

GSBA 524 Managerial Statistics

Syllabus – Fall 2015

Class meets: 3:30 – 4:50 PM on T and TH

First class: 08/25/2015

Last Class: 12/03/2015

Final Exam: 2:00 – 3:50 PM 12/15/2015

Professor: Dr. Robertas Gabrys

Office: BRI 401 O

Office Phone: 213-740-9668

Office Hours: W 4:00 – 5:30 PM

E-mail: gabrys@marshall.usc.edu

Course Description

With today's technology companies are able to collect tremendous amounts of data with relative ease. Indeed, many companies now have more data than they can handle. However, the data are usually meaningless until they are analyzed for trends, patterns, relationships, and other useful information. The objective of this course is to learn how to think critically about data, how it was collected and analyzed, and its uses in addressing interesting questions. Knowledge is almost always incomplete, yet decisions must be made. This course explores uncertainty and its relation to decision making, with a particular focus on the ways we can quantify uncertainty and how science is applied in order to improve decisions. The course emphasizes realistic business examples and the processes business analysts actually use to analyze business problems. The course has been designed to be very example-based and practical. Three features define this course: visualization; engagement with real life data; and a modern hands-on approach. All topics in the course are motivated by real life data. No concept will be fully understood until it comes alive in a real life situation. Each topic opens with a business case that frames a question and motivates the content of the topic. We return to the chapter as we present the statistical and analytical methods that provide answers to the question posed in the opening case. Data are used in virtually every major function in business, government, health care, education, nonprofit organizations, etc. For example:

- Annual reports summarize data about companies' profitability and market share both in numerical form and in charts and graphs to communicate with shareholders.
- Accountants conduct audits and use statistical methods to determine whether figures reported on a firm's balance sheet fairly represents the actual data by examining samples of accounting, such as accounts receivable.
- Financial data analysts collect and analyze a variety of data to understand the contribution that a business provides to its shareholders. These typically include profitability, revenue growth, return on investment, asset utilization, operating

margins, earnings per share, economic value added (EVA), shareholder value, and other relevant measures.

- Marketing researchers collect and analyze data to evaluate consumer perceptions of new products.
- Operating managers use data on production performance, manufacturing, quality, delivery times, order accuracy, supplier performance, productivity, costs and environmental compliance to manage their operations.
- Human resources managers measure employee satisfaction, track turnover, training costs, employee satisfaction, turnover, market innovation, training effectiveness, and skills development.
- Within the federal government, economists analyze unemployment rates, manufacturing capacity and global economic indicators to provide forecast and trends.
- Hospitals track many different clinical outcomes for regulatory compliance reporting and for their own analysis.
- Schools analyze test performance and state boards of education use statistical performance data to allocate budgets to school districts.

Data support a variety of company purposes, such as planning, reviewing company performance, improving operations, comparing company performance with competitors' or "best practices" benchmarks. This course introduces and illustrates in a practical way a variety of methods, from simple to complex, to help you analyze data sets and uncover important information and helps develop the quantitative thinking and skills needed for managerial data analysis. In many business contexts, data analysis is only the first step in the solution of a problem. Acting on the solution and the information it provides to make good decisions is a critical next step. Therefore, there is a heavy emphasis throughout the course on analytical methods that are critical in decision making. This course covers fundamental statistical techniques in a managerial setting, featuring case studies and conceptual exercises. Statistical topics include effective use of numerical and graphical summaries, estimation and confidence intervals, hypothesis testing, regression and time series. An important 'big picture' goal of this course is to think about the process of decision making under uncertainty which is applicable not only in classical statistical problems but in many related questions. We will also be looking to tie in the decision making process with material you have been learning in other core and elective courses. Microsoft Excel has been thoroughly integrated throughout this course.

Learning Objectives and Learning Outcomes

The primary objectives of the course are:

- **to learn** fundamental techniques that data analysts use to visualize and summarize data; methods for determining what can be due to chance and what cannot, and the limitations as well as the value of these procedures;
- **to develop** a thorough understanding of business analytics;
- **to equip** you with decision-making tools that you can apply in your business careers;
- **to learn** how to manage data and apply analytical techniques to optimize, simulate and make better decisions;
- **to develop** an application-oriented understanding of statistical inference and regression analysis and their use in decision making.
- **to provide** students with lots of hands-on experience with real problems for which statistics can be applied and challenge them to develop their intuition, logic, and problem solving skills.
-
- **to develop** sophisticated professionals' ability to critically analyze business data.

