MIS2502.001 – Data Analytics

**Fall 2014 (CRN 19561)**

About the Instructor:

Kartik Ganju (Kartik.Ganju@temple.edu)
201F Speakman Hall
http://community.mis.temple.edu/kartikkganju/

Phone: 215-204-5617

Office hours: 11:00 – 12:00 Tuesday and Thursday (at Main Campus)

Class Location and Time:

Alter Hall 232 12:30 – 13:50, Tuesday and Thursday

On the web: <http://community.mis.temple.edu/mis2502sec001f14>

Prerequisites:

Grade of C or better in MIS2101.

Course Description:

The course provides a foundation for designing database systems and analyzing business data to enhance firm competitiveness. Concepts introduced in this course aim to develop an understanding of the different types of business data, various analytical approaches, and application of these approaches to solve business problems. Students will have hands-on experience with current, cutting-edge tools such as MySQL and SAS Enterprise Miner.

Course Objectives:

* Articulate the key components of an organizations’ information infrastructure.
* Create data models based on business rules.
* Create a transactional database from a model using SQL.
* Create an analytical data store by extracting relevant data from a transactional database.
* Perform extract, transform, load (ETL) functions such as data sourcing, pre-processing, and cleansing.
* Discover trends in analytical data stores using the data mining techniques of clustering, segmentation, association, and decision trees.
* Present data visually for clear communication to a managerial audience.

Required Textbook:

There is no required textbook for this course.

**Evaluation and Grading**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|

|  |  |
| --- | --- |
| Item | **Percentage** |
| Exams (3)  | 70% |
| Assignments (10) | 25% |
| Participation | 5% |
|  |  |
|  |  |
|  |  |

 |

|  |
| --- |
| **Scale** |
| 94 – 100 | A | 73 – 76 | C |
| 90 – 93 | A- | 70 – 72 | C- |
| 87 – 89 | B+ | 67 – 69 | D+ |
| 83 – 86 | B | 63 – 66 | D |
| 80 – 82 | B- | 60 – 62 | D- |
| 77 – 79 | C+ | Below 60 | F |

 |

Exams

### There will be three exams during the semester. The date of the first exam is September 30, 2014 and the date of the second exam is October 28, 2014. The final exam is scheduled for Thursday December 11, 2014 from 10:30 to 12:30. Missed exams cannot be made up, regardless of the reason for absence.

Late Assignment Policy

An assignment is considered late if it is turned in after the beginning of class. No late homework assignments will be accepted without penalty. All assignments will be assessed a 20% penalty (subtracted from that assignment’s score) for each of the **first two calendar days** they are late. **No credit will be given for assignments turned in more than two calendar days past the due date.**

However, you must submit all assignments, even if no credit is given. If you skip an assignment, an additional 10 points will be subtracted from your **final grade** in the course. For example, if you do not turn in an exercise, you will receive no credit for that assignment plus a 10 point penalty, reducing your maximum grade in the course to an 87.5.

***Equipment failure is not an acceptable reason for turning in an assignment late****.*

Assignments

There will be ten assignments. They are to be done individually and should represent your own work. If you need help, you may consult with your instructor or your ITA.

|  |  |  |
| --- | --- | --- |
| **#** | **Assignment** | **Due** |
| 1 | Identifying Entities  | September 4, 2014 |
| 2 | ER Modeling | September 11, 2014 |
| 3 | SQL #1 – Getting Data out of the Database | September 23, 2014 |
| 4 | SQL #2 – Putting Data into the Database | October 7, 2014 |
| 5 | Pivot Tables in Excel | October 21, 2014 |
| 6 | Data Analytics Challenge | October 30, 2014 |
| 7 | SAS #1 – Introduction to working with SAS | November 6, 2014 |
| 8 | SAS #2 – Decision Trees | November 13, 2014 |
| 9 | SAS #3 – Clustering | November 20, 2014 |
| 10 | SAS #4 – Association Rules | December 4, 2014 |

Classroom Etiquette

The environment you and your fellow students create in class directly impacts the value gained from the course. To that end, the following are my expectation of your conduct in this class:

* Arrive on time and stay until the end of class.
* Turn off cell phones, pagers and alarms while in class.
* Limit the use of electronic devices (e.g., laptop, tablet computer) to class-related usage such as taking notes. Restrict the use of an Internet connection (e.g., checking email, Internet browsing, sending instant messages) to before class, during class breaks, or after class.
* During class time speak to the entire class (or breakout group) and let each person “take their turn.”
* Be fully present and remain present for the entirety of each class meeting.

Participation

Participation will be evaluated in two ways. First, a question will be posted to the Community Site each week about some aspect of the material we have just covered. Leave an answer to the question as a comment. You can also respond to other students’ comments, as long as you also add your own insight to the discussion. You are expected to contribute something to each week’s discussion.

Second, involvement during class is also important. Being present in class to ask and answer questions is essential to the learning process. While you’re not expected to say something in every class meeting, simply showing up for class does not qualify as participation.

Plagiarism and Academic Dishonesty

Plagiarism and academic dishonesty can take many forms. The most obvious is copying from another student’s exam, but the following are also forms of this:

* Copying material directly, word-for-word, from a source (including the Internet)
* Using material from a source without a proper citation
* Turning in an assignment from a previous semester as if it were your own
* Having someone else complete your homework or project and submitting it as if it were your own
* Using material from another student’s assignment in your own assignment

If you use text, figures, and data in reports that were created by someone other than yourself, you must identify the source and clearly differentiate your work from the material that you are referencing. There are many different acceptable formats that you can use to cite the work of others (see some of the resources below). You must clearly show the reader what is your work and what is a reference to somebody else’s work.

