HC 647
Statistical Foundations of Data Analytics
Summer 2017

Professor: Dean Poeth, Ph.D., P.E., C.Mfg.E.

Class: Online, June 12-August 22

Office: Graduate Center, Room 224

Office Hours: Tuesdays and Thursdays, 6:00-7:00 Eastern online and by appointment. Please email me if you would like to set up a time to talk in person, online, by telephone, or via Skype. My Skype ID is XXXXXXXXXXX

E-Mail: XXXXXXXXXXX

When sending me an email, always start the subject line with "HC647" followed by your last name and then the topic. Note that I am unable to check my email during the day. I normally check email in the early evening, and I will respond to your email as quickly as possible. I will, at a minimum, make every effort to respond to emails within 24 hours.

Telephone: (XXX) XXX-XXXX (this is my home phone – please do not call after 9:00 pm Eastern).

Description: This course emphasizes statistical data analysis to support managerial decisions. Where applicable the course material is related to the Six Sigma management philosophy, finance, marketing, or health management applications. With the help of a statistical package, case studies from these areas of application are analyzed and interpreted. Setting the proper problem statement, examining the type of data required and available, analyzing these data, then drawing and discussing possible conclusions for each case are emphasized. Focus on mathematical formulas and statistical theory is minimized.

It is anticipated that each student will need to dedicate at least 12-15 hours per week to course participation, readings, activities, and assignments.

Because this course will be conducted entirely online, it is important for all students to keep up with the required readings and assignments; I have found that once a student falls behind it can be extremely difficult to catch up. There are weekly deadlines to ensure that everyone moves through the course at the same pace. Plan your time accordingly.

Required Software: JMP Version 13 software (available for download via Moodle).

Required Text: *JMP Means Business: Statistical Models for Management*, J. Schmee and J.E. Oppenlander, SAS Institute, 2010.
ISBN 978-1-59994-299-5.

Resources: The SAS Institute offers live webinars to help you get started with the JMP software at no cost. To view the schedule and register go to https://www.jmp.com/en_us/events/getting-started-with-jmp/overview.html Other JMP resources such as tutorials and demos are available on the JMP website.

Prerequisite: IS 502 Introduction to Probability.

Course Technology Policy:

Because this is a web-based course, all students are expected to be familiar with basic computer technology including using email, reading email attachments, using web browsers, reading pdf files, and using word processors such as Microsoft Word. It is also expected that students regularly check their email accounts and Moodle accounts at least once every 24 hours. Any changes to the course schedule or assignments will be made via email and/or posted on Moodle.

It is also expected that all students will install, on their own computers, JMP 13. This is a statistical software package that will be required to complete course assignments and is available for download on Moodle. If you encounter difficulties downloading or installing JMP, immediately contact me or the Clarkson Information Technology Director Bob Keenan.

While it is possible to use Excel to complete some simple assignments early in the course, in the long run this will slow your progress and therefore its use is highly discouraged.

Course Objectives: By completion of this course you should be able to:

- Effectively communicate the use of and results from statistical methods as applied to business problems and decision making.
- Identify certain standard problems in business and finance.
- Synthesize numerical and graphical results of statistical analysis and communicate them in written reports.
- Identify problems and analyze data that require simple comparisons of means (either two-sample, paired and ANOVA designs) as they apply to a variety of situations.
- Estimate and evaluate simple and multiple regression and time series models, especially for forecasting, to find important predictor variables (X) to change or control a response variable (Y).

- Identify problems and analyze data using measures of association to establish empirical "cause and effect."

Competencies Taught:

- Analytic abilities: Defining the problem, gathering and analyzing appropriate data.
- Decision-making skills: Selecting actions based on evidence and analysis.
- Communicate effectively in written form.
- The ability to work effectively in groups.
- The ability to synergistically accomplish common goals.

Course Outline:

The following is the outline of the course including graded assignments. Any changes to this plan will be announced by email and/or on Moodle. I suggest using this as a checklist for class activities and assignments.