At the end of the course, the student will be able:

- **to summarize** and **display** data using R and Excel's features;
- **to describe** the place of data in the decision-making process and incorporate data into a decision;
- **to convert** raw data into actionable data that yields business intelligence and insights;
- **to develop** optimization models using R and Excel's Solver;
- **to critically analyze** data and **evaluate** information gathered for decision making in the local, regional and global business environment;
- **to scientifically transform** data into insights for better decision making.

Required Materials

Statistics for Business: Decision Making and Analysis, Robert Stine and Dean Foster, Pearson Custom Business Recourses, **Second Edition**, Boston: Pearson Education, 2013.

Course Notes:

I will be using slides and handouts which will be posted on the Blackboard website before each lecture.

Reading

Reading assignments, consisting of chapters from the book and handouts, are indicated in the table below (see **Topics and Tentative Schedule**). This material will be discussed in

the lectures. I would encourage you to do the required reading before and after each lecture; this will help you grasp the concepts better.

Course website:

Every student is required to have an access code to www.mystatlab.com. I will use **www.mystatlab.com** as the main correspondence and information area for the course. The site will provide schedule updates/changes, e-take-home quizzes, homework assignments, practice problems for exams, additions to reading requirements, etc. It is your responsibility to check the mystatlab site and your email daily to ensure class preparation.

Grading Policies:

Your grade will be based on a score out of 1,000 points, made up of 300 points for the final exam, 150 points for homework, 200 points for quizzes, 200 points for midterm and 150 points for project:

<u>ASSIGNMENTS</u>	<u>TOTAL POINTS</u>	<u>% OF GRADE</u>
COMPREHENSIVE FINAL EXAM	300	30.0%
HOMEWORK ASSIGNMENTS	150	15.0%
REGULAR TAKE-HOME QUIZZES	200	20.0%
MIDTERM	200	20.0%
PROJECT	150	15.0%
TOTAL	1000	100.0%

Final course letter grades will represent how you perform in the class relative to other students, i.e. your grade will be based on **your performance** and a **mandated target of the GPA of 3.3**. Four items are considered when assigning final grades:

1. Your average weighted score as a percentage of the available points for all assignments (the points you receive divided by the number of points possible).
2. The overall average percentage score within all cores.
3. Your ranking among all students in all cores.
4. Target course GPA of 3.3.

You may track your performance on blackboard. **There will be no extra credit.**

Take-home Quizzes (Individual Assignments)

There will be weekly **online** take-home quizzes to supplement the in-class learning. You will have a week to complete each quiz. The quizzes are designed around the specific class topics.

Homework Assignments (Group Assignments)

There will be 4 **group** homework assignments. This is a "hands-on" course: most of the learning will be derived from doing the homework assignments. Most of the homework will involve some computer work. Homework assignments should be typed neatly with necessary computer output and graphics placed in order with each corresponding homework exercise. Figures (including fonts) should be clear and readable. Assignments (hard copies) must be turned in on the due date at the beginning of each class. Each group submits one hard copy. Group members should be listed on the front sheet of each HW assignment. Any assignment turned in late, will receive a grade deduction (20 % reduction per one day). If you are unable to attend class on that day, make arrangements for it to be delivered to the classroom by the start of class. **Late or not, however, you must complete all required assignments to pass this course.**

Project (Group Assignments)

Students will work in teams of five to analyze a case. This case will require you to apply a variety of data analytics you've learned throughout the semester to a complex problem. Your project will involve a write-up summarizing and justifying your recommendations, a 15 minute presentation to the class of your findings, and providing constructive feedback on other team's presentations and analyses. Details of the case and requirements for the project, including grading expectations, will be distributed later in the semester.

Midterm:

There will be one midterm. The midterm will take place on Thursday, October 15 in class. The midterm will be closed book. Sample midterm with solutions will post on the mystatlab website.

Technology Policy

We will use **Excel** and **R** to perform statistical data analysis and modeling. Students are required to bring their laptops to class.

Working Together

Students are encouraged to work together on homework assignments and while preparing for exams. Students who work together generally are more productive, learn more quickly from mistakes, and have more fun. Working together only works when each student in the group is an active participant in the learning process.

