**Omaha Data Science Academy (NECC)**

**Data Manipulation and Management**

Instructor(s):  Jeremy Bergmann, BS/MS

Phone: 402-321-7530; Email:  [jeremy@jbanalyticsconsulting.com](mailto:jeremy@jbanalyticsconsulting.com); Office: By Appointment

GitHub: <https://github.com/jbergmann56/OSDA-DataManipulationAndManagement>

**COURSE DESCRIPTION:**

This class will teach the student how to design, store, clean, query, and access data, specifically to be used within data science projects.  It will also teach database design, SQL programming, Relational Database modeling & schemas, Python & NoSQL, data extract, cleaning/transformation and load (ETL) techniques, and cloud-based platform offerings.    
  
Additionally, this class will include a “class project”, where students create an end-product, using raw data, a MySQL database and BI Visualization tool (such as Tableau). The class project will build from knowledge gained in the course and provides the student a working example, similar to what one could expect within a business environment.

.

**Overall Approach:**

The overall approach to the course is to provide the background needed for data manipulation, with a focus on data science projects, using the SQL programming language and relational & non-relational databases. The topics include: Data Manipulation basics, writing SQL Queries, Manipulating/Managing data and leveraging tools to assist and automate the building and deployment of data-science projects.

**Class Schedule: (20 Hours – 5 Classes)  
Class 1 - Data Manipulation - Basics**

Hour 1 - Introductions, Class Overview, Data Science Philosophy, Tool Setup

Hour 2 – Lecture - Data Manipulation Background, Definitions & Database Basics

Hour 3 – Lecture - SQL Queries, Data Types, Functions & Sub-queries

Hour 4 - MySQL Workbench Examples, Class Project Overview & Exercises

**Class 2 – Manipulating Data**

Hour 1 - Class 1 Quiz, Lecture - Create Databases, Create/Alter/Delete Tables, Constraints & Views

Hour 2-3 - SQL Joins – Lecture & Working Examples

Hour 4 – Exercises & Class Project - Joining & Storing Data for Analysis

**Class 3 – Cleaning & Managing Data**

Hour 1-2 – Class 3 Quiz, Lecture – Enterprise Data Management, Database Admin, Data Gov. & Quality

Hour 2-3 – Lecture - Case Statements, SQL Operators & Logical Functions, Data Cleansing

Hour 4 - Exercises & Class Project - Cleaning Data for Analysis, Measuring Data Quality

**Class 4 – Structuring & Modeling Data**

Hour 1 - Class 2 Quiz, ER & Relational Diagrams, Data Models

Hour 2 - Normal Forms – Lecture

Hour 3-4 – Class Project - Data Modeling – Creating ER & Relational Diagrams using draw.io

**Class 5 - Tools for Data Manipulation and Management**

Hour 1 – Cloud Platforms (AWS, Azure & Google) – Lecture

Hour 2 - NoSql, ETL & Workflow Tools - Lecture

Hour 3-4 – Class Project: Data Visualization using MySQL - Working Session