**DATA CLEANING AND WRANGLING ASSIGNMENT**

1. Steps involved in data wrangling
   1. Data discovery involves getting familiar with a given dataset i.e exploring the data to understand it better.
   2. Data structuring involves restructuring the data to fit into a particular format e.g. changing from a csv to Jason format to be used for analysis, changing from text format to number format.
   3. Data cleaning involves removing incorrect, incorrectly formatted, duplicates and incomplete data from a dataset.
   4. Data enriching involves merging data together when the have a common description so as to be able to make informed decisions.
   5. Data Validation ensures the quality and accuracy of a given dataset.
   6. Data publishing involves storing the wrangled dataset in a folder where it can easily be accessed.
2. Data wrangling is important so as to be able to create tidy data to be used for analysis
   1. When raw data is accurately wrangled it ensures that quality data is passed down into the downstream process
   2. A wrangled data leads to accurate result which enhances decision making
   3. Data wrangling is a stage that prepares data for the data mining process

**Practicals using MySQL**

1. SELECT \*

FROM customers

WHERE points < 500;

1. SELECT \*

FROM customers

WHERE state = 'VA';

1. SELECT \*

FROM employees

UPDATE employees

SET salary = 100000

WHERE employee\_id = 37270;

SELECT \*

FROM employees

WHERE employee\_id = 37270;(updating using manual method)

1. SELECT \*

FROM employees;

DELETE FROM employees

WHERE employee\_id = 63196;