Plagiarism and cheating are serious offenses. Penalties for such actions are given at my discretion, and can range from a failing grade for the individual assignment, to a failing grade for the entire course, to expulsion from the program.

Student and Faculty Academic Rights and Responsibilities

The University has adopted a policy on Student and Faculty Academic Rights and Responsibilities (Policy # 03.70.02) which can be accessed through the following link:
<http://policies.temple.edu/getdoc.asp?policy_no=03.70.02>

**Schedule***(Keep in mind that all dates are tentative – check the Community site regularly for changes in the schedule!)*

You are expected to review the assigned material for each class. Additional, supplementary material may be assigned throughout the course of the semester.

|  |  |  |  |
| --- | --- | --- | --- |
| **Day** | **Topics** | **Course Materials** | **Assignments** |
| **Week 1** |
| Aug. 26 | Course Introduction and SyllabusThe Things You Can Do with DataThe Information Architecture of an Organization | PowerPoint: The Things You Can Do With Data PowerPoint: Information Architecture |  |
| Aug. 28 | Data Modeling Gathering requirementsIntroducing The Entity-Relationship Diagram**In-class exercise: Identifying entities** | PowerPoint: Relational Data Modeling |  |
| **Week 2** |
| Sept. 2 | More on ERDs:Relationships, cardinality**In-class exercise: Creating an entity relationship diagram** | PowerPoint: Relational Data Modeling |  |
| Sept. 4 | From ERDs to Schemas:Normalization, primary/foreign keys, joins**In-class exercise: Converting ERDs to schemas** | PowerPoint: Relational Data Modeling | Assignment 1 Due: Identifying Entities |
| **Week 3** |
| Sept. 9 | Getting data out of the database:SQL SELECT, DISTINCT MIN, MAX, COUNT, and WHERE*Make sure you’ve done the MySQL tutorial and reviewed the MySQL PowerPoint deck.***In-class exercise: Pen-and-paper SQL exercise** | PowerPoint: SQL 1 |  |
| Sept. 11 | Getting data out of the database:Joining tables, SQL subselects, LIMIT | PowerPoint: SQL 1 | Assignment 2 Due: ER Modeling |

|  |
| --- |
| **Week 4** |
| Sept. 16 | **In-class exercise: Working with SQL, part 1** |  |  |
| Sept. 18 | Creating and updating the databaseSQL CREATE, DROP, and ALTERSQL INSERT, UPDATE, and DELETE | PowerPoint: SQL 2 |  |
| **Week 5** |
| Sept. 23 | **In-class exercise: Working with SQL, part 2** |  | Assignment 3 Due: SQL #1 |
| Sept. 25 | Review for Exam 1 |  |  |
| **Week 6** |
| Sept. 30 | **Exam 1** |  |  |
| Oct. 2 | Principles of Data Visualization | PowerPoint: Principles of Data Visualization |  |
| **Week 7** |
| Oct. 7 | **In-class exercise: Data Visualization** |  | Assignment 4 Due: SQL #2 |
| Oct. 9 | Turning transaction data into analytical data: Overview of the Dimensional Model The structure of the Dimensional Model: The Star Schema  | PowerPoint: Dimensional Data Modeling |  |
| **Week 8** |
| Oct. 14 | Working with Dimensional Data: Pivot Tables in Excel **In-class exercise: Pivot Tables in Excel** |  |  |
| Oct. 16 | Getting data into the warehouse and cube:The Extract, Transform, Load processData quality: Best practices, data cleansing, and integration | PowerPoint: ETL |  |
| **Week 9** |
| Oct. 21  | **In-class exercise: ETL** |  | Assignment 5 Due: Pivot Tables in Excel |
| Oct. 23 | Review for Exam 2 |  |  |

|  |
| --- |
| **Week 10** |
| Oct. 28 | **Exam 2** |  |  |
| Oct. 30 | Introduction to Advanced Analytics and SAS Enterprise Miner**In-class exercise: Descriptive Statistics Review**  | PowerPoint: Advanced Analytics – Introduction | Assignment 6 Due: Data Analytics Challenge |
| **Week 11** |
| Nov 4 | **In-class exercise: Introduction to SAS Enterprise Miner/Preparing Data for Analysis** |  |  |
| Nov 6 | Analysis Scenario: Determining customer behavior based on a profile (decision trees)**In-class exercise: Interpreting Decision Tree Output**  | PowerPoint: Classification using Decision Trees | Assignment 7: Intro to SAS |
| **Week 12** |
| Nov 11 | **In-class exercise: Decision trees in SAS Enterprise Miner**  |  |  |
| Nov 13 | Analysis Scenario: Identifying similar customers (clustering and segmentation) **In-class exercise: Interpreting Clustering Output**  | PowerPoint: Clustering and Segmentation |  |
| **Week 13** |
| Nov 18 | **In-class exercise: Clustering and Segmentation in SAS Enterprise Miner** |  |  |
| Nov 20 | Analysis Scenario: What products are purchased together? (Association Rules) **In-class exercise: Interpreting Association Rule Mining Output** | PowerPoint: Association Rule Mining | Assignment 8 Due: Decision Trees  |
| **Week 14** |
| Fall Break **Nov 25th: Assignment 9 Due: Clustering** |
| **Week 15** |
| Dec 2 | **In-class exercise: Association Rule Mining in SAS Enterprise Miner** |  |  |
| Dec 4 | Review for Final Exam |  | Assignment 10 Due: Association Rules |