2014

STAT 500 Business Statistics

Brett Stowell
stowellbs@xavier.edu

Follow this and additional works at: http://www.exhibit.xavier.edu/statistics_syllabi_summer_2014

Recommended Citation
http://www.exhibit.xavier.edu/statistics_syllabi_summer_2014/2

This Restricted-Access Syllabus is brought to you for free and open access by the Statistics Syllabi 2014 at Exhibit. It has been accepted for inclusion in Statistics Syllabi Summer 2014 by an authorized administrator of Exhibit. For more information, please contact exhibit@xavier.edu.
Instructor: Brett S. Stowell  Email: stowellbs@xavier.edu  Phone: 812/290-3584
Office / Mail Drop: Smith 102  Office Hours: By appointment before or after class
Text:  Business Statistics: Communicating with Numbers; Jaggia / Kelly
       Numbers Rule Your World by Kaiser Fung

Williams College of Business Mission: “We educate students of business, enabling them to improve organizations and society, consistent with the Jesuit tradition.”

My Vision: The information gleaned from applying the statistical tools you will be learning in this course impacts each of us in our daily lives. As managers and executives, you will be requesting, generating, and acting upon reports that have used these tools to take some “glob” of data and transform it into information. Given the nature of this thing called statistics, it is possible either by intent, negligence, or legitimate error to come up with the “wrong” answer doing all of the right things. It is my hope that by the end of this course you will have a sufficient understanding of the subject matter to be able to ask more pertinent questions of the data you have available and to be able to rationally evaluate the validity of information derived from that data. The study of statistics is multidisciplinary in nature, and I guarantee you that some or all of this material will appear again in your studies and career.

Course Description: This course is designed to familiarize you with some of the basic statistical tools and techniques used to transform raw data into actionable information. As such, class time will be divided between reviewing homework assignments, introducing new material, discussing real-world applications, and employing newly introduced skills.

Course Materials: In addition to the text, you are responsible for materials posted on Canvas and McGraw-Hill Connect. Homework assignments, supplemental reading, and other information will be posted regularly. Students without access to a PC running Excel should make use of the Virtual Lab for completing assignments, projects and exams. (http://www.xavier.edu/ts/students/Virtual-Desktop.cfm)

Grades:  Your final grade will consist of two exams (50%), a final project (30%) and regular homework and cases (20%). Possible grades are A (95-100), A- (90-94.9), B+ (85-89.9), B (80-84.9), B- (75-79.9), C+ (70-74.9), C (65-69.9), F (Below 65). You must earn at least a “C” on each of the exams and the project to pass the class; the instructor will provide anyone receiving a failing grade an opportunity to demonstrate the required competency, replace the failing grade with a 65%, and thus pass the class. No other extra credit is available.

Due Dates: Due dates and test dates are firm. Students are expected to contact the instructor prior to an expected absence to make arrangements. Late assignments will receive no credit.

Academic Honesty: Do your own work. Review Xavier’s Academic Honesty policy for more information.

Xavier University Policies: Xavier University policies regarding privacy rights, incomplete work and attendance, and academic honesty will be strictly enforced. Please see the current University catalog for more on these policies.
Tentative Course Calendar

