**University of Houston – C.T. Bauer College of Business**

**Department of Decision and Information Sciences**

# MIS 7397/4397 – Predictive Analytics & Business Intelligence

**Fall 2016 – Wed 6:00 to 9:00pm**

Instructor: Dr. Norm Johnson
Office: 280C Melcher
Office hours: W - 5:00-6:00pm;

Email: njohnson@bauer.uh.edu
Office phone: 713-743-4164

**Required Text: Data Mining for Business Analytics: Concepts, Techniques, and Applications** with JMP PRO- by Galit Shmueli; Peter C. Bruce, Mia L. Stephens, & Nitin R. Patel; Wiley Pub.

***Prerequisite: STAT 6360/3331, BZAN 6310 or 6320 (with a grade of B+ or better)***

***Predictive Analytics*** describe the uses of a variety of techniques to analyze data so as to make predictions about future events or to discover meaningful patterns and rules. These techniques are related to ***Data Mining,*** which is the process of extracting useful information from large data sets. As a result of the technologies they use, companies now amass such data sets. For example, every day Wal-Mart uploads about 20 million point-of-sale transactions to a massively parallel system with 483 processors running a centralize database. There is much for companies to learn from the data they amass so as to improve knowledge of customers and markets. There is also an urgent need for people with skills to support companies in their efforts and to help them make good data-driven decisions. Some common business questions that one might address with Analytics and Data Mining are: (1) From a large list of prospective customers, which are most likely to respond to ad campaigns? (2) Which customers are most likely to default on payments? (3) Which customers are most likely to abandon a subscription service? This course covers a variety of techniques, from simple to complex, that are used in Analytics and Data Mining. There is a heavy emphasis throughout the course on analytical methods. The methods are taught by an example-based approach. In this regard, emphasis is placed on realistic business problems. There is a set of three software packages* that will be used in this course – ***JMP Pro from SAS and SAS Enterprise Guide.**

**Copies of these software applications will be provided by the instructor.*

## Teaching Methods:

1. Lectures and Discussions: Important material from the text and outside sources will be covered in class. This coverage will be mostly in the form of hands-on problem solving. Students will work together in teams. Students should also plan to take careful notes on topics that are presented by the instructor.

2. Assignments: Problems and readings might be periodically assigned to help support and supplement material found in the text. In particular, end of chapter questions could be assigned and, in such event, must be turned in on time.

3. Exams: Exams/Quizzes will be closed book/note and will test assigned readings and material discussed in class. Review sheets will be provided prior to the exam day. The final exam will not be cumulative, but may include some ideas, which are relevant to the topics for the final exam, that were previously covered. All relevant ideas will be noted on exam review sheets.

4. ***NOTES ONLINE - Blackboard:*** Most material will be distributed on the Internet, using the **Blackboard** application. It is assumed that students know how to access the content on Blackboard.

The lecture slides will be available in the CLASS NOTES section of the class web site. The slides will be there at least one day prior to each class. But, on some occasions, I might make a few minor changes to them **up to one hour before class begins**. If you want to have the slides with you during the class, please print them out and bring them with you. Printouts will not be available in class.

5. **Announcements** regarding the class such as schedule changes, assignments, projects, and so on will be made in class during the first 10 minutes as well as on the web at the Announcements page. You are responsible for being in class ON TIME to hear the announcements and for checking the class web site for announcements regularly.

6. **Contacting the Professor:** You can reach me by telephone or email. If you try to reach me, and you are unable to do so, then leave a message for me. I will try to get back to you within 48 hours.

7. **Grading:**

   1. **Quiz #1 – 20%**
   2. **Quiz #2 – 25%**
   3. **Class Exercises & Home Work Questions: 15%**
   4. **Project – 40%**

   Final course letter grade follows the numeric-letter grade system used at University of Houston.

