

ST 625

REGRESSION ANALYSIS

- Instructor:** Dr Dominique Haughton
Morison 380
Tel: 781-891-2822
dhaughton@bentley.edu
- Objectives:** This course is an introduction to regression analysis, one the fundamental tools in analytics, using the leading statistical packages SAS and SPSS, as well as a data mining tool XLMiner. The class will learn the tricks of the regression trade first with data sets collected at a fixed point in time and will then move on to regression analysis with time series, where data vary over time. Typical problems include building an equation to explain the average SAT score in various states in the United States, or to explain unemployment rates over time. This is a technology-intensive class, so students will become proficient at running analyses in SPSS and in SAS via SAS code and will also gain experience with the symposium software Centra. Participants will get exposed to a data mining package, XLMiner.
- Textbook:** *Data Mining for Business Intelligence*, by Shmueli, Patel and Bruce, second edition, inclusive of license to XLMiner; www.dataminingbook.com .
Reference: A Second Course in Statistics: Regression Analysis, by Mendenhall and Sincich, most recent edition.
- Main software products:** **SAS** on the network at Bentley and/or get a copy for home use (to be discussed in class). **SPSS** on the network; you can also get a copy for home use (to be discussed in class). **XLMiner**, license included with book *Data Mining for Business Intelligence*. If you already own the book and your XLMiner license has expired, please email Duane Lincoln at sales@solver.com with your current license info to get a new license. Centra for class meetings, for office hours and remote team meetings. Earphones and microphone for home use are recommended.
- Homework:** We will schedule approximately one homework per week. Feel free to discuss these with your classmates, but you will need to run and write your own analyses.
- Final public presentations:** At the end of the semester, we will organize ourselves into teams, and each team will choose an interesting problem or case that it

would like to make a public presentation on. This is our chance to show everyone just how great you are!

Course grade:

For each course participant, the final grade will rely on the following: homeworks, inclusive of two “Monster Drills” (35%), the final presentation (20%), class participation (40%), and Centra proficiency (5%). **Please note that in thirty years I have never had an employer ask me for the grade someone received in grad analytics courses. Employers are looking for a certain set of proven skills, not grades.**

Meeting

Topics

1	Introduction to Centra Review of simple linear regression; introduction to SPSS and XLMiner
2	Regression Diagnostics Inference in regression Introduction to SAS
3	Multiple Regression 1
4	Multiple Regression2
5	Monster drill 1
6	Transformations 1; squares, dummy variables
7	Transformations 2; logarithms
8	Multicollinearity; interactions
9	Influence; partial residual plots
10	Monster drill 2
11	Time series 1
12	Time series 2
13	Time series 3
14	Time series 4
15	Public presentations