Wharton
UNIVERSITY *of* PENNSYLVANIA

DEPARTMENT OF STATISTICS

STAT 621   Fall 2020

Accelerated
Regression Analysis for Business
Syllabus

Instructor**:**   Richard Waterman   waterman@wharton.upenn.edu

   TA: Mauricio Daros   daros@wharton.upenn.edu

Class meeting times (Monday/Wednesday) all on Zoom with links available through Canvas.

> Section 1: 9:00am – 10:20am.
> Section 2: 10:30am – 11:50am
> Section 3: 1:30pm – 2:50pm.

Waterman's office hours will also be through Zoom at this url:
https://upenn.zoom.us/j/5136434021

Mauricio's office hours will also be through Zoom at this url:
https://upenn.zoom.us/j/9467591304

Source material

Required
- Class Notes. These can be downloaded from the class Canvas website. It is recommended that you use these to annotate the lectures during class.
- JMP 15 (software), SAS Institute, free download from the class Canvas website.
- Stine and Foster, *Statistics for Business*, Third Edition, Pearson.[1] References throughout are to the third edition, though the second edition is very similar.

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*, 10th 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 will use extensively for statistical calculations and graphics. 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 very useful and is used in other classes here at Wharton.

---

[1] Various purchase options are available at
http://www.mypearsonstore.com/bookstore/statistics-for-business-decision-making-and-analysis-9780134763736.
The pure eBook edition us fine too: https://www.pearson.com/store/p/statistics-for-business-decision-making-and-analysis/P100002245171/9780134506715?creative=460795233349&keyword=&matchtype=&network=g&device=c&gclid=CjwKCAjw4rf6BRAvEiwAn2Q76lfm8_u7tGyJByxoa58iPniEN3Q5Pi-iJmnx6xUrPYNpQSvohfPlMBoCpPIQAvD_BwE

Course Overview

The Class Notes are organized into 12 modules.

Module	Title	Readings (SF)	Data Analysis Exercises
1	*Getting Started*	Review Background Notes Stat 613 [2]	
2	*Fitting Linear Equations to Data*	19	19: 39, 41, 43, 47
3	*Fitting Nonlinear Equations to Data*	20 (skip 20.3)	20: 33, 35, 37
4	*The Simple Regression Model*	21.1-2	21; 23,25,27
5	*Detecting and Dealing with SRM Violations*	22	22: 37, 39, 45, 49
6	*Inference with the Simple Regression Model*	21.3-4	21: 39, 41, 43, 47, 49
7	*Multiple Regression*	23.1-2	23: 25, 32
8	*The Multiple Regression Model*	20.3, 23.3-5	23: 39, 41, 43, 47
9	*Collinearity in Multiple Regression*	24	24: 33, 35, 37, 41
10	*Categorical Explanatory Variables*	25.1-4	25: 39, 41, 43, 47
11	*Comparing Several Groups*	25.5	25: 46
12	*Building Regression Models*	SIA p 815-820	

The course assumes that you *know* the material covered in the first half of Stat 613, namely basic descriptive and inferential statistics.[3] With this material 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.

Class Preparation, Review and Exercises

As soon as possible, you should download and install JMP. Before each class, you should review the material from the previous class and 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 listed above. 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

---

[2] This folder on the Canvas website contains the first half of the Stat 613 lectures.
[3] Without this background, you are strongly advised to enroll in Stat 613 rather than Stat 621.

exercises in every chapter; they review notation and basic properties of the methods covered in class. In addition, the course outline above identifies additional "you do it" exercises that require data analysis or computation related to the examples and topics in the lecture notes. These exercises *will not be collected*, but they are very helpful for the learning process. The textbook supplies brief answers to the odd numbered questions and office hours are available for further questions.

Quizzes and Final Exam

- There will be 5 short 8-minute in class quizzes (drop the lowest score).
- There will be a two-hour final exam on October 15.

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 report on in two installments. Installment 1 is due in by midnight on Sep 27. Installment 2 is due on Oct 18.

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.

Late homeworks will be penalized by 25% for the first 24 hours and the 50% for up to 48 hours. More than 2 days late and they score 0.

Office Hours

Richard waterman (Instructor):   Monday/Wednesday 3-4:00 pm online at https://upenn.zoom.us/j/5136434021.

Mauricio Daros (TA):   Wednesday/Friday 4-5:00pm online at

https://upenn.zoom.us/j/9467591304

Grading

Grades for the course will be based on the following components

Final Examination	55%
In-class Quizzes	20%
Project Installments (2)	25% (9% and 16%)

Grade queries must be received within 1 week of solutions being posted and sent to Mauricio.

Class meeting dates and quiz schedule

1.   9/2/2020
2.   9/9/2020
3.   9/14/2020 (Quiz 1)
4.   9/16/2020
5.   9/21/2020 (Quiz 2)
6.   9/23/2020
7.   9/28/2020
8.   9/30/2020 (Quiz 3)
9.   10/5/2020
10.   10/7/2020 (Quiz 4)
11.   10/12/2020
12.   10/14/2020 (Quiz 5)