STATISTICS FOR THE BEHAVIORAL AND SOCIAL SCIENCES I

APSTA.GE.2001
Course Syllabus

<u>Professor</u>:
Sean P. Corcoran
246 Greene Street, Room 314E (IESP)
Phone: (212) 992-9468
Email: sean.corcoran@nyu.edu

Course description
This is the first of a two-semester sequence designed to introduce students to the basic concepts and tools of statistics, the practice of data analysis, and the use of statistical software (e.g., Stata or SPSS). Emphasis is on practical application and the acquisition of conceptual understanding rather than mathematical proof. Accordingly, the course is not appropriate for students seeking to learn the mathematical theory underlying the concepts covered. Specific topics are listed in the outline below.

Prerequisites
None, however a firm understanding of basic algebra and PC skills are expected. The class begins at a basic introductory level, but progresses relatively quickly.

Textbook
The following textbook is required:

> *Statistical Methods for the Social Sciences, 4th edition* by Alan Agresti and Barbara Finlay, Pearson Prentice-Hall, 2009.

You may also benefit from the following basic introduction to Stata (not required):

> *A Gentle Introduction to Stata, Revised 4th edition* by Alan C. Acock, Stata Press, 2014.

We will primarily use Stata in this course. For students who expect to use SPSS in the future, I recommend the following text, typically used in the 2nd course of this sequence at Steinhardt. You will not need it for this course, however.

> *Statistics Using SPSS: An Integrative Approach, 3nd edition* by Sharon Weinberg and Sarah Abramowitz, Cambridge University Press, 2015.

Course requirements
Your grade for this course will be determined as follows:

- *Nine* written problem sets (3% each for a total of 27% - all are weighted equally)
- A statistical project (23%)
- Midterm (25%) and final exam (25%)

The problem sets are designed to give you practice with the analytical tools introduced in class, and are critical to your success. They are due <u>before class</u> on the indicated date (see the attached class schedule). Except in rare circumstances, assignments will NOT be accepted after solutions are posted. You may drop your lowest score at the end of the semester (10 are assigned, 9 are counted).

Please submit your completed problem set as a <u>PDF</u> document via email (<u>sean.corcoran@nyu.edu</u>). Use your last name and problem set number as the filename (e.g., *Smith Problem Set 2.pdf*). Doing so will allow me to grade your assignment quickly and return it to you electronically.

Both the midterm and final exams will be given in class. You may use a calculator and one 8.5 x 11" sheet of notes (2-sided). Additional details on the statistical project are provided below.

Computer lab and software
Access to Stata and SPSS is possible through any of three methods: (1) the Virtual Computer Lab, (2) the (real) computer labs, and (3) purchase.

> NYU operates a service called the Virtual Computer Lab (VCL) which provides access to university-licensed software from anywhere with an NYU student login. You can access the VCL through <u>NYUHome</u> or: <u>https://vcl.nyu.edu/vpn/index.html</u>. Currently, Stata version 14.1 SE is accessible through the VCL.

> As a student you have access to campus computer labs with your ID. Lab attendants are not typically experts in Stata or SPSS, but they can answer system-level questions about opening files, saving, printing, etc. <u>NYU Data Services</u>, located on the 5th floor of Bobst Library, offers consulting to students who need assistance with statistical software. Contact them for more information, or to make an appointment. Data Services offers occasional tutorials on Stata, SPSS, and other software. Click <u>here</u> for a list of labs at NYU that provide Stata.

> You may be interested in buying Stata for your own computer. Stata version 14 can be purchased at a discounted student rate (the "<u>Campus GradPlan</u>"). "Small" Stata is the least expensive ($38 for six months or $54 for a year), but is limited in the size of datasets it can manage. (Note: Small Stata may no longer be available). For this class I recommend Stata/IC ($125 for a year, or $198 for a perpetual license). For most purposes, you will notice few differences between versions 11-14. However, be aware that minor differences do exist.

Other class information

1. <u>NYU Classes</u>: All materials pertaining to this course (lecture notes, readings, problem sets, data) will be made available via <u>NYU Classes</u>. Enrollment in the course should automatically give you access to the class site. Check in frequently for new materials and announcements. Lecture notes will generally be posted 1-2 days in advance of class. However, occasional (hopefully rare) delays are to be expected.

2. <u>Lab etiquette</u>: The class is held in a computer lab. To help promote a productive learning environment, please keep all other internet activities (e.g. email) to a bare minimum. Please do <u>not</u> use Facebook, instant messaging, or other such services while in the lab, and do not use class time to work on your problem sets (unless I formally give you class time).

3. <u>Absences</u>: Please see me immediately if you have any conflicts with scheduled assignments and/or exams, or if you anticipate being absent due to religious observances.