8. **Project Description**

   The class project will give you an opportunity to apply most of the techniques you learn in this course. Around the sixth week of class, you will be given a set of data on which your project is based. The data will be what we consider "big." You will be required to get an understanding of the variables in your data set. You will use the software applications for the course to perform various predictive analytic techniques on your data. In your project, you will attempt to show how well you can apply the techniques, ***define multiple problems and solve them using appropriate techniques.*** **You will need to** ***interpret all your results. And, say how your results help a stakeholder to make good business decisions, which you will also define***.

**Course Policies:**

Missed Classes: The student is responsible for obtaining material, which may have been distributed on class days when he/she was absent. This can be done through contacting a classmate who was present or by contacting the instructor during his office hours or other times. Missed or late exams cannot be made up under any circumstances, unless an official excuse is provided. **Any uncoordinated, unexcused missed exam will result in a score of 0 for that exam.**

Assignments: All assignments are due at the beginning of class on the date due. As you know, this class has lab sections. In these sections, you will do exercises that correspond to topics already covered. As a result, this course outline is also a guide for the lab sections.

Academic Dishonesty: Plagiarism and cheating are serious offenses and may be punished by failure on exam, paper or project; failure in course; and or expulsion from the University. For more information, refer to the "Academic Dishonesty" policy in the University's Catalog. The University of Houston Academic Honesty Policy is strictly enforced by the C. T. Bauer College of Business. No violations of this policy will be tolerated in this course. A discussion of the policy is included in the University of Houston Student Handbook, http://www.uh.edu/dos/hdbk/acad/achonpol.html. Students are expected to be familiar with this policy.

Need for Assistance: If you have any condition, such as a physical or learning disability, which will make it difficult for you to carry out the work as outlined in this document, or which will require academic accommodations, please notify me as soon as possible. I will recommend that you contact the Center for Students with Disabilities. The contact person is Justin Dart in the CSD building #568, room 110. The numbers for the CSD office are Ph: 713-743-5400; TDD: 713-749-1527; Fax: 713-743-5396 or email: uhcsd@uh.edu.

## Tentative Lecture Outline

This outline is tentative. We will go over the topics described below in an applied way. That is, in solving problems our focus will be on the topic of the day. But keep in mind that the order in which these topics are covered can change in the event of unforeseen class disruptions.

| Date | No. | Topic | Chapter |
|---|---|---|---|
| | | | |
| 8/24 | 1 | Introduction to Course | Chap One |
| | | | |
| 8/31 | 2 | The Data Mining Process | Chap Two |
| | | Core Ideas | |
| | | Supervised & Unsupervised Learning | |
| | | Steps in Data mining | |
| | | Model Building | |
| | | Using JMP Pro | |
| | | | |
| 9/07 | 3 | Data Management – Queries, Tables, & Data Visualization | Chap Three |
| | | | |
| 9/14 | 4 | Dimension Reduction | Chap Four |
| | | | |
| 9/21 | 5 | Classification and Predictive Performance | Chap Five |
| 9/28 | 6 | Multiple Regression | Chap Six |
| | | ***Harvard Business Case- Harrah's Entertainment*** | |
| | | | |
| 10/05 | 7 | **Quiz #1** | |
| | | | |
| 10/12 | 8 | K-Nearest Neighbor | Chap Seven |
| | | | |
| 10/19 | 9 | Classification and Regression Trees | Chap Nine |
| | | | |
| 10/26 | 10 | Logistics Regression | Chap Ten |
| | | ***Harvard Business Case- Pilgram Bank*** | |
| | | | |
| 11/02 | 11 | Neural Nets | Chap Eleven |
| | | Discriminant Analysis | Chap Twelve |
| | | | |
| 11/09 | 12 | Association Rules | Chap Thirteen |
| | | | |
| 11/16 | 13 | Cluster Analysis | Chap Fourteen |
| | | ***Predictive Analytics & Business Strategy*** | |
| 11/30 | 14 | **Quiz #2 & Final Project Due** | |