

# Northeastern University

## College of Professional Studies

### ALY 6030 Data Warehousing and SQL Module 1 Assignment

---

#### Class ALY6030 – Data Warehousing and SQL

#### Module 1 Assignment: Data Modeling and Basic SQL

##### Overview and Rationale

In this assignment you will install required tools and perform basic data modeling using a flat file. At least three business questions will need to be defined and then answered using basic SQL commands.

##### Assignment Instructions:

##### Part 1: Install MySQL Server

- Install MySQL Server on your computer.
- Document every step.
- Include screen shots of **each** step and **comments for each** screen shot explaining what it is.

##### Part 2: Install MySQL Workbench

- Install MySQL Workbench on your computer.
- Document every step.
- Include screen shots of **each** step and **comments for each** screen shot explaining what that is.

##### Part 3: Data Modeling and SQL

- Choose **one** of the following five (CSV) data files provided in the Module 1 Introduction section:
  - Sacramento: 985 real estate transactions over a 5-day period.
  - Cars: various makes of cars and their attributes.
  - EBay: 5-day auctions for Xbox consoles on EBay.
  - Movies: Films, with leading actor/actress and other attributes.
  - Airport: Traffic (no. of flights) for various US airports in one month.
  - TechCrunch: Start-up companies raising funds events.
- For dataset of your choice
  - Manually create new schema (database) in MySQL. Provide SQL script and screen shots with results.

# Northeastern University

## College of Professional Studies

### ALY 6030 Data Warehousing and SQL Module 1 Assignment

---

- Manually create new empty table according to the structure of data set you've selected. Provide SQL script and screen shot with results
- Import data from the data set flat file into MySQL using MySQL Workbench.
- You must insert the data into a pre-existing table you created.
- Provide SQL script and/or screen shots with each step and results

*Include into your paperwork the following information:*

- What are the columns (attributes)?
- What is the appropriate MySQL data type for each column? There may be more than one right answer! Explain your choices.
- Based on the nature of your particular dataset, propose three (or more!) meaningful business questions you can ask.
- Answer each question using SQL.
- Between your three questions you should cover each of the following language elements at least once:
  - WHERE (in combination with a comparison operator BETWEEN, (NOT) IN, IS (NOT), etc.)
  - ORDER BY
  - MAX, MIN, AVG, COUNT
  - GROUP BY,
  - HAVING
  - AS
- Show all SQL code and the results by taking a screen shots of your Workbench session.

#### **Submission requirements:**

- Submit paperwork using .docx file in Microsoft Word format named as ALY6030\_Assignment\_1\_*[your name]* .docx
- Submit SQL script .sql file in text format names as ALY\_6030\_Assignment\_1\_*[your name]*.sql

**Paperwork file must meet the following requirements:**

# Northeastern University

## College of Professional Studies

### ALY 6030 Data Warehousing and SQL Module 1 Assignment

---

- Title Page must have:
  - Class name
  - Class number
  - Class CRN number
  - Assignment name
  - Your full name as registered in Canvas
- Introduction section
  - explain what this assignment is about
  - set goals and expectations
  - explain the selected dataset
  - introduce tools and languages you plan to use
- Final conclusions section
  - Explain if goals and expectations were met or not
  - Explain cons and pros of using the tools, methods and techniques.
  - Explain what would you do differently
  - References (optional)

#### **SQL Script file must meet the following requirements:**

- Only ,sql text format will be accepted
- Each line of the SQL code must have a meaningful comment
- Code should be executed on other computers without any modifications (add extra steps in the comments if necessary)