Date	Week	Topics	Assignment and Due Date
6/12/17	1	Course Introduction Introduction to JMP Module 1: Describing Data	Course Introduction Quiz opens 6/12/17 closes 6/15/17 (9pm) Discussion 1. First post due 6/17/17 (noon) closes 6/20/17 (9pm)
6/19/17	2	Module 2: Probability Distributions	Discussion 2. First post due 6/24/17 (noon) closes 6/27/17 (9pm) Assignment 1 Data description memo due 6/27/17 (9pm)
6/26/17	3	Module 3: Single Variable Analysis	Discussion 3. Data visualization. First post due 7/1/17 (noon) closes 7/5/17 (9pm)
7/3/17	4	Module 4: Comparing Two Means	Assignment 2 PowerPoint Slides due 7/5/17 (9pm)
7/10/17	5	Module 5: One-way Analysis of Variance	Case Study #1 Sherri's Bakery due 7/11/17 (9pm) Discussion 4. First post due 7/15/17 (noon) closes 7/18/17 (9pm)
7/17/17	6	Module 6: Two-way Analysis of Variance	Discussion 5. Olive Oil. First post due 7/22/17 (noon) closes 7/25/17 (9pm)
7/24/17	7	Module 7: Methods for Nominal Variables	Roundtable Discussion (ungraded) opens 7/24/17 closes 7/28/17 (9pm) Discussion 6. First post due 7/29/17 (noon) closes 8/1/17 (9pm)
7/31/17	8	Module 8: Simple Linear Regression Introduction to case study: Challenger Disaster	Assignment 3 Startup due 8/1/17 (9pm)
8/7/17	9	Module 9: Multiple Regression Analysis	Discussion 7. First post due 8/12/17 (noon) closes 8/15/17 (9pm)
8/14/17	10	Module 10: Ethics of Statistics and Big Data Introduction to Predictive Analytics	Discussion 8. First post due 8/19/17 (noon) closes 8/22/17 (9pm) Case Study #2: Challenger Disaster due 8/22/17 (9pm)

Course Grading:

In this course statistical models are used to analyze data to support business decision making. Management expects the analyst to provide correct results obtained from appropriate analysis methods. In addition, the data should be thoroughly examined and summarized. The grading criteria for each assignment will follow these management expectations for quantitative analysis. Even if you do not have to analyze data in your future, you may well supervise those who do and will need to be prepared to critically evaluate their work.

The grading criteria are:

- Presentation quality, which includes use of correct spelling and grammar.
- Correct results.
- Application and interpretation of statistical methods.
- Examination and description of the data.

These criteria form the essential components of a quantitative analysis, which begins with examination and description of data. The analyst must understand the data source and the quality and pedigree of the data and examine the data for potential outliers. Data is summarized by providing graphs and numerical statistics. Once a thorough understanding of the data is obtained, the correct statistical method or model that addresses the business problem is applied. The analyst shall be prepared to explain the method in layman's terms (i.e., without using statistical jargon) along with any necessary assumptions and interpret the results. In business, important decisions may be influenced by the results of quantitative analysis, therefore the analyst must ensure the results are correct. Finally, the results and information must be presented clearly and concisely in the context of the business problem.

Final course grades will be based on the following:

Item	Weight (%)
Course Introduction Quiz	5
Discussions	30
Assignment 1 Data Description Memo	10
Assignment 2 PowerPoint Slides	10
Assignment 3 Business Startup	15
Case Study #1 Sherri's Bakery	15
Case Study #2 Challenger Disaster	15

Discussion Grading

Students are expected to be active participants in this course. Active participation includes asking and answering questions in discussion forums. An initial post is due by Saturday noon EST and at least two more responses to other posts are the minimum expected.

All posts must add substantively to the topic. For example, a post such as, "I agree with what John posted last week," is not substantive and will not count toward the frequency of participation component of the grade.

I primarily want to see that you are taking the time to first work through the assignment independently, read and reflect on what others in the class are saying, and in addition posting your own comments, questions, and thoughts. A post made at the end of the discussion period does not allow sufficient opportunity for your classmates to read and respond to your submission and will have a lower weight.

Stick to the subject and write thoughtfully and concisely. Be sympathetic and encouraging to others. Avoid disrupting the flow by introducing unrelated subjects. Instead, wait until the current topic winds down.

Each discussion has a closing date and time. Late submissions will not count toward your discussion grade.

Each discussion will be graded on a 10 point scale, and each discussion will account for 3.75% of the final course grade. Please note that discussion grades are for the thoughtfulness, active engagement, and professionalism shown by your comments and responses, not their technical correctness.

Discussion grading rubric

Criteria	Possible points
Your initial post (thoughts & answers to assignment, having taken the time to work through the assignment independently first)	5
Posted at least two additional times (comments and responses to what others have posted, answers to my questions, additional ideas)	5

Graded Assignments

Graded assignments (with the exception of PowerPoint presentations) must be double-spaced, Arial 11 font. All assignments are due on or before the due date and time in pdf format only. All assignments shall use this file format:

HC647_assignment name_your last name.pdf Writing quality (including spelling and grammar) as well as content will be evaluated. Do not write in the first person voice, and always explain statistical terminology to the reader. Assignments shall follow a clearly recognizable format and structure and incorporate graphs and only relevant computer results from JMP. Assignments must be within the prescribed page limits. Please write your name on the first page of each submission. Students may work together on assignments, but each must turn-in a separate and original submission for grading.

Late Assignment Policy

All students are required to submit assignments on or before the due date and time. It is acceptable to submit assignments early. However, due to the fast pace of this course, late assignments will not be accepted and extensions will not be granted so please do not ask. Plan your schedule accordingly. There is no extra credit available in this course.

Getting Acquainted with the Course Assignment.