Class Attendance

Class attendance will help you understand the subject better. You will have an opportunity to earn participation points and question ideas that are confusing. I encourage, as well as expect, questions during the lectures. Besides the homework and test grades, this is really my only source of information about how comfortable you are with the pace of the course. You need to give me feedback or I will assume that everyone is comfortable with the material and will keep moving forward. I have no problem going over a concept multiple times. If you feel uneasy bringing up your questions in class, there are plenty of opportunities to speak with me one-on-one. I will be more than happy to speak with you before or after class or during office hours.

Study hints

I recommend the following practice:

1. Before the lecture read the entire chapter, taking careful notice of the key words that are in bold type.
2. During lecture, take notes and try to understand the examples and cases I am discussing.
3. After the lecture, review your notes, and then read the chapter again solving the end-chapter problems.

Statement for Students with Disabilities

Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to me (or to your TA) as early in the semester as possible. DSP is located in STU 301 and is open 8:30 a.m.–5:00 p.m., Monday through Friday. The phone number for DSP is (213) 740-0776. For more information visit www.usc.edu/disability.

Statement on Academic Integrity

USC seeks to maintain an optimal learning environment. General principles of academic honesty include the concept of respect for the intellectual property of others, the expectation that individual work will be submitted unless otherwise allowed by an instructor, and the obligations both to protect one's own academic work from misuse by others as well as to avoid using another's work as one's own. All students are expected to understand and abide by these principles. *SCampus*, the Student Guidebook, (www.usc.edu/scampus or <http://scampus.usc.edu>) contains the University Student Conduct Code (see University Governance, Section 11.00), while the recommended sanctions are located in Appendix A.

Students will be referred to the Office of Student Judicial Affairs and Community Standards for further review, should there be any suspicion of academic dishonesty. The Review process can be found at: <http://www.usc.edu/student-affairs/SJACS/>. Failure to

adhere to the academic conduct standards set forth by these guidelines and our programs will not be tolerated by the USC Marshall community and can lead to dismissal.

Emergency Preparedness/Course Continuity

In case of a declared emergency if travel to campus is not feasible, USC executive leadership will announce an electronic way for instructors to teach students in their residence halls or homes using a combination of Blackboard, teleconferencing, and other technologies.

Incomplete Grades

A mark of IN (incomplete) may be assigned when work is not completed because of a documented illness or other “emergency” that occurs after the 12th week of the semester (or the twelfth week equivalent for any course that is scheduled for less than 15 weeks).

An “emergency” is defined as a serious documented illness, or an unforeseen situation that is beyond the student’s control, that prevents a student from completing the semester. Prior to the 12th week, the student still has the option of dropping the class. Arrangements for completing an IN must be initiated by the student and agreed to by the instructor prior to the final examination. If an Incomplete is assigned as the student’s grade, the instructor is required to fill out an “**Assignment of an Incomplete (IN) and Requirements for Completion**” form (<http://www.usc.edu/dept/ARR/grades/index.html>) which specifies to the student and to the department the work remaining to be done, the procedures for its completion, the grade in the course to date, and the weight to be assigned to work remaining to be done when the final grade is computed. Both the instructor and student must sign the form with a copy of the form filed in the department. Class work to complete the course must be completed within one calendar year from the date the IN was assigned. The IN mark will be converted to an F grade should the course not be completed.

About the Instructor

Professor Gabrys teaches undergraduate, MBA and PhD courses ranging from Applied Business Statistics through to advanced Foundations of Statistical Inference. Currently, he teaches Applied Business Statistics, Data Analysis for Decision Making and Business Forecasting in the undergraduate business program; Managerial Statistics, Essentials of Data Analysis and Modeling, Time Series Analysis for Forecasting, and Applied Modern Statistical Learning Methods in the MBA Program. Dr. Gabrys’ research is concerned with data analysis, statistical modeling, forecasting and inference. Areas of current interest include big data analytics, functional data analysis and its applications to environmental sciences, financial econometrics, and space physics; statistical computing, and actuarial mathematics. Dr. Gabrys’ work has been published in the Journal of the Royal Statistical Society, Journal of Multivariate Analysis, Journal of the American Statistical Association, Austrian Journal of Statistics, and Journal of Time Series Econometrics. Prior to joining USC, Professor Gabrys was an actuary at Skandinaviska Enskilda Banken (SEB), one of the leading Nordic and Baltic corporate banks.

