

Course Objective and Related Courses: STA 6126 provides introduces statistical methods most commonly used in the social sciences. There is no prerequisite. The first half of the course introduces descriptive statistics and inferential methods (confidence intervals and significance tests) for means and proportions. The second half introduces bivariate methods for qualitative and quantitative variables (e.g., contingency table analysis, regression) and ends with an introduction to multiple regression. The follow-up course (STA 6127) presents more advanced regression methods (including multiple regression and its use for analysis of variance and covariance, and logistic regression for binary data). Other courses in the Statistics Department that are appropriate for students who have taken STA 6126 include STA 5503 (Categorical Data Methods), STA 5507 (Applied Nonparametric Methods), and STA 5701 (Applied Multivariate Methods).

Instructor: Alan Agresti, 204 Griffin-Floyd Hall, 392-1941, ext. 234, e-mail AA@STAT.UFL.EDU

Office Hours: Tues. and Thurs. 1:30-3:30 pm, or by appointment.

Teaching Assistant: Youngkyoung Min, 212 Griffin-Floyd, 392-1941, ext. 226, e-mail ymin2@stat.ufl.edu. Office hours Mon, Wed, and Fri 2-4 pm or by appointment (send e-mail), plus some weeks in a computing lab Mon 4-5 pm and Wed 5-6 pm.

Course web page: www.stat.ufl.edu/~aa/sta6126/

Course Policy:

Text: *Statistical Methods for the Social Sciences*, by A. Agresti and B. Finlay, Prentice-Hall (3rd ed., 1997).

Exam: Three in-class exams, each of which contributes 1/3 of the final grade. Make-ups will not be given unless arrangements have been made prior to the exam, and then only for illness or family emergencies.

Exam Dates:	Exam 1	September 30
	Exam 2	November 4
	Final Exam	Thurs, Dec. 18, 12:30-2:30

Homework: I recommend that you work the textbook problems listed on the next page to improve your understanding and help prepare for exams. Feel free to work with other students on these and on course study. A homework solutions manual for STA 6126 is available at University Copy Center (1620 W. University Avenue). Use this to check your solutions and help you with problems you find difficult. Please keep a neat file of your solutions to the homework problems. (These should be more detailed than those in the solutions manual!) Hand in this file when you take each exam. Each exam will contain 30 points (out of the 100 total) for homework credit; 15 points will be based on two data analyses for which you'll use software (see below), and the other 15 points will be based on the other exercises. For each of these parts, grades will be 15 - complete or nearly so, 7 - incomplete, 0 - not handed in or most solutions missing.

Software: When we cover computationally complex statistical methods, I'll show examples in class of the use of statistical software (mainly, SPSS) for the analyses. The homework exercises in boldface numbers require software. These refer to a data set at <http://web.stat.ufl.edu/~aa/social/data.html> based on a survey of STA 6126 students in 1996. Students are encouraged to become familiar with the use of a software package, which is used more extensively in follow-up courses such as STA 6127. Those who have not had previous exposure to any statistical software may want to work in a team with 1 or 2 other students on exercises that require software. SPSS and SAS are available at many computing labs on campus. A student version of SPSS 11.0 for Windows costs about \$80 at the campus bookstore (technology hub, first floor). A more advanced version that also handles all methods of STA 6127 (such as logistic regression) costs about \$200. One year student licenses are available of SAS and SPSS (see <http://www.software.ufl.edu/>).

Outline: The pages and problems listed in parentheses refer to optional extra material not covered in class or on exams but recommended for those who want to extend further their statistical knowledge.

<u>Topic</u>	<u>Text Pages</u>	<u>Problems</u>
1. Introduction to Statistics	1-9	1, 4, 6, 9
2. Sampling and Measurement	12-21 (21-28)	1-3, 5, 7, 11, 16, 18, 23 (20, 24, 27)
3. Basic Descriptive Statistics	35-67	1, 9, 11, 14, 17, 20, 26, 37, 38, 40, 42, 50 43, 49, 56, 59, 61 (62-64, 66, 68, 69)
4. Probability Distributions		
Normal distributions	80-94	1, 7-13, 18, 20, 22, 24
Sampling distributions	94-111	27, 31, 34-38, 46-50 (51, 54, 55, 57, 58)
5. Inference: Estimation		
Confidence interval for means	121-131	3, 5, 7, 8, 10, 36 , 45, 49-53
Confidence interval for proportions	131-135	13-15, 19
Sample size determination	135-141 (141-144)	26, 28, 29, 32 (46, 48, 54, 55, 56)
6. Inference: Testing Significance		
Mean: Hypotheses, test statistic, P-value	154-167	1, 3, 7, 9, 11, 51, 59-61
Proportion: Steps of test	167-173	12, 15, 16, 56
Decisions and types of errors	173-180	21, 22, 55, 57, 58, 64, 65, 67, 68
Small samples: <i>t</i> -test	180-187 (187-195)	23-25, 28, 29 (27, 44, 48, 70, 72)
7. Comparison of Two Groups		
Comparing means	210-216	1, 2, 44, 55, 35
Comparing proportions	216-220	9a, 16a-d, 54
Matched samples	226-229 (220-226, 229-233)	30, 31 (45, 48, 57, 58)
8. Association Between Categorical Variables		
Contingency tables, Chi-squared	248-264	1, 4-6, 8-10, 14, 33a
Summarizing association	265-268, 272-281 (268-272, 282-285)	11a-c, 15acd, 27, 28, 31, 39, 45, 47, 50 (3, 16, 22-24, 51-54, 57)
9. Linear Regression and Correlation		
Regression model, least squares fit	301-318	1, 3, 5, 6
Correlation and r^2	318-326	8, 9, (11, 17: SPSS users), 33-37, 40, 41
Inference, assumption, influence	326-341	10, 20, 22i , 42, 43 (44-46, 49-51, 54)
10. Introduction to Multivariate Relationships		
Association and causality, statistical control	356-365	2, 5-12, 32, 35, 36
Relationships, statistical interaction	369-372	23-24, 39, 40 (15, 17, 20-22, 25, 31, 33, 37, 38)
11. Multiple Regression		
Multiple regression model	382-394	1-3, 25i , 38, 39, 45
Multiple correlation and R^2	394-398	5a-e, 9a-c, 35
Inference, F distribution	398-404	5h-j, 7, 9de, 14 (17, 18, 47, 48, 51, 55)