Wharton
University of Pennsylvania

DEPARTMENT OF STATISTICS

STAT 621 Fall 2015

Accelerated Regression Analysis for Business
Syllabus

Instructor:

Richard Waterman waterman@wharton.upenn.edu 443 JMHH

Source material

Required

- Class Notes. These can be downloaded directly from the Stat 621 Canvas e-room.
- JMP 12 (software), SAS Institute, downloadable from upenn.onthehub.com
- Stine and Foster, *Statistics for Business*, Second Edition, Pearson.

Optional (on reserve at Lippincott Library)

- Sall, Creighton, Lehman, *JMP Start Statistics*, 5th Edition, SAS Institute.
- Freedman, Pisani and Purves, *Statistics*, 4th edition, Norton.
- Keller, *Statistics for Management and Economics*, 8th edition, South-Western Cengage Learning.

The fundamental material for the class is contained in the Class Notes, which will be discussed and elaborated in the class lectures. The Stine and Foster (SF) textbook elaborates on most (but not all) of the Class Notes. Links to the relevant readings in SF appear throughout the Class Notes. For those who would like further background materials, we recommend Sall, Creighton and Lehman (SHL), Freedman, Pisani and Purves (FPP) and Keller (K). SHL is an example-rich guide to statistical analysis with the statistics package JMP. FPP is a highly verbal and conceptual book - an excellent introduction both for "poets" who are unfamiliar with technical readings and for "quants" who would like a better sense of the reasoning behind statistics. K is in the style of a traditional "reference manual" and explains details and provides many formulas for statistical procedures that are not covered in class.

JMP is the computer package we'll use extensively for statistical calculations and graphics. In particular, an essential component of 621 will be project work requiring substantial use of JMP. Although JMP is merely a tool and not the central point of the course, it is sufficiently useful that you will need it.

Class Preparation

As soon as possible, you should obtain and install JMP. Before each class, you should review the material from the previous class and you should skim the Class Notes that will be covered. This is a course that builds upon itself and it is crucial to not fall behind. The classes focus on critical interpretation of results and analysis of assumptions. We use JMP to carry out the computations, although the software itself is not the main focus of the course.

You should also read the relevant sections of the SF textbook as annotated throughout the Notes and shown in this syllabus. We strongly recommend that you review the exercises that conclude each chapter. The exercises in each chapter of the SF textbook begin with matching, true/false, and conceptual questions. You should routinely skim these exercises in every chapter; they review notation and basic properties of the methods covered in class. In addition, the course outline identifies specific additional "you do it" exercises that require data analysis or computation that is related to examples and topics of lectures. These exercises will not be evaluated but are useful for review.

Course Overview

The course assumes that you are familiar with the material covered in the first half of Stat 613 which is basic probability and inferential statistics. With this as a foundation, the course critically explores the use of the key statistical methodology known as regression analysis for solving business problems. These methods and their application will reappear in many other MBA classes and are part of the basic "tool kit" expected of all MBAs in their careers.

Lecture Date	Key Topics	Reading (SF)	Exercises
1 Aug 26	*Course overview* *Review of the key ideas from inferential statistics to be used regression*	Chapters 1-18.	
2 Aug 31	*Fitting lines to data* Slope and intercept, fitted values and residuals, r-squared	19	19.39, 41, 43, 47
3 Sep 2	*Fitting curves to data* Transformations (logarithm, reciprocal), elasticity	20	20.31, 33, 35
4 Sep 9	*Simple regression model* Parameters, assumptions, basic diagnostics	21.1-2	
5 Sep 14	*Inference for the Simple Regression Model* Tests, confidence intervals, prediction intervals	21.3-4	21.39,41,43,47
6 Sep 16	*Remedies for common problems* Nonlinearity, dependence, heteroscedasticity, outliers	22	22.37,39,45 4M (p572)
7 Sep 21	*Multiple regression* Scatterplot matrix, marginal and partial slope, path diagram	23.1-2	
8 Sep 23	*Multiple regression model* R^2, *F*-statistic, model profile, diagnostic plots	23.3-5	23.39, 41, 43, 47
9 Sep 28	*Collinearity in multiple regression* Variance inflation factor	24	24.33, 35, 37, 41
10 Sep 30	*Using categorical variables in regression* Dummy variable, partial *F*-test, model profile	25.1-4	25.39, 41, 43, 47
11 Oct 5	*More categorical predictors*	25.5	
12 Oct 7	*Review: building a regression model* Stepwise regression, data mining	SIA p736	
Oct 14	***Final Exam 6-8pm***		

Attendance

Attendance is an important aspect of the Wharton commitment. Wharton students are admitted in part because of the experiences they bring to the community that they can add to class discussions. Without attending, learning as a collaborative process cannot exist. Accordingly, absences are only appropriate in cases of personal emergency. In addition, late arrival is disruptive to the learning environment and promptness is expected. Please make note of the start of the term and the time of deliverables and exams as you make travel plans. In case of illness, we require a letter of confirmation from Student Health Services. If you find yourself in a conflict due to your career search or recruiting activity, you should work with the MBA Career Management Office to find a resolution.

Absences due to recruiting are not excused. Employers are prohibited from requiring recruiting-related activities (e.g., interviews, events or travel) that conflict with a student's academic commitments. An employer's inflexibility on this issue is a violation of Wharton's recruiting policies.

Exercises, Quizzes and Exam

There will be weekly exercises as indicated in the course syllabus. These exercises will not be collected, but they are essential for the learning process and you should treat them as a requirement. The textbook supplies brief answers to these questions and office hours are available for further questions.

There will be four 10 minute in-class quizzes. Quizzes will take place on Mondays (Sep. 14, 21, 28 and Oct 5).

There will be a two-hour final exam.

Learning Team Project

A project will be assigned to each learning team during the course. It will entail the statistical analysis of data for a business application that your team will describe in two installments. Installment 1 is due in on Sep 21. Installment 2 is due in on Oct 12.

This project must reflect the work of only your learning team. You are strictly forbidden from discussing this project with anyone outside your learning team.

Teaching Assistants (TAs)

TAs for Stat 621 will hold office hours throughout the course. Times and locations will be posted in the 621 Canvas e-room.

Classroom Expectations - Concert Rules

- Class starts and ends on time.
- Sit according to the seating chart (posted on-line in Canvas).
- Late entry or reentry only under exceptional circumstances.
- Name tents displayed.
- Phones, laptops and other electronic devices turned off. We do permit the use of tablets (e.g., an iPad) for taking notes in class.

Grading

Grades for the course will be based on the following components

Final Examination	50%
In-class Quizzes (4)	20% (5% each)
Project	25% (10% and 15%)
Concert rules, including attendance	5%

Attendance is mandatory. One unexcused absence is allowed during the quarter without penalty; beyond that, each unexcused absence removes a ½ percentage point from your total grade.