Carnegie Mellon University, Heinz College  
Business Intelligence & Data Mining with SAS Suite  
Spring 2016 (94-832)

Course Information

**Course**

*Course Number/Section*  
94-832

*Course Title*  
Business Intelligence & Data Mining with SAS Suite

*Dates and Room*  
Tue/Thu 3:00pm-4:20pm (Section B3), HbH 1502  
Tue/Thu 4:30pm-5:50pm (Section A3), HbH 1502

**Professor Contact Information**

*Professor*  
Beibei Li <beibeili@andrew.cmu.edu>

*Office Location*  
HbH 3026

*Office Hours*  
Tue 6:00pm-7:00pm (or by individual appointment)

**TA Information**

Yingjie Zhang yingjie2@andrew.cmu.edu  
Wed noon-1:30pm, Fri 2:30pm-4:30pm (HBH 3005)

Bhawna Gaba bgaba@andrew.cmu.edu  
Thu 2-4pm, Sat 12-1:30pm

**Course Pre-requisites, Co-requisites, and/or Other Restrictions**

No official prerequisite, preferably some knowledge in statistics, economics and database.

**Course Description**

With the proliferation of Web 2.0 making inroads into the enterprises and industries, the ability to understand, analyze and interpret businesses from Big Data has become increasingly more important today. This class aims to equip you with highly demanded business analytics skills in the current job market. The course will focus on extracting business intelligence by leveraging firm's business data as well as online social media content for various applications, including (but not limited to) search engine marketing, social media analytics, crowd-sourcing management, market analysis and demand estimation, social network analysis, customer segmentation, customer relationship management (CRM), web mining and health care management.

The class will be hands-on and the emphasis will be placed on the "know-how" aspect - how to extract and apply business intelligence to improve business decision making and marketing strategies. We will analyze real-world business data from Fortune 500 companies using various business intelligence tools, primarily SAS Enterprise Miner. Time permits, we will also introduce some advanced economic and predictive models in analyzing digital markets. Prior programming skill is not required. Throughout the class business insights from several market leaders such as Google, Microsoft, Amazon, Travelocity, TripAdvisor, Netflix and Facebook will be revealed.

**Class Objectives**

- Differentiate, design and assess various business intelligence (BI) and data mining models.
- Identify and translate real-world business problems into BI and data mining problems.
- Exhibit ability in pre-preparing and visualizing the right data towards these problems.
- Implement efficient BI strategies to solve these problems.
- Develop proficiency in BI software (SAS Enterprise Miner).
- Enhance knowledge and skills in the current trends in the management and use of BI.
Recommended Textbooks and Materials

Recommended Textbook (temporary)

Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management, 3rd Edition

Gordon Linoff and Michael Berry, 2011, Wiley,

ISBN: 0470650931

Additional Readings (TBD)

Deliverables and Grading

Assignment submission instructions

You will submit your assignments (in the required file format with a simple file name and a file extension) by using the Assignments tool at Blackboard. For the team project assignment, one group member will submit the assignment for the group and all group members will be able to view the results and feedback once it’s been graded.

Instructions for each assignment will be posted to Blackboard and the rubric will be provided in class. Assignments are due on the dates stipulated by the instructor on the syllabus or in class. Assignments will not be accepted past the due date and time unless a religious observance or a documented medical condition prevents on-time submission and the student has consulted with the instructor in advance for approval of an alternate deadline.

Class participation

Your class participation is extremely important for your final grade. The grade I assign for your participation is a careful, subjective assessment of the value of your input to classroom learning. I keep track of your contributions towards each class, and these contributions include (but are not restricted to):

- Attend class on time.
- Participate in class discussions of case studies and assigned readings.
- Respond to general and individual questions based on assigned readings.
- Raise questions that make your classmates think.
- Provide imaginative yet relevant analysis of a situation.
- Contribute background or a perspective on a classroom topic that enhances its discussion.
- Answer questions raised in class.
- Class presentation. We’ll have a list of extremely useful readings/cases on which students are highly encouraged to lead the discussions in class. Each team will also get an opportunity to present your group projects in class.

Emphasis is placed on the quality of your contribution, rather than merely on its frequency. A lack of preparation, negative classroom comments, or improper behavior (such as talking to each other, sleeping in the classroom or walking in and out of the class while the lecture is in process) can lower this grade. Regular attendance and participation are very important.
Grading

<table>
<thead>
<tr>
<th>Components</th>
<th>Grades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class Participation</td>
<td>10%</td>
</tr>
<tr>
<td>Assignments</td>
<td>40%</td>
</tr>
<tr>
<td>Assignment 1</td>
<td>10%</td>
</tr>
<tr>
<td>Assignment 2</td>
<td>10%</td>
</tr>
<tr>
<td>Assignment 3</td>
<td>10%</td>
</tr>
<tr>
<td>Assignment 4</td>
<td>10%</td>
</tr>
<tr>
<td>Comprehensive Quiz</td>
<td>20%</td>
</tr>
<tr>
<td>Final Project</td>
<td>30%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

At CMU Heinz we seek to teach challenging courses that allow students to demonstrate their mastery of the subject matter. In general, students can expect a grading distribution where:
- 25-35% of students can expect to receive A’s for excellent work
- 50-70% of students can expect to receive B’s for good or very good work
- 5-15% of students can expect to receive C’s or less for adequate or below work

Note that while the School uses these ranges as a guide, the actual distribution for this course and your own grade will depend upon how well you actually perform in this course.