This 10-week course moves at light-speed. It is extremely important that you begin the course as soon as the term starts and get acquainted with its general layout and structure. The first assignment will help you accomplish this and it is due by 9pm on June 15. The assignment will be a quiz and will ask about information you can easily find in the syllabus and will ask questions ensuring you have completed certain tasks.

Study Problems:

Ungraded problems are given to reinforce the methods and concepts presented and to prepare for the graded assignments. Problem solutions will be posted to Moodle. Students are encouraged to work the problems prior to studying the solutions.

Module	Problems	Objective(s) of Study Problems
1	Chapter 3 - 3, 4, 9	Use numeric and graphic statistical analysis to critically evaluate and summarize data.
2	Chapter 5 - 4, 5 Chapter 6 - 4, 7	Apply basic probability operations and models to business problems.
3	Chapter 7 - 2, 9 Chapter 8 - 6, 10	Apply one-sample hypothesis tests and confidence intervals to business problems.
4	Chapter 9 – 3, 5	Apply two-sample hypothesis tests and confidence intervals to business problems.
5	Chapter 10 – 2, 3	Compare the means from several groups using one-way ANOVA.
6	Chapter 11 – 2, 9	Compare the means from several groups using two-way ANOVA.
7	Chapter 12 – 2, 3, 6 Chapter 13 – 1, 2	Conduct tests and compute confidence intervals for proportions and Chi-square tests for independence.

8	Chapter 14 – 3 Chapter 15 – 1	Focus on interpretation of simple regression analysis results: meaning of coefficients, recognition of poor fit and model inadequacy, outlier detection. Describe the association between continuous variables through simple correlation analysis.
9	Chapter 16 – 3 Chapter 17 – 3, 5 Chapter 18 – 7	Focus on constructing and interpreting multiple regression analysis: meaning of coefficients, assessing model adequacy, multicollinearity, and importance of predictors. Create regression models containing both nominal and continuous variables.

Supplemental Information:

Examination of data, application of statistical models and methods and the effective communication of statistical results are emphasized in this course. Written work shall be clear, concise, and focused on the application of statistical methods to the business problem.

We will be using JMP extensively throughout the course. Most of the answers will rely on some computer results, both numerical and graphical. All computer packages provide you with comprehensive output; much of it is unnecessary in answering the question posed in your problem statement. Therefore select only those results that support your analysis. Do not append the totality of the output at the end; it will earn you a reduction in your grade. Incorporate the output into your answers through cutting and pasting.

Be careful if you decide to share your homework with a colleague before the assignment deadline. It's fine to work together if you choose, but remember that each student is required to submit a separate and original submission for grading.

Materials used in this course may be subject to copyright protection.

Roundtable

The content of this course is the direct product of prior students' comments and suggestions for improvement over my many years of teaching. This feedback has primarily occurred through roundtable discussions that are held during every class I teach. Your input is critical to the continued success of this course, and I hope you will participate in this optional discussion.

Academic Integrity

You are expected to practice academic honesty in every aspect of this course. Make sure you are familiar with the Clarkson University Regulations, especially Section IV Academic Integrity. Students who engage in academic misconduct are subject to university disciplinary procedures, as well as consequences with regard to this course.

School of Management Grading Guidelines:

The faculty of the School of Management has adopted the following grading guidelines. Though they are not meant to be a rigid set of rules, each professor is expected to use these guidelines in combination with his/her assessment of the students' abilities and performances to arrive at final grades.

Grade Meaning Guidelines

(A) *Exceptional* performance: Consistently displays original thinking; good organization; capacity to analyze and synthesize; superior grasp of subject matter with sound critical evaluations; evidence of extensive knowledge base.

(**A-**) Similar to A, but exhibits occasional gaps in knowledge or critical thinking skills

(B+) Similar to B, but on occasion, displays superior knowledge or critical thinking skills.

(B) *Competent* performance: evidence of grasp of subject matter; some evidence of critical capacity and analytic ability; reasonable understanding of relevant issues; evidence of familiarity with the literature.

(B-) Similar to B, but occasionally fails to exhibit average understanding or thinking skills and occasionally fails to produce minimally acceptable work.

(C+) Similar to C but with more frequent displays of competent knowledge and thinking skills.

(C) *Unacceptable performance:* Displays a general lack of understanding of the subject matter; frequently fails to develop solutions to simple problems in the material; often produces uninspired work that is faulty and lacking style and rigor. Without compensating higher performance in other classes, such students are generally deemed to be unfit to graduate.

(F) Lack of competence or willingness to complete work is evident to the point that the student should be immediately declared unfit to graduate.

Typical Grade Distributions   It is anticipated that final grades in an average class of reasonable size (10 or more) will normally result in an average class GPA in the 3.2 – 3.7 range with grade distributions falling within the following broad limits:

Grade A or A-	30% - 60% of the class
Grades B+, B, or B-	40% - 70% of the class
Grade C+, C, or F	0 - 20% of the class