Data Analysis for Managers : 91-801  
Fall 2014  Wednesday, 6:00 – 8:50 p.m.

Course Information

**Instructor:**  
Kathy Smith  
3027 Hamburg Hall  
Phone: (412) 268-4408  
Email: ks54@andrew.cmu.edu

**Office Hours:**  Wednesday 4:30 – 5:30 or by appointment

**Teaching Assistants:**  
Rocco Pacella  
e-mail: roccoapacella@gmail.com

Joel Acie  
e-mail: joel7754@yahoo.com

**Prerequisites:**  
none

**Course Materials:**  

You may use an older edition. However, all page numbers, chapters, and problem numbers on the syllabus, assignments and outline sheets refer to the current edition. If you are using an older edition it is your responsibility to be certain that you are using the correct information.

**Course Description:**  
This course will provide an introduction to the principles of data collection, description and analysis. You will learn the basic tools of statistical inference and modeling, as well as some fundamentals of designing a statistical study, how to sample and collect data, and which statistical techniques are appropriate. You will also learn how to interpret statistical output, and how not to be fooled by statistical studies.

**Recitation and TA Office Hours:**  
Saturday Recitation: 8:30 – 9:50 a.m. Room 1001  
TA Office Hours: 10:00 – 11:00 a.m. on Saturday (following Recitation)

Recitation will consist of a review of the material taught in class the previous week, with time allowed for questions and clarifications. The TA will come to recitation with examples prepared. Attend this session if you want to rehash and clarify the material taught in class.
At office hours, the teaching assistants will answer specific questions that you have that are related to the assignments or to something that was done in class. They may elaborate and clarify topics as they come up at these review sessions. You should come to office hours prepared with your questions. TAs will not answer the assignment questions for you, nor will they tell you if your answer is correct. They will answer general questions on the assignment, or may do a similar problem with you.

**Course Objectives:**
At the end of this course, you should be able to
- Present data visually in tabular and graphic form
- Summarize a set of observations by reporting a measure of center and dispersion
- Explain how and why sample data can be used to estimate descriptive measures of populations when census data is unavailable, and how we measure the accuracy and precision of the estimate
- Apply the basic rules of probability
- Find and interpret the probability for a random variable which has a normal distribution
- Explain how to take a proper scientific sample that can be used to make inferences about the larger population
- Explain what sampling error is and why it exists
- Classify data by type (quantitative or quantitative, discrete or continuous) and use the proper summary statistics and tests for the data type
- Interpret the p-value, test statistic and other Minitab output from a test of hypotheses, confidence interval, and linear regression
- Explain what it means if test results or poll differences are statistically significant
- Apply the concepts of sampling, estimation, and hypothesis testing to real world examples from polls and surveys, clinical trials and observational studies.

**Course Organization**
I have divided the course into four units. **Unit 1** focuses on tools for summarizing and describing a set of data. Graphical, tabular and numeric tools will be examined. We will also study probability in this unit, which is the tool that allows us to make scientific, sound estimates and conclusions from sample data.

**Unit 2** looks at how, when properly collected, we can take observed data from a sample and use it to make estimates of characteristics of the larger population from which it came. We will look at some actual polls and surveys as an application of statistical estimation. Exam #1 will be given at the end of Unit 2.

Following the midterm exam, we will continue to use sample data to make inferences about the larger population from which it was taken. But in **Unit 3**, we start with a hypotheses about some unknown characteristic of a population, and we use sample data to test that hypothesis. Several types of hypothesis tests will be studied. Actual clinical trials and population studies, which are real-life applications of tests of hypotheses, will also be discussed.

In the final unit, **Unit 4**, we will use regression analysis to model relationships between variables. Correlation analysis and Analysis of Variance (ANOVA) will also be studied. The final exam will cover the material in Units 3 and 4.
Class Format and Preparation
At the beginning of each class, I will distribute Class Examples that will be done that evening, along with Class Notes. Article Packs and other supplementary materials may also be distributed. All of your materials will be in a folder with your name on it; you should pick up your materials from this folder as you come into class each week. Most class materials will also be available on Blackboard at http://www.cmu.edu/blackboard/

Laptops, Cell Phones and Other Messaging Devices
No laptops, cell phones, tablets or other electronic devices may be used during class unless I have explicitly stated so. If you bring these with you, they should be put away, off your desk, and any ringers turned off. If you need to have access to your cell phone or laptop during class, please let me know before class.

Performance Measures
Assignments:
Eleven assignments will be made during the term. Homework is due at the beginning of class. No late assignments will be accepted unless either you have made prior arrangements with me or an emergency occurs. I will drop the lowest assignment score at the end of the semester.

Article Analysis:
Choose a current article from a periodical, newspaper, journal, the internet, or other published source. Write up a short (approximately 3 - 5 pages) analysis of the article similar to what I do in class. This will be graded on form, including style and organization, and on the completeness, correctness, and clarity of your analysis. It is due on November 12 and will contribute 10% of your final course grade (i.e. the same amount as five weekly assignments). As such, you should plan to put an equivalent amount of effort into this project. The article analysis can be either an individual or group project. More details will be given later.

Exams:
Two exams will be given during the semester. You will be permitted to bring one 8-1/2 x 11 inch ‘note sheet’ to use during each exam. The first exam will be given in class during the first 80 minutes of class on October 8. The final exam will be given on the last night of class, December 10 for the full class period.

Make-up Policy:
You must take the exams during the specific times unless you have made prior arrangements with me, and these will only be granted in extreme, unusual, unavoidable situations. Travel plans that conflict with the exam period are not considered to be unavoidable.

