DEPARTMENT OF STATISTICS AND DATA SCIENCE

STAT 6130 Fall 2025

Regression Analysis for Business
Syllabus

Professor Abraham (Adi) Wyner ajw@wharton.upenn.edu
303 Academic Research Building
Office Hours: Mondays 4:15-5:15 pm (when class meets)
215 898-2439 (you can leave a voice message here)

Class Room and Sections

004 A. Wyner MW 8:30- 10:00 G60 JMHH
005 A. Wyner MW 10:15- 11:45 G60 JMHH
006 A. Wyner MW 1:45- 3:15 G60 JMHH

Teaching Assistants:

- **Head Teaching Assistants:** Joey Rudoler (jruloder@wharton.upenn.edu)
- **MBA Teaching Assistants:** TBD

Your usual, frequent and main tool to contact the staff about anything related to the course is the Ed Discussion tool on canvas. There are tabs related to various subjects; we encourage you to make your questions public but if private questions are needed you can make your post private.

Only use email if you have a question that requires a specific member of the staff.

Source material

Required

- JMP Student Edition (Accessible for free through canvas)
- Pekoz, Erol *Managers Guide to Statistics* available at the bookstore https://upenn.bncollege.com/my-account/rental-order/orders
- Class Notes (canvas)

Additional Reading Suggestions

- 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, 2014, South-Western Cengage Learning.
- Stine and Foster, *Statistics for Business*, Addison Wesley.

Course Pedagogy

The traditional way of teaching a statistics class is lecture. The professor presents formulas, justifies the formulas and then perhaps illustrates their use by solving problems in class. Students ask questions about confusing points.

Although there will be plenty of lecture (some asynchronous), actual class time will also feature in-class activities, student problem-solving, and discussions.

For the course to work as well as possible, it is important that you **prepare for class.** Before every class there is a reading and self-study problems (not submitted). It is highly recommended that you prepare for class by do the reading first and then try to solve the problems. Solutions to all the self-study problems are posted to Canvas. If you properly prepare the lecture will be super clear and easy to follow. Towards the end of class there will be breakout problems that will be solved in groups during class and submitted individually. These will be graded automatically and used as part of you participation score.

The class is organized into modules on canvas where all the assignments posted.

Background

The statistics background for students in STAT 6130 is quite variable. Some of you have never had any courses and others have had many. *I will assume this is your first course in statistics*.

The Wharton MBA program has a strong analytical focus, and the Statistics 6130 class has always reflected that orientation. Sometimes the urgency to teach skills obscures the need for conceptual understanding. To that end, I am assigning the book "**Manager's Guide to Statistics"** by Boston University Statistics Professor Erol Pekoz. The book is highly conceptual; it is in many ways a managerial version of class text "Statistics" by Freedman, Pisani and Purves). If you are new to statistics (or it has been years and years) the reading will be an essential part of the learning process.

JMP (pronounced "jump") is the computer package we'll use extensively for statistical calculations and graphics. An essential component of 6130 will be homework that use JMP.

Why use JMP?

There are many statistical packages, including SAS, Minitab and even Excel (if you are generous). JMP has a growing user base, but it is not likely to be the tool of choice when you return to the work force. But JMP has many advantages. It is **extremely** powerful. Indeed, its most recent versions are equipped with an unsurpassed suite of artificial intelligence tools including natural language processing, machine learning and internet data acquisition. It can be used entirely through a "point and click" interface which is super easy and highly conducive to exploration. **Thus, you can concentrate on understanding.** You will also be doing very powerful analyses very quickly. It is also used in many upper-level stat classes at Wharton and UPenn.

Can I learn R?

The book has optional material that teaches you to use the free statistical programming language R. If majoring in Statistics is something you are seriously thinking about, then learning to use R will be important. Use of R is optional in this course, but it will be supported by the head TAs. You may install and experiment with the free R software package and the corresponding R studio editor using the two links here:

- R: https://cran.rstudio.com/
- R studio: https://www.rstudio.com/products/rstudio/download/

Can I use Chat GPT 4 with Code Interpreter?

Data analysis is well suited for AI. You can use it do your analysis but BE CAREFUL. It can do strange things. If you use AI and your analysis, looks good, but is all wrong, you will not do well.

You cannot use Chat GPT for any written language work. If you want to do any data analysis you can try but caveat emptor. You can submit graphs, tables, equations, etc… made by ChatGPT; JUST NOT WORDS. These you must write on your own.

Class

Classes will begin with a lecture and often a problem or two sometimes based on the reading. Then we will "flip" the class after about 50-60 mi and break into groups to complete a worksheet of problems. We will then regroup and time permitting a group will present their solution to the class. You have been assigned seats according to your

learning teams, but you may work with any students you wish on your breakout problems. When you submit, include the names of the students in your group on that day,

Attendance and Videos

Use the attendance app to record your participation. There is no check-in when there is no-class and not on the dates of the in-class exam.

Lecture Date	Key Topics
1 August 25	*How not to be fooled with data!* Causality, confounding and Controlled Studies. The method of comparison.
2 August 27	*The Histogram* Portraying variation in the data. Summary measures of the center and spread of the Data (median, mean and IQR). SD and Empirical Rule
3 Sep 3	*More Data Visualization:* Box plots, mosaic plots, *More Data Summary Examples*
4 Sep 08 QUIZ 1	*Normal Distribution:* A very useful tool for understanding data and making predictions.
5 Sep 10	*Telling Statistical Stories with Graphs:* The graph-builder in JMP.
6 Sep 15	*Correlation*
7 Sept 17 QUIZ 2	*More Correlation*
8 Sep 22	*Regression Equation, Regression Residuals*
9 Sept 24	*Multiple Regression Continued* Fitting hyperplanes to data. Interpreting Regression Equations. Marginal and Partial Slope. **ROSH HASHANNA: Class will be virtual.**
10 Sep 29	*Multiple Regression Continued* Dummy Variables, Interactions
11 Oct 1	*Multiple Regression Examples. Transformations*
12 October 6	Exam I: In class.

