**FDN 112J /MAT111**

**Spring 2016**

**Probability and Statistics**

**Instructor Information**

Dr. Homer S. White, 115 ASC, extension 8307, email hwhite0.

Office Hours: MWF 9-9:50am; TuTh 2:25-3:25 pm.

**Course Description**

This course is designed to build upon students’ foundational skills in academic inquiry, analysis, argument, critical thinking and discussion, and expression of ideas, as derived from FDN 111. The course will contain modestly interdisciplinary content, though it may be taught within a specific department.

The section of the course in which you have enrolled is an introduction to statistics, including graphical and numerical descriptive techniques, design of observational studies and experiments, basic probability in connection with the distribution of sampling statistics, and inferential procedures. Applications are taken from the social and life sciences. Statistical techniques are studied in an open-source computational environment that thoroughly integrates statistical analysis with the emerging paradigms of literate programming and reproducible research. In layman’s terms, this means that you learn to use totally free software to produce compelling data-analysis reports easily and with a minimum of fuss, in a way that facilitates collaboration with team members and permits colleagues to verify every step of your analysis.

**Text and Materials**

* **Required Text and Materials:** All materials are free and are available on the course website (<http://statistics.georgetowncollege.edu>) or on the Server.
* **Strongly Recommended:** Your own computer. You will have a desktop computer during class, but if you want to bring your own laptop, that’s fine. Everything has to be done by computer, so you will find it a great convenience to have your own machine for out-of-class work. If you do not have your own machine, then you can use any computer that is connected to the Internet.
* **Access to Server Course Materials.** Course materials are stored on the R Studio server: rstudio.georgetowncollege.edu. (Look under … ,then enter **/mat111**.

**Course Objectives**

Both content and skill-cultivation will be important in this course. This course aims to develop the five skills taught in the Foundations 111 course. Recall that the FND skills are described as follows:

* READ (that is, to examine carefully and evaluatively) written and other kinds of texts for content and meaning and, to some degree, to attend to questions of structure and form as they impact and/or shape meaning.
* COMPOSE logical, original, and critical (that is, objective, evaluative, and analytical) responses to texts, problems, and questions.
* ARGUE coherently and persuasively, displaying knowledge of the elements and inner workings of argument—that is, using logical organization, adequate support, independent thought, and sound reasoning, as modeled in many of the arguments in course readings.
* ENGAGE, deeply and meaningfully, with course material—including texts, problems, and questions—and with others in critical discussions of that material, both in the sense of participating in college‐level class discussions, and in the sense of entering into written and other kinds of larger cultural discourses at a pre‐disciplinary level.
* SELF-ASSESS by reflecting thoughtfully on both their understanding of course content and their development of target skills; in other words, self-assessment here refers to the skill of “knowing when we know something.”

This course aims to development Foundations skills in the setting of data analysis rather than the study of classical texts. In particular students will learn:

* to **Read** a dataset, and the various graphical and numerical summaries of the dataset:
  + in light of its *structure* (e.g., the nature of the variables recorded, the structure of the graphical devices, etc., and even the structure of computer-programming objects associated with data);
  + in *contex*t (e.g., where, how and why the data was collected, its relationship to other data and the questions attending upon that other data);
  + for *interpretation* (e.g., to summarize the data and to discover patterns present within it);
  + in a spirit of *critical engagement*, by asking such questions as:
    - Does the pattern I found in the data exist in the population from which the data was gathered, or could the pattern reasonably be ascribed to chance variation in the process of collecting the data?
    - Was the data collected in a way that makes it likely to be representative of the population?
    - Are the numerical and graphical techniques at my disposal sufficient for exploration of the data, or should new techniques be developed? How might I modify an existing technique to develop a technique suitable for this data?
    - Are there interesting features of the data that suggest research questions other than the questions that motivated the collection of the data in the first place?
* to **Argue** about patterns present in data, and about what inference might be drawn from the data. Specifically, we aim develop students’ ability to:
  + *Describe* arguments by others made concerning data;
  + *Evaluate* those arguments;
  + *Construct* their own arguments. For example, we aim to develop facility in the chains of statistical reasoning that underlie point estimation, interval estimation, tests of significance and the evaluation of predictive models.
* to **Compose** analyses of data in a well-organized way, incorporating numerical and graphical evidence with textual description and argumentation. Elements of this skill – such as literate programming, production and interpretation of graphs, interpreting numerical output of statistical procedures, etc. – will be developed in two out-of-class homework assignments that take the form of miniature but multi-faceted data reports, as well as in a more elaborate team-based data analysis project.
* to **Engage** course material with energy and persistence, through active participation in class and development of independent study habits (see Course Requirements below).
* to **Self-Assess**by reflecting thoughtfully on both their understanding of course content and their development of targeted skills (see Course Requirements below).

