

**Syllabus and Course Description for
SURV 616 and SURVMETH 686
Statistical Methods II
Winter 2012**

Instructor: James Wagner
Meeting Time and Places: Friday 12:30 – 3:00pm
1208 Lefrak Hall and 300 Perry

CTools Chat: Tuesdays 12:00-1:00pm

Office Hours: By appointment, 4040 ISR

Office Phone: 734-647-5600

Email: jameswag@isr.umich.edu

Class Website: <https://ctools.umich.edu/portal>
All the course-related material will be posted on the website: JPSM students need to create U Michigan friend's accounts to access the ctools website. If you have not received email about this, please contact Jill Esau (jesau@isr.umich.edu) ASAP.

Video site: <http://www.jpsmclasses.umd.edu>
Class sessions can be taped upon request. Please let me know if you need to miss a class and I will request that the session be taped. Check the following website for information on viewing the recordings:
<http://www.jpsm.umd.edu/support/View%20Recordings.pdf>.

Instructions for remote access to JPSM network for running SAS at home:

www.jpsm.umd.edu/support.htm

Text (required): Heiberger, R. M. and B. Holland (2004). *Statistical analysis and data display : an intermediate course with examples in S-plus, R, and SAS*. New York, Springer Science+Business Media.

Class Notes. Steve Miller and Richard Valliant.

Singer, J. D. (1998). "Using SAS PROC MIXED to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models." *Journal of Educational and Behavioral Statistics* 23(4): 323-355.

Singer, J. D. and J. B. Willett (1993). "It's About Time: Using Discrete-Time Survival Analysis to Study Duration and the Timing of Events." *Journal of Educational and Behavioral Statistics* 18(2): 155-195.

Other references (optional): *An Introduction to Categorical Data Analysis* (2007)
by Alan Agresti, John Wiley & Sons.

Applied Multivariate Statistical Analysis (2002)
by Richard A. Johnson and Dean W. Wichern, Prentice Hall

Applied Linear Regression, (2005), 3rd edition (Weisberg).
Statistical Methods, 8th ed (Snedecor and Cochran)

Grading:	Homework	55%
	Project	30%
	Final Exam	15%

Each assignment will be will receive a numerical score, which corresponds to the following letter grades:

A+	[98,100]
A	[93,98)
A-	[90,93)
B+	[87,90)
B	[83,87)
B-	[80,83)
C+	[77,80)
C	[73,77)
C-	[70,73)
D+	[67,70)
D	[63,67)
D-	[60,63)
F	< 60

Homework and projects that are turned in late will be subject to the following policy:

- Assignments turned in late but within 1 week of the due date: 50 percentage points off. Within 1 week of the due date means the assignment is turned-in at or before the start of the first class after the due date.
- Assignments turned in more than 1 week late will not be accepted.
- Late Project Abstracts and Projects will not be accepted without the prior permission of the instructor.

The final course grade is determined by assigning the letter grade corresponding to the result of the weighted average of the numerical scores for each of the assignments (50% for the Homework, 35% for the Projects, and 15% for the Final Exam).

OVERVIEW

The purpose of the class is to build on the concepts discussed in Statistical Methods I by introducing the student to five major groups of statistical methods :

- I. Multivariate Analysis*
- II. Categorical Data Analysis*
- III. Statistical Methods in Epidemiology.*
- IV. Time Series Analysis*
- V. Analysis of Panel Data*

The emphasis will be to understand and apply the methods presented, and develop a feel for how problems in data analysis can be viewed from several different perspectives. In all cases the emphasis will be on understanding the techniques, rather than deriving their theoretical properties. The student will be expected to apply the techniques on weekly homework problems, projects, and a final exam.

OUTLINE

Note: All readings associated with a given class should be completed before the class meets, except for the first class.

Class	Date	Topic	Reading	Homework Due
1	January 6	Random Effects	Class Notes #1	
2	January 13	General Linear Model	Class Notes #2, Singer (1998)	HW #1
3	January 20	Multivariate Analysis	Class Notes #3; Optional: Wichern and Johnson, Chps. 2, 4, 5.1-5.3	HW #2
4	January 27	Categorical Data Analysis	Class Notes #9; Optional: Agresti Chps. 1-2;	HW #3
5	February 3	Categorical Data Analysis	HH Chp. 15; Class Notes #10; Optional: Agresti, Chp. 7	HW #4
6	February 10	Epidemiology	Class Notes #13	Project Abstract
7	February 17	Categorical Data Analysis	Class Notes #11	HW #5
8	February 24	Categorical Data Analysis	HH Chp. 17.1-17.4 Class Notes #12, Optional: Agresti, Chps. 4-5;	HW#6
Spring Break – No class				
9	March 9	Categorical Data Analysis	HH Chp 17.5-17.10	
10	March 16	Epidemiology	Class Notes #14; Singer and Willett (1993)	HW #7
11	March 23	Time Series	HH Chp 18.1-18.6; Class Notes #4;	HW #8
12	March 30	Time Series	HH Chp. 18.7-18.8; Class Notes #5, #6	
13	April 6	Panel Data	Class Notes #7	Project
14	April 13	Panel Data	Class Notes #8	
	April 20	Final Exam		

PROJECT

Projects are designed to allow each student some individual creativity in developing and expressing their own approach to data analysis, within the framework of the stated goals of SURV 616/SURVMETH 686.