<table>
<thead>
<tr>
<th>Class</th>
<th>Topics Covered</th>
<th>Assignments Due</th>
</tr>
</thead>
</table>
| 8/4     | Introduction to Course  
Introduction to Xavier's Library  
Introduction to Virtual Desktop  
JK 1.1 – 1.3  
JK 2.1, 2.2, 2.4  
JK 3.1 - 3.4, 3.6-3.8 | Initial read of *Numbers Rule Your World*            |
| 8/11    | JK 4.1-- 4.3  
JK 5.1, 5.2, 5.4 – 5.6  
JK 6.1, 6.2 | Project Proposal  
Connect Homework  
Case Study 5.1 |
| 8/18    | JK 6.3  
JK 7.1-7.5  
JK 8.1-8.5 | Connect Homework  
Case Study 7.2 |
| 8/25    | JK 9.1-9.4  
JK 10.1-10.3 | Exam 1 (CH 1-8)  
Connect Homework |
| 9/8     | JK 13.1, 13.3, 13.4 | Exploratory Analysis Submission  
Connect Homework  
Case Study 10.2 |
| 9/15    | JK 14.1-14.4 | Connect Homework  
Case Study 13.3 |
| 9/22    | JK 15.1, 15.4  
JK 16.1, 16.2 | Connect Homework  
Case Study 14.2 |
| 9/29    | JK 17.1, 17.2  
Regression Review | Final Project Paper Draft Submission  
Connect Homework  
Case Study 16.2 |
| 10/6    | Topics from JK 18-20 | Exam 2 (CH 9,10, 13-17)  
Final Project Paper Submission  
Connect Homework  
Case Study 17.2 |
| 10/13   | Project Presentations | Final Project Oral Defense |

**Class Preparation:** Students should review designated learning objectives for each class session prior to arrival.

**Homework / Cases:** Homework and solutions to be completed for student learning are provided for each learning objective. Students are encouraged to assist one another in the completion of homework and cases provided each student submits his or her original work for credit.

**Exams:** All examinations are to be completed at home. Class notes, Excel help files and the textbook are appropriate resources to use while completing the exams. Assistance from any third party, except for the instructor, is prohibited.

**Project:** See the project addendum at the end of this document.

**Learning Objectives:** Students completing this course will be able to evaluate the veracity of reports, claims and decisions based on statistical evidence, develop frameworks for utilizing appropriate statistical tools to locate and clarify problems, evaluate alternatives and draw evidence-based conclusions, and communicate relevant findings to stakeholders.

---

1 MBA Learning Goals: Strategic Thinking and Leadership
2 MBA Learning Goals: Critical Thinking
3 MBA Learning Goals: Effective Written and Oral Communication
Final Project Guidelines

Overview

The objective of this project is to use several\(^4\) statistical tools to produce one of the following options:

1. A short chapter that could be appended to *How to Lie with Statistics*. This option works under the assumption that you are a trained statistician who is reading a newspaper/blog and encounters an analysis of some data set (impact of government policy comes to mind of late) and critiques the analysis provided in the source piece using a proper understanding/use of statistical tools and additional data sources.

2. A short chapter that could be appended to *Numbers Rule Your World* where data from your own business unit (or another business problem) is analyzed. (You will want to confirm with a manager that it is appropriate to use any data presented.) You will identify a problem to be solved, collect the appropriate data, perform the analyses, and report.

3. An alternative project mutually agreed upon by student and instructor.

Roadmap

The due dates listed on the roadmap are firm. You are encouraged to submit the proposal and the exploratory analysis early as these steps are intended to provide you with feedback for improving your project and are not graded.

- **Project Proposal Submission**
  - Prior to beginning the project, you must submit a brief (1-3 sentence) description of the research to be conducted. This abstract must include, at a minimum, the piece to be critiqued (option 1; attach a pdf or include a link), the business problem to be solved (option 2), or a lengthier proposal for an alternative project (option 3), and an initial overview of the problem\(^5\) and data source(s)\(^6\) you propose to use in the analysis.

- **Exploratory Analysis Submission**
  - Once your proposal is approved, the next step is to perform some combination of the following:
    - Visualize the data (graph, histogram, box-plot, etc.)
    - Summarize the data (Five number summary, distribution, etc.)
    - Hypothesize
    - Analyze Variations
    - Identify dependent variable and proposed independent variables to test.
  - These items, plus an outline (think “bullet points”) of your analysis should be submitted for this component.

- **Rough Draft Submission**

- **Final Paper Submission**
  - Your final paper should not exceed 1500 words (about 5 typewritten pages) excluding figures.

- **Oral Presentation and Defense**
  - You will make a presentation of your paper before the class.
    - The presentation should include the problem to be solved and your findings. Plan on it being about 5 minutes; you will be cut off at 7.
    - Following the presentation the instructor and students will critique your paper. You must be ready to field questions and defend your project.
    - This is a presentation to a manager, governing body, etc. Academic work often focuses on process; this presentation should focus on result e.g. “In evaluating x I found y which indicates z.”