**Course Requirements**

Your grade will be determined by your performance on quizzes, homework, tests, a final exam, a project, and an engagement component.

***Engagement: Daily Assignments (DAs)****.* On each class day you will turn in a small assignment on the R Studio Server. Near the end of the semester I will check them. Most are graded as either turned in or incomplete/missing, but a few are selected for closer scrutiny to make sure you are putting in the proper effort. **DAs are worth 5% of your course grade.** Skills addressed: Read, Argue, Engage.

***Online Quzzes (OQs).*** These are posted on Moodle. There are very strict deadlines, so if you have to be away from the course due to travel or some other reason, make sure to work on the quiz well ahead of time. You are allowed three attempts on each quiz, with your highest score counting toward your grade. **OQs count for 15% of your course grade.** Skills addressed: Read, Argue, Engage, Self-Assess (through feedback on multiple attempts at quizzes).

***Preliminary Data Analysis Reports (DARs)***. After you have attained a certain level of proficiency in R and in basic descriptive statistics, you will begin to write small data analysis reports. Each DAR addresses the same Research Questions pertaining to a particular dataset. When you write the second DAR you are working on the same Researech Question and dataset, but by that time you will have additional analytical tools at your disposal. Thus the first DAR plays the role of a “first draft” for the second DAR. The feedback you get from me on your first DAR will be helpful in producing the best possible second DAR. **This is the “process writing” that is an essential component of any Foundations class.** Although the DARs both address the same Research Question they are graded separately. Hence you want to do the best job you can with your first DAR. You turn in each DAR as an R Markdown document by leaving a saved copy in the Submit directory in your Home folder on the R Studio server, and you turn in a printed hard-copy of a “knitted” pdf version in class.

You may consult with me, with fellow students and with Help Session Leaders about the DAR assignments, but everyone must write up his/her own DARs without ANY copying from others. Strong evidence of copying from others results in a zero grade for all parties involved.

Together the two **DARs are worth 10% of your course grade.** You should also be aware that you must submit the “knitted” pdf version the second DAR to Live Text for purposes of assessment. This is separate from, and in addition to, placing it in the Submit folder on the server.

Skills addressed: Read, Argue, Compose, Engage, Self-Assess.

***Tests (T1 and T2)***. There will be two semester tests, **each worth 15% of your course grade**. If you must miss a test,then you must let me know in advance, if it is humanly possible for you to do so. If I judge that your excuse is acceptable, then we will try to arrange a time for you to take the test separately, although usually this is not possible. In the event that you are unable to make up the test in this way, then your final will count an additional 20%. Skills addressed: Read, Argue, Engage.

***Final (FE)*.** The final exam is worth **20% of your course grade**. It is cumulative, but is slanted somewhat toward material covered since T2. Skills addressed: Read, argue, Engage.

***Final Project (FP)***. The final project is substantial data-analysis report. You will work in team of approximately four people per team. Part of your grade on the project depends on your team-members’ assessments of your effort. More details, including the specific project options, a timeline and a grading rubric that addresses the **Read**, **Argue** and **Compose** skills, will be made public in due course. The final project is worth **20% of your course grade.** Skills assessed: Read, Argue, compose, Engage, Self-Assess.

