MIS2502.012– Data Analytics

Summer 2017 (CRN 7631)

About the Instructor

Jaclyn M. Hansberry (jaclyn.hansberry@temple.edu)  
755 Wachman Hall

Phone: 215-204-1132

Office hours: 3:00 – 4:00 Tuesday and Thursday

Class Location and Time:

Alter Hall 232 4:40 – 7:30, Tuesday and Thursday

On the web: <http://community.mis.temple.edu/mis2502s012summer17/>

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 R.

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) | 60% | | Assignments (9) | 30% | | In-class activities | 5% | | 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. A tentative Exam schedule is as follows:

### Exam 1: May 30, 2017

### Exam 2: June 13, 2017

### Exam 3 (Final Exam): June 22, 2017

### While there is some natural overlap in material between the exams, the exams are not intended to be cumulative.

### Make-up exams will not be given under most circumstances. Exceptions are granted at the instructor’s discretion and are typically limited to extreme circumstances such as documented hospitalization.

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.**

**Please note:**

* Equipment failure is not an acceptable reason for a late submission.
* For the assignment to be considered “on time,” you must send the submission to the correct email address by the due date. If I don’t receive the email because you sent it to the wrong address, the usual late penalty applies.
* For the assignment to be considered “on time,” you must attach all necessary files specified in the assignment instructions by the due date. For any revisions or additional documents received after the due date, the usual late penalty applies.

**Assignments:**

There will be nine assignments. Each assignment is to be done individually and should represent your own work. You may not work together or share answers! If you need help, consult your instructor for the course.

|  |  |
| --- | --- |
| **#** | **Assignment** |
| 1 | ER Modeling |
| 2 | SQL #1 – Getting Data out of the Database |
| 3 | SQL #2 – Putting Data into the Database |
| 4 | ETL in Excel |
| 5 | Pivot Tables in Excel |
| 6 | Introduction to working with R |
| 7 | Decision Trees |
| 8 | Clustering |
| 9 | Association Rules |

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.

Class Presence and 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 | | | |
| May 16 | Course Introduction and Syllabus  The Things You Can Do with Data.  The Information Architecture of an Organization.  Data Modeling  Gathering requirements Introducing The Entity-Relationship Diagram  **In-class exercise: Identifying entities** | PowerPoints:  The Introduction and Information Architecture  Relational Data Modeling |  |
| May 18 | More on ERDs: Relationships, cardinality  From ERDs to Schemas:  Normalization, primary/foreign keys, joins  **In-class exercise: creating an entity relationship diagram**  **In-class exercise: Converting ERDs to schemas** | PowerPoint: Relational Data Modeling |  |
| Week 2 | | | |
| May 23 | Getting data out of the database:  SQL SELECT, DISTINCT MIN, MAX, COUNT, and WHERE  Getting data out of the database:  Joining tables, SQL subselects, LIMIT  *Make sure you’ve reviewed the guide for setting up a connection in MySQL Workbench and reviewed the MySQL PowerPoint deck.*  **In-class exercise: Pen-and-paper SQL exercise**  **In-class exercise: Working with SQL, part 1** | PowerPoint: SQL 1 |  |
| May 25 | Creating and updating the database  SQL CREATE, DROP, and ALTER  SQL INSERT, UPDATE, and DELETE  **In-class exercise: Working with SQL, part 2** | PowerPoint: SQL 2 | Assignment 1 Due: ER Modeling |
| Week 3 | | | |
| May 30 | **EXAM 1** |  | Assignment 2 Due: SQL #1 |
| June 1 | Principles of Data Visualization  Getting data into the analytical database: The Extract, Transform, Load process  **In-class exercise: Data Visualization** | PowerPoint: Data Visualization  PowerPoint: ETL | Assignment 3 Due: SQL 2 |

|  |  |  |  |
| --- | --- | --- | --- |
| Week 4 | | | |
| June 6 | Turning transaction data into analytical data: Overview of the Dimensional Model  The structure of the Dimensional Model: The Star Schema  Working with Dimensional Data: Pivot Tables in Excel  **In-class exercise: Pivot Tables in Excel** | PowerPoint: Dimensional Data Modeling | Assignment 4  Due: ETL |
| June 8 | Working with Dimensional Data: Pivot Tables in Excel  Introduction to Advanced Analytics and R  **In-class Exercise: Descriptive Statistics Review & getting familiar with R and RStudio** | PowerPoint: Advanced Analytics and R | Assignment 5 Due: Pivot Tables in Excel |
| Week 5 | | | |
| June 13 | **Exam 2 Due**\*\*  Analysis Scenario: Determining customer behavior based on a profile (decision trees)  Analysis Scenario: Identifying similar customers (clustering and segmentation)  **In-class exercise: Chi-Squared Statistics**  **In-class exercise: Decision trees in R** | PowerPoint: Classification using Decision Trees  PowerPoint: Clustering and Segmentation | Assignment 6 Due: Intro to R |
| June 15 | Analysis Scenario: What products are purchased together? (Association Rules)  **In class exercise: Clustering in R** | PowerPoint: Association Rule Mining | Assignment 7 Due: Decision Trees in R |
| Week 6 | | | |
| June 20 | Connecting to a MySQL Database using R  **In-class exercise: Computing Confidence, Support, and Lift**  **In-class exercise: Association Rule Mining in R** |  | Assignment 8 Due: Clustering in R |
| June 22 | **Exam 3** |  | Assignment 9 Due: Association Rules |