## Assignment 3 – Data Manipulation **Objectives**

Understand and learn how to use pandas properly.

**Datasets**

patient-data.csv

**Assignments**

Provide Python code using Pandas to read the above file and carry out exercise as per Preparation & Reporting sections. You may install any package as required

**Data Preparation:**

1. Provide New Column BMI-Value (BodyMassIndex) using the given formula   
   BMI = Weight (KGs) / Height2 (Meters)
2. Provide New Column BMI-Label based on BMI-Value as per follows

Underweight Less Than 18.50

Normal 18.50 to < 25.00

Overweight 25.00 to < 30.00

Obese 30.00 And Above

1. Review the Dataset For Errors In Data. Provide solution to rectify the error.
2. Review the Dataset For Missing Data. Provide solution to rectify the same.
3. Convert the HealthGrade of the records as per follows

Good Health 1

Normal 2

Bad Health 3

**Reporting:**

1. Display top 10 records based on BMI-Value.
2. Display bottom 10 records based on BMI-Value.
3. Provide frequency / counts of   
   Gender > Race
4. Provide max, min and average values for BMI-Values as per following   
   Race > Gender.
5. Display All Records for all people who are dead.
6. Display All Records for “Hispanic Females”
7. Provide 7 sample records from the Dataset. Use seed(707)

**Submission Guidelines**

1. You must submit the final version of your .py program as   
   “a2-StudentName.py.
2. Please ensure that your .py is well documented.
3. In case you make any assumptions, please note the assumptions at the start of the .py file.
4. Your .py file should be emailed, as an attachment.
5. The email must be sent by the end of day on which the assignment is due.