Course Content

PART ONE: VARIATION
1. Getting, Describing and Summarizing Structured and Unstructured Data Using Excel.
2. Business Analytics on Spreadsheets
3. Describing Categorical Data
3.1 Looking at Data
3.2 Charts of Categorical Data
3.3 The Area Principle
3.4 Mode and Median
4. Describing Numerical Data
4.1 Summaries of Numerical Variables
4.2 Histograms
4.3 Boxplot
4.4 Shape of a Distribution
5. Association between Categorical Variables
5.1 Contingency Tables
5.2 Lurking Variables and Simpson's Paradox
5.3 Strength of Association
6. Association between Quantitative Variables
6.1 Scatterplots
6.2 Association in Scatterplots
6.3 Measuring Association
6.4 Summarizing Association with a Line
6.5 Spurious Correlation
PART TWO: PROBABILITY
7. Probability
7.1 From Data to Probability
7.2 Rules for Probability
7.3 Independent Events
8. Conditional Probability
8.1 From Tables to Probabilities
8.2 Dependent Events
8.3 Organizing Probabilities
8.4 Order in Conditional Probabilities
9. Random Variables
9.1 Random Variables
9.2 Properties of Random Variables
9.3 Properties of Expected Values
9.4 Comparing Random Variables

10. Association between Random Variables
10.1 Portfolios and Random Variables
10.2 Joint Probability Distribution
10.3 Sums of Random Variables
10.4 Dependence between Random Variables
10.5 IID Random Variables
10.6 Weighted Sums
11. Probability Models for Counts
11.1 Random Variables for Counts
11.2 Binomial Model
11.3 Properties of Binomial Random Variables
11.4 Poisson Model
12. The Normal Probability Model
12.1 Normal Random Variable
12.2 The Normal Model
12.3 Percentiles
12.4 Departures from Normality
PART THREE: INFERENCE
13. Samples and Surveys
13.1 Two Surprising Properties of Samples
13.2 Variation
13.3 Alternative Sampling Methods
13.4 Questions to Ask
14. Sampling Variation and Quality
14.1 Sampling Distribution of the Mean
14.2 Control Limits
14.3 Using a Control Chart
14.4 Control Charts for Variation
15. Confidence Intervals
15.1 Ranges for Parameters
15.2 Confidence Interval for the Mean
15.3 Interpreting Confidence Intervals
15.4 Manipulating Confidence Intervals
15.5 Margin of Error
16. Statistical Tests
16.1 Concepts of Statistical Tests
16.2 Testing the Proportion
16.3 Testing the Mean
16.4 Significance versus Importance
16.5 Confidence Interval or Test?
17. Comparison

17.1 Data for Comparisons
17.2 Two-Sample z-test for Proportions
17.3 Two-Sample Confidence Interval for Proportions
17.4 Two-Sample T-test
17.5 Confidence Interval for the Difference between Means
17.6 Paired Comparisons
18. Inference for Counts
18.1 Chi-Squared Tests
18.2 Test of Independence
18.3 General versus Specific Hypotheses
18.4 Tests of Goodness of Fit
PART FOUR: REGRESSION MODELS
19. Linear Patterns
19.1 Fitting a Line to Data
19.2 Interpreting the Fitted Line
19.3 Properties of Residuals
19.4 Explaining Variation
19.5 Conditions for Simple Regression
20. Curved Patterns
20.1 Detecting Nonlinear Patterns
20.2 Transformations
20.3 Reciprocal Transformation
20.4 Logarithm Transformation
21. The Simple Regression Model
21.1 The Simple Regression Model
21.2 Conditions for the SRM
21.3 Inference in Regression
21.4 Prediction Intervals
22. Regression Diagnostics
22.1 Changing Variation
22.2 Outliers
22.3 Dependent Errors and Time Series
23. Multiple Regression
23.1 The Multiple Regression Model
23.2 Interpreting Multiple Regression
23.3 Checking Conditions
23.4 Inference in Multiple Regression
23.5 Steps in Fitting a Multiple Regression
24. Building Regression Models
24.1 Identifying Explanatory Variables
24.2 Collinearity

24.3 Removing Explanatory Variables
25. Categorical Explanatory Variables
25.1 Two-Sample Comparisons
25.2 Analysis of Covariance
25.3 Checking Conditions
25.4 Interactions and Inference
25.5 Regression with Several Groups
26. Analysis of Variance
26.1 Comparing Several Groups
26.2 Inference in ANOVA Regression Models
26.3 Multiple Comparisons
26.4 Groups of Different Size

The schedule above and procedures described in the syllabus may vary at the discretion of the instructor and/or the overall progress of the class. **It is your responsibility to be aware of any important announcements the instructor makes in class or via email.**