*Course Grade Calculation*. Your course percentage is computed as follows:

**Course = 0.05\*DA + 0.15\*OQ + 0.10\*DAR + 0.15\*T1 + 0.15\*T2 + 0.20\*FE + 0.20\*FP**

The grading scale is as follows:

A 92.5% and above

A/B 87.5% to 92.5%

B 82.5% to 87.5%

B/C 77.5% to 82.5%

C 70% to 77.5%

D 60% to 70%

F below 60%

**Final Exam Date**

Saturday, May 3, 9-11 AM, in LRC 012.

**Course Outline**

The following list indicates the course content and the approximate placement of tests within that content:

* Chapter One: Introduction to R, R studio, and the Basic Ideas of Statistics
* Chapter Two: Basic Descriptive Statistics (numerical and graphical) for:
* one factor variable
* relationship between two factor variables
* one numerical variable
* relationship between one factor and one numerical variable
* Chapter Three: More on Two Factor Variables
* Further concepts for Descriptive Statistics
* Introduction to Inference, with the Chi-Square Test
* Chapter Four: Relationship Between Two Numerical Variables
* Scatterplots
* Linear Models (descriptive approach only)
* **Test #1**
* Chapter Five: Sampling to Gather Data
* Chapter Six: Design of Studies (Experiments and Observational Studies)
* Chapter Seven: Probability for Statistics
* Basic Ideas of Probability
* Random Variables
* Expected Value and Standard Deviation
* Two Important types of Random Variables: the Binomial and Normal Families
* Chapter Eight: Random Variables in Sampling
* Five Basic Parameters
* one mean
* difference of two means
* mean of differences
* one proportion
* difference of two proportions
* Estimators for these parameters, and their probability distributions
* Chapter Nine: Confidence Intervals for the Basic Five Parameters
* **Test #2**
* Chapter Ten: Tests of Significance for the Basic Five Parameters
* Chapter Eleven: Inference for One Factor Variable
* The Chi-Square test for Goodness of Fit
* (As time permits)
  + - Getting Data
* Multiple Regression
* Inferential Aspects of Linear Models
* **In-Class Project Presentations**
* **Final Exam**

**Miscellaneous Course Policies**

* **Attendance**. Attendance will be taken regularly, and is part of your Engagement Grade. Late arrivals may be counted as absences. Tests and quizzes missed due to unexcused absences or lateness cannot be made up, and homework turned in late is either penalized, or – at the discretion of the instructor – not accepted at all.
* **Academic Honesty**. On in class-assessments (in-class quizzes and tests) you may use only materials that are explicitly permitted. For out-of-class assessments (Electronic HW, out-of class quizzes, and the Project) you consult with me, with fellow students and with Help Session Leaders, but everyone must write his/her own assignment without copying from others. Strong evidence of copying results in a zero grade for all parties involved.
* **Office hours** (see under "Instructor Information") are very important to college professors. During my posted office hours I am in my office for no other purpose than to help you. Therefore you never need to make an appointment to see me during office hours: just stop by. If you cannot make my office hours, consider stopping by anyway. Quite often I will be too busy to see you, but sometimes I can put off my other tasks. Remember: to the faculty, you are much more interesting than any paperwork they have to do. You can also contact me to set an appointment if you need to be sure of seeing me outside of office hours. To summarize: don't be shy, and if you start to have troubles, come to me for help right away—not weeks later when it will probably be too late.
* A **Peer Instructor** will run Help Sessions several times per week, time and location TBD. Details will be provided in class.
* **Check your College email daily for course announcements.** Updates on quizzes and homework, as well as other announcements, are announce4d via email.

**Disclaimer**

I hope that the foregoing has given you a good idea of what the course will be like. It should not, however, be construed as a contract or legal document of any sort. In particular, the course content and policies mentioned herein are subject to reasonable modification in response to changing circumstances and events. I will, however, endeavor to notify you well in advance of any needed changes.