Re-grade Policy:
If you think that an exam or assignment was graded incorrectly or unfairly, written appeals will be accepted up to one week after the paper was returned to you. Your appeal should include the number of the question and the specific reason why you believe that it should have been marked as correct.
Individual and Group Work / Academic Integrity
You may work with other students on the assignments. If you work with other students, I realize that your analyses may essentially say the same thing. However, the words that you submit must be your own, and written in the style that you use on exams and other items. If anything is copied from the text or another published source, it must be documented. If this is not done, it is plagiarism. If anything is copied from (that is, you use the same words as) another student or any other source, it is considered to be cheating.

You are not permitted to be in possession of any assignments or exams from another student either from the current semester or from past semesters whether they are electronic or paper. Furthermore, you not permitted to share your work with other students. Possession of or sharing such files constitutes an infraction of the academic integrity policies of this course. Sharing your assignments with any other student in any form (a paper document, an electronic file, photo image, or any other format) will be considered cheating.

If you submit plagiarized or copied work, you will receive a grade of zero on the assignment or exam. In addition, all cheating and plagiarism infractions will be reported in writing to the Associate Dean for the program, the Dean of Student Affairs, and the Dean. They will review and determine if expulsion should be recommended. The report will become part of the student’s permanent record. Further information about cheating and your rights can be found on the CMU website at http://www.cmu.edu/policies/documents/Academic%20Integrity.htm

Performance Evaluation
Your final grade for the course will consist of grades from the four items above. They will be weighted as follows:

- 20% Assignments
- 10% Article Analysis
- 35% Exam # 1
- 35% Final Exam

Each grade will be multiplied by the appropriate weight, and you will end up with a score between 0 and 100. (Numbers will be rounded in the usual way.) Your final grade will be determined by your placement on the following scale:

- 98 - 100 A+ 70 - 74 B-
- 90 – 97 A 65 - 69 C+
- 85 - 89 A- 60 - 64 C
- 80 - 84 B+ 55 - 59 C-
- 75 - 79 B 0 - 54 R

I do not “curve” grades, nor do I grade to the average. You must master a minimum of 55% of the material (as measured by the points that you receive on assignments, quizzes and exams) in order to pass the course, 75% to receive a grade of a B, etc.
## Syllabus Overview – as of August 27, 2014

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Topics</th>
<th>Readings</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Aug. 27</td>
<td>Describing Data: Graphs, Charts, and Tables</td>
<td>Ch. 1: All Ch. 2: 2.1 - 2.4 Article Pack #1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Sept. 3</td>
<td>Describing Data: Numeric Summaries</td>
<td>Ch. 3: 3.1 – 3.4 Article Pack #1</td>
<td>Assignment #1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Standardized Values / z-scores</td>
<td>Ch. 2: 2.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ch. 5: 5.1 – 5.5, 5.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ch. 6: 6.1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Sept. 10</td>
<td>Probability Theory</td>
<td>Ch. 8: 8.1, 8.2, 8.4</td>
<td>Assignment #2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Probability Distributions</td>
<td>Ch. 9: 9.1 – 9.3, 9.6</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Sept. 17</td>
<td>Sampling Distributions</td>
<td></td>
<td>Assignment #3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Estimating π with Confidence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Sept. 24</td>
<td>Determining Sample Size</td>
<td>Ch. 8: 8.3</td>
<td>Assignment #4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Confidence Intervals μ</td>
<td>Ch. 9: 9.4 – 9.5, 9.7</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Oct. 1</td>
<td>Sampling and Bias</td>
<td>Ch. 4: 4.1 – 4.3, 4.6 – 4.8 Article Pack #2</td>
<td>Assignment #5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Polls and Surveys</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Oct. 8</td>
<td>Normal Distribution</td>
<td>Ch. 7: 7.1 – 7.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Standardized Values</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Oct. 15</td>
<td>Hypothesis Tests for Proportions – One Sample</td>
<td>Ch. 10: 10.1 - 10.2, 10.6 Distributed Material</td>
<td>Assignment #6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Oct. 22</td>
<td>Two Sample Test for Proportions</td>
<td>Ch. 11: 11.1, 11.6 Article Pack #3</td>
<td>Assignment #7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Clinical Trials / Experiments</td>
<td>Ch. 4: 4.4</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Oct. 29</td>
<td>One and Two Sample Tests for Means</td>
<td>Ch. 10: 10.3 – 10.5 Ch. 11: 11.1 – 11.3, 11.5 Ch. 2: 2.3 (Dotplot only)</td>
<td>Assignment #8</td>
</tr>
<tr>
<td>11</td>
<td>Nov. 5</td>
<td>Analysis of Variance</td>
<td>Ch 12: 12.1 – 12.4 Article Pack #4</td>
<td>Assignment #9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Observational Studies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Nov. 12</td>
<td>Correlation Analysis</td>
<td>Ch. 2: 2.5 Article Analysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simple Regression</td>
<td>Ch. 3: 3.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ch. 15: 15.1 – 15.5</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Nov. 19</td>
<td>Multiple Regression</td>
<td>Ch. 16: All Article Pack #5</td>
<td>Assignment #10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ch. 17: 17.4</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Dec. 3</td>
<td>More on Multiple Regression and Correlation Analysis</td>
<td>Distributed Material Article Pack #5</td>
<td>Assignment #11</td>
</tr>
<tr>
<td>15</td>
<td>Dec. 10</td>
<td>FINAL EXAM</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>