13 Oct 20	*Probability: a "language" to reflect uncertainty* *Conditional Probability and Bayes Rule*
14 Oct 22	*Random Variables and their Expected Value and SD*
15 Oct 27	*Sampling variation:* The law of averages and the standard error of the sample percentage.
16 Oct 29 QUIZ 3	*Central Limit Theorem of Statistics*
17 Nov 3	*Confidence Intervals*
18 Nov 5	*Hypothesis Testing* One Sample tests, the p-value.
19 Nov10	*Hypothesis Testing:* Two Sample tests: The "A/B" test
20 Nov 12 QUIZ 4	*Confidence Intervals in Single Variable Regression* The p-value of a regression.
21 Nov 17	*Confidence Intervals in Multiple Regression*
22 Nov 19	B*uilding a regression model* Stepwise regression, data mining.
Nov 24	Optional Virtual Class: Problem Session.
23 Dec 1 **Quiz 5** (optional)	*How to diagnose problems in regression and what to do about them.* *Project guidance.* *Wrap-Up !!*
24 Dec 03	*Exam II*
Dec 16	***Final Project Due at 11:59 PM.***

Homework, Quizzes and Exams

- There will be 6 homework assignments and one final exam-data-analysis project. (5-10 hours of work). Brevity is encouraged! The final data analysis project is individual. Assignments **3 and 5 are individual.** Assignments 1,2,4, and 6 are with your teams. Due dates are September 9th , 18st , **October 2th** and 28th, **November 11th** and December 1nd.
- There will be 4 short quizzes. These will be multiple choice. There will be a fifth quiz that is a make-up for anyone who missed a quiz earlier in the semester.
- There will be two 1.5 hour exams in class (October 6th and December 3th)
- One week grade query maximum from the time work has been handed back (go to the TA).

Many problems will be from the book. All datasets required are posted to canvas. Some of the homework assignments must be done individually and submitted online by the due date and time. You may discuss the problems with other students in the class, but **your final write-up should be done individually or team** and then uploaded to the "Assignments" section of the course website.

Group assignments are submitted once and there is one grade for the entire group. T*he content of your homework is to be generated by you (or your team) and you alone. If you copy work from another team or from another source you not only will get severely penalized, you (and your team) will be referred to the Office of Student Conduct for review.*

Pre-Class Work: Self-Study and Reading

Before most classes, there is a reading assignment from the text and "self-study problems" many based on the readings and lectures. These will not be graded, and they will be simple and they do not need to be submitted.

Breakout Problems

In every lecture, there will be breakout problems. In class you can work in a group. You are required to submit the class's breakout problem quiz **individually** online. Submissions are due on the day they are assigned by 5:00 PM. You have been assigned seats according to your learning teams, but you may work with any students you wish on your breakout problems. You do not have to get everything correct and although your submission will be graded (by computer) the grade is just feedback and doesn't count to your final grade.

Teaching Assistants (TAs)

TAs for Stat 6130 will hold office hours throughout the course. Times and locations will be posted on canvas. You can refer questions to any of your TAs. Questions about regrades should be taken to your PhD student TAs.

Class Participation Information

I will construe class participation broadly. There will be plenty of opportunity to participate in class, you don't need to do them all equally well.

1. Answer questions and ask questions during lectures.
2. Submit breakout problems
3. Attend class consistently (use the attendance app to record your presence).

Grading

1. The tests and the quizzes are in-class, taken on a laptop using the Lockdown browser. They are closed book, closed notes except that for each exam (not quizzes!) you may use both sides of one sheet of handwritten notes.
2. Tests last one hour and thirty minutes
3. Quizzes are 10-12 minutes in length. There will be a “fifth” quiz that is optional and can be used to replace the lowest quiz or used as a make-up, if an in-class quiz is missed.
4. Late homework is penalized by 25% for up to 24 hours late and 50% for up to 48 hours late. At that point solutions will be posted and no further submissions will be accepted. Any homework grade queries must be made via Ed Discussion within one-week of the solutions being posted.
5. There are no extra credit opportunities in the course.

Grades for the course will be based on the following components

- Test 1: 20%
- Test 2: 25%
- Quizzes: 15%
- Homework: 10%
- Take-home final exam data-analysis project: 20%.
- Attendance and class participation (including breakout problems completeness): 10%.

Makeup

There is no reason to miss a test or a quiz as they are scheduled in class, and students should not schedule anything that conflicts with class time, especially on quiz and exam dates. Please plan accordingly. Exceptions will only be made for true unforeseen emergencies and major life events like illness, marriages, deaths etc. The pre-arranged make-up date for quizzes is December 1nd.

Miscellaneous Notes

Failure to complete the final project results in an automatic failure of the course.

Individual homework must be your own work. Students who submit work that is not solely their own effort or the effort of their learning team, will automatically fail the class.

If you check-in as present for a class that you are not attending, you will get a 0 on your attendance and participation grade for the entire semester. Warning: The attendance record is displayed on my computer monitor so it is not difficult to quickly check for discrepancies.

Cell phones and computers are no allowed in class. You may take notes on a tablet computer or iPad. If you use your phone in class, you will get a 0 on your attendance grade for the semester. A second infraction will lower your grade to not higher than a C. If you must use your phone for an emergency, leave the room. You may use computers during the breakout sessions but not cell phones.