

Statistics 586: Interpretation of Data I
Spring 2014

Textbook: Cabrera, J., McDougall, A. (2002). *Statistical Consulting*. New York, NY: Springer-Verlag.

Room: Hill 552 Wednesday 6:40 to 7:50 p.m.; 8:10 to 9:30 p.m.

Web Page: sakai.rutgers.edu

Instructor: Lynn A. Agre, MPH, PhD, RBS, 100 Rockefeller Road, Livingston Campus, Room 0526

Email: agre@rci.rutgers.edu; lagre@rutcor.rutgers.edu

Office Hours: By Appointment

Objectives: Learning goals of this course are how to create data sets, distinguish among different types of variables, decide which methods to apply given the constraints of the data, conduct analyses and report findings. Other key objectives include: computing using statistical/mathematical software, data analysis, preparation tables, graphs, and summarizing results suitable for submission in refereed publications. Oral communication and presentation of research hypotheses to peers is also a fundamental component of data interpretation.

Grading:

Project 1:	Data Set Creation 5 variables; Hypotheses; Analysis; Results; Discussion 5-7 pages; can work in groups of two. (20%)
Project 2:	Data Set Creation 7-8 variables; Hypotheses; Analysis; Results; Discussion 7-10 pages; can work in groups of two. (30%)
Project 3:	Data Set Creation 10 variables; Hypotheses; Analysis; Results; Discussion 10-15 pages; must be an individual project. (40%)
Attendance:	Participation in class and presentation of findings. (10%)

Note: Two Different Data Sets (but can use Data Set 1 and add variables given the unit of analysis is the same).

Reports: Format will follow general structure as delineated below:

- A. Introduction
 - Describe the research question or hypothesis addressing:
 - Who
 - What
 - Where
 - When
 - Why
 - How
 - Literature Review to support argument (i.e. using existing evidence on which to build hypothesis and now why testing)
- B. Methods
 - Explain variables .
 - Justify why chose that analysis.
 - What goal to be achieved
- C. Results
 - Tables
 - Graphs
- D. Discussion
 - Draw a conclusion; address limitations of study; implications for further research
- E. Literature Cited - References, i.e. Bibliography

<u>Date</u>	<u>Topic</u>	<u>Location</u>
January 22, 2014	Syllabus, Review Statistical Concepts - Chapter 1 - Statistical Consulting and Scientific Method - Introduction/Demonstration in R - Summary Statistics - Mean and Median; Histogram; Box Plot - Sample Presentation Data Analysis	HILL 552
January 29, 2014	Creating Data Sets Chapter 2- How to Describe Data Results; Exploratory Data Analysis - Univariate Statistics Summary Statistics in Graphical Display: - Stem and Leaf Displays - Shape of Distribution with Q-Q Plots; Histograms; - Standard Deviation - Writing up findings and discussion	HILL 552
February 5, 2014	Assignment Due: Data Set for Project No. 1 Created Chapter 3 - Hypothesis Testing and Use of Methodology to Answer Research Questions Using Measures of Dispersion to make inferences about data	HILL 552
February 12, 2014	Assignment Due: Abstract (Summary) for Project No. 1 Analysis in R Chapter 3 - Methodology - Continued - Box Plots - Power Transformations for variance stabilization - Spread vs. Level Plot	HILL 552
February 19, 2014	Assignment Due: Outline of Analysis for Project No. 1 Chapter 3 - Types of Transformation; When to Apply and Why	HILL 552
February 26, 2014	Project No. 1 - Due - Univariate Analysis w/Tables Chapter 7 - Bivariate Correlation; Cross-Tabulations - Simple Linear Regression; - Scatter Plots of relationship; - Least Squares vs. Line Fitting; Creation of Data Set for Project No. 2	HILL 552
March 5, 2014	Assignment Due: Data Set for Project No. 2 Created - ANOVA with two and three groups; - Fitting equations to data; - Power Transformations and non-linear fits; - Examining Residuals	HILL 552
March 12, 2014	Assignment Due: Abstract (Summary) for Project No. 2 - Contingency Tables - t- test -Two-way tables and Chi-Square	HILL 552
March 19, 2014	No Class - Spring Break	