Re-Grading

The process of assigning grades is intended to be one of unbiased evaluation. Students are encouraged to respect the integrity and authority of the professor’s grading system and are discouraged from pursuing arbitrary challenges to it. If you believe that the grade you received was unjustified, you can appeal the grade. To appeal the grade you must write a one page explanation as to the reason for your appeal and hand it along with your graded assignment back to your TA. Please think twice before appealing a grade: the TA will completely re-grade the assignment, which may increase or lower your grade (if the TA caught more mistakes second time around). If again you consider that your grade was unjustified, you can appeal the grade with the instructor.

Individual Consultation

I encourage you to meet with the TA or with the instructor at any point during the semester to discuss your progress or any problems with the material or the assignments. I would prefer if you could come during the office hours, but if you cannot, please see the instructor after class, or send an e-mail, and we can schedule a time to meet. In the event that you feel the need to email us directly with a question, please make sure to put BIDM.SAS at the beginning of the subject line, so that we recognize that the email is from one of you, and so that our spam filters do not accidentally delete your message. Please talk to us if you have questions or problems. We are here to help.

Academic Honor Code

Students are subject to Carnegie Mellon University’s policies on academic integrity. Plagiarism is a serious offense and can result in failing the course and other disciplinary action. Plagiarism includes, but is not limited to:
- Presenting another writer’s work as your own;
- Cutting and pasting content verbatim without using quotation marks to indicate a direct quote or paraphrasing content without citing the source in-text using parenthetical references, footnotes, or endnotes in addition to listing each source on the Works Cited, References, or Notes page in a manner consistent with the format detailed in an approved style guide;
- Providing incomplete or incorrect information about the source cited.
Any attempt to represent the work of others as your own will be considered plagiarism and will be referred to the CMU Discipline Committee. Penalties determined by this committee range from academic probation to expulsion. It is in your best interest to submit nothing or a partial assignment, rather than an assignment copied in violation of the honor code.

**Blackboard**

Much of the class information can be found at the Blackboard Portal. Students will use their CMU account to login to the course at: [http://www.cmu.edu/blackboard/](http://www.cmu.edu/blackboard/). The data which can be found there include the class schedule, assignments, the lecture slides and class news. This is also where you will access the Discussion Board. Posting comments on the discussion board will be counted towards class participation.

**Class Schedule**

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Topic</th>
<th>Work Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/12</td>
<td>Course Overview; Data mining theory &amp; methodology</td>
<td>Install SAS EM;</td>
</tr>
<tr>
<td></td>
<td>1/14</td>
<td>Translating business problems into data mining problems; BI &amp; DM overview; Model prediction &amp; assessment; BI case studies;</td>
<td>Get Familiar with SAS EM Environment;</td>
</tr>
<tr>
<td></td>
<td>1/19</td>
<td>SAS EM Analytics Scheme: SEMMA palette for BI.</td>
<td>SEMMA palette;</td>
</tr>
<tr>
<td></td>
<td>1/21</td>
<td>Lab 1: SEMMA</td>
<td>HW1</td>
</tr>
<tr>
<td>2</td>
<td>1/26</td>
<td>Data Exploration and Pattern Discovery: Developing intuition about data; Market Basket Analysis; Association Rules; Clustering and segmentation; Using SAS EM for pattern recognition.</td>
<td>HW2</td>
</tr>
<tr>
<td></td>
<td>1/28</td>
<td>Lab 2: Association Mining and Link Graph</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2/2</td>
<td>Predictive Modeling (1): Classification and predictive modeling; Decision trees as classification tools; Using SAS EM to build decision trees.</td>
<td>HW3</td>
</tr>
<tr>
<td></td>
<td>2/4</td>
<td>Lab 3: Predictive Modeling (Decision Trees, Model Comparison)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2/9</td>
<td>Predictive Modeling (2): Regression Models (Linear &amp; Logistic); Using SAS EM to build regression models.</td>
<td>HW4</td>
</tr>
<tr>
<td></td>
<td>2/11</td>
<td>Lab 4: Predictive Modeling (Regression, Variable Selection, Missing Value, Transforming Variable)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2/16</td>
<td>Comprehensive Quiz: Advanced Topics for BI: Social Media Analytics &amp; Recommendation Systems; Geo-Retailing &amp; Mobile Advertising;</td>
<td>Quiz</td>
</tr>
<tr>
<td></td>
<td>2/18</td>
<td>Guest Lecture</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2/23</td>
<td>Final Presentation</td>
<td>Final Presentation</td>
</tr>
<tr>
<td>7</td>
<td>2/25</td>
<td>Final Report Due</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Exam</td>
<td>No Final Exam</td>
<td></td>
</tr>
</tbody>
</table>

Course Syllabus