4. <u>Academic integrity</u>: NYU Steinhardt policies on academic integrity will be *strictly enforced* in this class. You can find the school's official statement on academic integrity here. You are encouraged to study and work together on problem sets, but all submitted work must be that of the individual student.

5. <u>Withdrawal</u>: If you wish to withdraw from the course, please do so formally with the University Registrar. If you withdraw without authorization, you are at risk for receiving a failing grade for the course. *September XX is the last day for graduate and undergraduate students to withdraw without receiving a "W" on their transcripts.*

6. <u>Accommodations</u>: Any student requiring an accommodation due to a chronic psychological, visual, mobility and/or learning disability, or who is Deaf or Hard of Hearing, should register with and consult with the Moses Center for Students with Disabilities at 212-998-4980, 726 Broadway, 2nd floor (http://www.nyu.edu/life/safety-health-wellness/students-with-disabilities.html). Of course, I am happy to provide any and all accommodations recommended by the Moses Center.

STATISTICAL PROJECT INSTRUCTIONS

Overview
The statistical project is an opportunity for you to apply your knowledge of statistics to real-world data, and answer a specific research question. You will analyze your data using Stata, interpret your results, and communicate your findings in writing. (If you have a strong preference for using SPSS, I will accept projects using that statistical package).

Requirements
For the statistical project, you may choose one of the following options: (1) provide a written report on a topic to be determined by me, or (2) provide a written analysis of a research question of your choice, using appropriate data. I will provide data for the former, and guidance on the latter.

Your written analysis should range between 12 – 15 pages, including tables and graphs. You may wish to include some graphs and tables in the main body of the paper, and relegate others to an appendix (appendices do not count toward the page limit).

Deadlines
You must choose a topic by *November XX*. The project itself will be due on or before *December XX*.

Guidelines
Think of your statistical report as communicating a story through data analysis. Like any good story, it should have a beginning, middle, and end, and a "plot" that is clear, logical, and compelling. Be selective about the tables, graphs, and statistics you choose to include. Overwhelming your reader with output and statistical results is a surefire way to lose (and bore) them. It also makes you look like an amateur. Be sure to appropriately reference all figures and tables in the text itself.

You may use any technique learned in class in your analysis, unless otherwise specified (relevant in option 1). One of the objectives of this project is a demonstration that you know which techniques to apply in specific situations. In many cases there is more than one technique that can be used. Choose the one that most clearly communicates the result you are trying to convey to the reader. Show some range in your use of methods, and avoid overt repetition.

If you choose option (1), you must fully address all of the questions I assign. These will be available later in the semester. If you choose option (2), you may craft your own research question as you see fit—but the quantity and rigor of the analysis should match or exceed that required under option (1).

A sample project and "model answer" will be posted on NYU Classes, as an example of past work. Additional details on the project will be provided later in the semester.

Submission
Please submit your completed project as a PDF document via email (sean.corcoran@nyu.edu). Use your last name and the words "Statistical Project" as the filename (e.g. *Smith Statistical Project.pdf*). There are many free Word-to-PDF converters on the internet, and many NYU labs have Adobe Acrobat Professional which creates PDFs. Apple computers generally allow you to save documents in PDF format, as do recent versions of Microsoft Word for Windows.

COURSE OUTLINE

Lecture 1: Introduction to concepts of probability and statistics (ch. 1-2)	
Lecture 2: Describing univariate distributions I (ch. 3)	*Problem Set 1 due*
Lecture 3: Describing univariate distributions II (ch. 3)	*Problem Set 2 due*
Lecture 4: Probability and sampling distributions I (ch. 4)	*Problem Set 3 due*
NO CLASS—NYU FALL BREAK	
Lecture 5: Probability and sampling distributions II (ch. 4)	*Problem Set 4 due*
Review for midterm exam	*Problem Set 5 due*
MIDTERM EXAM (covers chapters 1-5)	
Lecture 6: Statistical inference: estimation (ch. 5)	
Lecture 7: Statistical inference: significance tests (ch. 6)	*Problem Set 6 due* *Project topic due*
Lecture 8: Hypothesis testing: comparing groups (ch. 7)	*Problem Set 7 due*
Lecture 9: Bivariate relationships and correlation (ch. 9)	*Problem Set 8 due*
Lecture 10: Linear regression (ch. 9)	*Problem Set 9 due*
Lecture 11: Introduction to multivariate relationships (ch. 10-11, selected pages)	*Problem Set 10 due*
NOTE: NYU classes meet on Monday schedule Review for final exam	*Project due*
FINAL EXAM—6:00 – 7:50 p.m. *note special time* **(emphasis on Chapters 5-7 and 9-11)**	