
make an appointment to speak with the instructors as soon as possible.

As a University of Georgia student, you have agreed to abide by the University’s aca-
demic honesty policy, “A Culture of Honesty,” and the Student Honor Code. All aca-
demic work must meet the standards described in “A Culture of Honesty” found at
www.uga.edu/honesty. Lack of knowledge of the academic honesty policy is not a rea-
sonable explanation for a violation. Questions related to course assignments and the
academic honesty policy should be directed to the instructor.