MIS2502: Exam 2 Study Guide (Spring 2019)

Instructor: Zhe Deng

The exam will be a combination of multiple-choice and short-answer questions. It is a closed-book, closed-notes exam. You will not be able to use a computer during the exam.

The following is a list of items that you should review in preparation for the exam. Note that *not every item on this list may be on the exam, and there may be items on the exam not on this list.*

**SQL Out: Advanced Queries (LIMIT, Joins, Subselects)**

* Given the schema of a database, be able to create the SQL statements that
	+ Limit the number of results to be returned
	+ Require a join of multiple tables
	+ Contain a Subselect statement
	(i.e., determine the customers with the highest sales)

**SQL In (CREATE, ALTER, INSERT, UPDATE, and DELETE)**

* Given the schema of a database, be able to create the SQL statements that
	+ Create a table based on a list of its metadata/schema using CREATE TABLE
	+ Change the structure of a table using ALTER TABLE
	+ Delete a table using DROP TABLE
	+ Add a record to a table using INSERT
	+ Update an existing record in a table using UPDATE
	+ Delete a record from a table using DELETE
* Be familiar with using WHERE conditional statements in the UPDATE and DELETE statements
* Be familiar with MySQL data types (INT, DECIMAL, BOOLEAN, DATE/DATETIME, VARCHAR, etc.)
* Know how to specify primary keys and foreign keys in CREATE TABLE statements
* Identify how to add records to a table created from a many-to-many relationship so that the new record associates two existing records in the associated tables
(i.e., add a record to a film-actor table that associates a particular film with a particular actor)

**ETL**

* What is it? Why is it important?
* Explain the purpose of each component (Extract, Transform, Load) in combining data sets
* The problems with using data from legacy systems
* How do inconsistencies in data get resolved?
* Tableau Prep
	+ How does it perform the extract, transform, and load operations?
	+ Major operations – splitting values, joining tables, calculated fields, grouping data

**Semi-Structured Data**

* What is semi-structured data? Examples? What does it mean to have no formal data model?
* What is unstructured data? Examples?
* Compare csv, XML, and JSON data formats and explain advantages/disadvantages of each
* Construct a csv, XML, and JSON data file from raw data
* Construct a structured data table from a csv, XML, or JSON file
* Explain how applications exchange data using semi-structured data

**And don’t forget…**

* ERDs, including cardinality, entities, and attributes
* Basic SQL, including GROUP BY, ORDER BY, WHERE, and functions
* The difference between a transactional database and an analytical data store

**\*\*Advanced Analytics and R will be covered in the final exam.**