---

\(^4\) You will define which tools you intend to use in your proposal. I will let you know if it meets the threshold of “several” statistical tools.

\(^5\) Either the problem with the analysis in the piece or the problem to be solved.

\(^6\) I highly recommend you visit the Xavier Library’s website and check out the databases.
Prompts

If you are stuck (in terms of selecting a topic) the TV show *Numb3rs* has demonstrates some of the statistical analyses we have used in class as applied to a real world problem:

<table>
<thead>
<tr>
<th>Episode</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>104</td>
<td>Histograms</td>
</tr>
<tr>
<td>109</td>
<td>Regression</td>
</tr>
<tr>
<td>202</td>
<td>Scatterplots</td>
</tr>
<tr>
<td>217</td>
<td>Binomial Theorem</td>
</tr>
<tr>
<td>224</td>
<td>Histograms, Relative Frequency, Probability</td>
</tr>
<tr>
<td>308</td>
<td>Least Squares Regression, Data Analysis &amp; Scatter Plots</td>
</tr>
<tr>
<td>311</td>
<td>Box-and-Whisker</td>
</tr>
<tr>
<td>316</td>
<td>Weighted Averages, Z-scores</td>
</tr>
<tr>
<td>316</td>
<td>Scatterplots, Least Squares Regression, Residuals</td>
</tr>
</tbody>
</table>

Nearly every episode deals with probability in some fashion. Not necessarily suitable for children.

Example

There are multiple acceptable projects you may choose. Here is an example of a proposal and the components of the analysis.

*Option 2 Proposal*

I want to develop a model to predict the rental rate of one bedroom apartments in different neighborhoods in Cincinnati. To develop this model I am going to collect data from cincyrents.com on one-bedroom apartments in Hyde Park, College Hill, Oakley, and Price Hill. I will be using the mean household income, from census records, as a proxy for the affluence of the neighborhood, the number of square feet in the apartment, the age of the building, what floor the apartment is on, and whether or not there is off-street parking available.

*Exploratory Analysis Components for Regression-Based Model*

1. A statement of the problem to be solved – expand upon the what and why that you included in your proposal.
2. A table containing your dependent and independent variables and data source(s).
3. Any notations about data transformations and the rationale behind them.
4. Visual representation(s) of the data including scatter plot(s), a regression line, box-plots, or any other appropriate graphical presentation. (Should be included in text with appropriate reference e.g. See figure 1 for…)
5. A table containing the regression output from Excel.
6. An interpretation of the Excel template containing, at a minimum, a description of how accurate the model will be in predicting future outcomes. This discussion should include the overall regression equation’s accuracy (R Square & Adjusted R Square), the probability that this output was not by chance
(ANOVA – Sig F), individual regression coefficient and y-intercept accuracy, and a visual analysis of residuals.

a. Note the difference between R Square, which tells how well the regression line approximates the real data, but will almost always increase when additional variables are introduced, and the Adjusted R Square, which is more conservative and only increases when additional variables increase the accuracy of the regression equation.

b. The significance of F indicates the probability that the regression output is a result of chance. An F = .040 indicates that there is only a 4% chance that the regression output was a chance occurrence.

c. The p-value of each coefficient indicates the likelihood that the coefficient or Y-intercept is valid. Larger p-values indicate a greater probability that the result occurred due to random chance.

d. The residuals are the difference between the regression’s predicted value and the actual value of the output variable. These should be plotted on a scatterplot chart. The more random and centered around zero the residuals appear to be, the more likely it is that the equation is valid.

7. A discussion of how the model could or could not be deployed to solve a “real life” problem based upon the results noted in item 6 above.

8. A discussion of how you could improve upon the model based upon the results noted in item 6 above. (See “multicollinearity” for a discussion about dividing explanatory power amongst correlated independent variables.)