Data

You can use datasets from your work or another source. The main requirement is that the statistical techniques covered in class can be applied to analyze the data.

The web site for the Inter-University Consortium for Political and Social Research, <http://www.icpsr.umich.edu/>, has many downloadable datasets that can be used for analysis. Some of the datasets are collected using sample surveys that include strata, clusters, and weights. Students in SURV 616/SURVMETH 665 are NOT expected to use the specialized techniques, taught in other courses, that have been developed for analyzing survey data. You can treat each dataset as if it were a simple random sample. However, keep in mind that if a dataset from a complex sample were analyzed appropriately, the relationships found could be different from those that you uncover.

Types of Analyses

The analyses you conduct should be ones that we have covered in the class. These include logistic regression modeling, time series, and the analysis of panel data. Although fitting of statistical models is expected, do not overlook the need for simple descriptive analyses to better understand the variables in your data. Your project goals should be to:

- (a) Demonstrate that you know how to apply the methods covered in class in a systematic and thorough way to investigate relationships in a dataset or to explore a particular problem, and
- (b) Present your results coherently in a written report.

Reports should include tables and figures that illustrate the points you want to make and the analyses you have done.

Format of the Abstract

- Loosely follow Structured Abstract Format
(Refer to http://research.mlanet.org/structured_abstract.html)
- 300 words limit
- Include the following three sections:
 - o Objectives: State the essential issue the project attempts to address.
 - o Data: Introduce the data source.
 - o Analysis: Provide an analysis plan.

Format of the Report

The project reports take the form of a Lab Report (as discussed in *Stat Labs: Mathematical Statistics Through Applications* by Deborah Nolan and Terry Speed). Appendix A of *Stat Labs* (by permission of the authors) will be posted on the class website, as a guide to writing project reports. Reports should be 10-15 pages, counting tables and figures but excluding the abstract and appendix. Use the following format:

- MS-Word
- Space and one-half (do not use single spacing)
- 1 inch margins
- Include page numbers
- Parts of report

- **Title page** with title of project, name, class, date
- **Abstract**—300 words limit; Closely follow Structured Abstract Format on http://research.mlanet.org/structured_abstract.html
- **Introduction**—state the problem you are addressing, the goals of your project, and your findings
- **Methodology**—describe your data, how they were collected, and the variables that you will be analyzing
- **Results**—present your findings along with a description of the techniques used for analysis. You can refer to particular sections of the class notes or text as part of your explanations. Graphs of different kinds are often a good way to display results (e.g., histograms, scatterplot matrices). There should be no more than 5 tables and figures (e.g. 2 tables and 3 figures=5) total.
- **Discussion**—what did you learn? Describe any limitations to your findings and follow-up work that might be appropriate.
- **Technical appendix**—this part should include SAS code and the parts of the output that are relevant to the discussion in the main body of the report. The appendix is not counted in the suggested length of 10-15 pages but should not itself exceed 15 pages. If, in the main body of the report, you refer to results in the Appendix, you should refer to specific pages, figure numbers, or tables of the Appendix for clarity.

Project Grading

The project will account for 30% of the grade in the class.

Grades will be based on complexity of the project, thoroughness of analysis, correctness of results, and the quality of the written report. Extremely simple projects will receive poorer marks than more complex ones. Points will be deducted if reports are poorly written or poorly organized.

GROUND RULES

Assignments:

- a. Homework should be completed in either R or SAS. Submit your code and output.
- b. Submit a PDF file. This can be prepared using one of the following software packages (but not done by hand):
 - a. MS-Word format. Use MathType or other math symbol typing software for mathematical expressions.
 - b. Latex
- c. Submit via the CTools website.
- d. Homework assignments are due **BEFORE** the following class. Class begins at 12:30 on Friday. Homework should be submitted before 12:30. Late homework assignments will be accepted until one week after they are assigned but will receive 90% of the points as outlined above.
- e. Late project abstracts and reports will not be accepted unless permitted by the instructor 3 weeks prior to the due date.
- f. Graded assignments will be returned via the ctools website a week after the due date.
- g. Naming convention for assignment files: ASSIGNMENT_LASTNAME_FIRSTNAME.xxx
 - o e.g., hw1_wagner_james.doc; abstract_wagner_james.doc
- h. Students are encouraged to work together on assignments but in a productive way.
 - o Limit what is shared.
 - o Copying other's work will **NOT** benefit your education.

Class etiquette:

- a. If you are at the remote site, stay within the camera angle.
- b. No ringing phones.
- c. Refrain from using personal computing/communication devices during lectures.
 - o This includes laptops, netbooks, iPod touch, iPads, smart phones and cellphones.
 - o Of course, feel free to use them during breaks.