March 26, 2014	Assignment Due: Outline of Analysis for Project No. 2	HILL 552
	Chapter 7	
	-Linear Models	
	-Generalized Linear Models	
	-Logistic Regression	
	-Random Effects Models	
April 2, 2014	Project No. 2 - Due - Bivariate Analysis	HILL 552
	Chapters 8 and 4	
	-Multivariate Analysis	
	-Data Visualization	
	-Principal Components Analysis - Factor Analysis	
	-Selection of Factors	
April 9, 2014	Assignment Due: Data Set for Project No. 3 Created	HILL 552
	-Rotations	
	-Discriminant Analysis and Classification Methods	
	-Cluster Analysis	
April 16, 2014	Assignment Due: Abstract (Summary) for Project No. 3	HILL 552
	- Data Mining	
	- Using Multivariate Analysis for Variable Reduction	
	- Segmentation and Subsetting of Large Data Sets	
	- Extracting Information from Large Data Sets	
	- Recursive Partitioning and Trees	
April 23, 2014	Assignment Due: Outline of Analysis for Project No. 3	
	Student Presentations - Final Project - 15 minutes each	HILL 552
	Start Time: 6:30 p.m.	
April 30, 2014	Student Presentations - Final Project - 15 minutes each	HILL 552
	Start Time 6:30 p.m.	
May 6-7, 2014	Reading Days	
May 14, 2014	Final Project No. 3 - Paper - Due - Univariate; Bivariate; Multivariate	

Selected Reference List
(Chapters and Articles Will Be Posted on Sakai, with Additions)

- Abramson, J.H. (1994). *Making Sense of Data: A Self-Instruction Manual on the Interpretation of Epidemiologic Data*. New York, NY: Oxford University Press.
- Allen. M.P. (1997). *Understanding Regression Analysis*. New York: Plenum Press.
- Bland, J. M. (1996). Statistics notes: Transformations, means and confidence intervals. *British Medical Journal*. 312:1079.
- DeMuth, J.E. (2006). *Basic Statistics and Pharmaceutical Statistical Applications*. 2nd Ed. Chapman and Hall, CRC Press.
- duPrel, J-B. Hommel, Gerhard, Rohring, Bernd, Blettner, M. (2009). Confidence interval or p-value? *Deutsches Arzteblatt*. 106 (19): 335-339.
- Heeringa, S.G., West, B.T., Berglund, P.A. (2010). *Applied Survey Data Analysis*. Boca Raton, FL: CRC Press Taylor and Francis Group.
- Horton, N. J., K. Kleinman. (2011). *Using R for Data Management, Statistical Analysis and Graphics*. Boca Raton, Florida: CRC Press, Taylor and Francis Group.
- Jacoby, W.G. (1997). *Statistical Graphics for Univariate and Bivariate Data*. Series: Quantitative Applications in the Social Sciences, v. 117. Thousand Oaks, CA: Sage Publications.
- Kachigan, S. K. (1991). *Multivariate Statistical Analysis*. New York: Radius Press.
- Maindonald J., Braun, W.J. (2007). *Data Analysis and Graphics Using R*. Cambridge, UK: Cambridge University Press.
- Miller, J.E. (2005). *The Chicago Guide to Writing about Multivariate Analysis*. Chicago, IL: University of Chicago Press.
- Motulsky, H. (1995). *Intuitive Biostatistics*. New York, NY: Oxford University Press.
- Pearl, J. (2000). *Causality. Models, Reasoning and Inference*. New York, NY: Cambridge University Press.
- Tabachnik, B.G., Fidel, L.S. (1996). *Using Multivariate Statistics*. New York, NY: Harper Collins.