**SSIE 637 Advanced Topics in Healthcare**
**Home Assignment 01**
**Fall 2022**

Instructor: *Daehan Won*

**Due 09/23/2022**

**Question 1 (40%)**
Let practice data processing and visualizations. For each problem, you must provide not only the answers but also codes in your answer sheet.
Note: (i) You must implement data loading module in your codes, (ii) You can choose any programming language except EXCEL

**Q1-1 (10%)**. Load the data; 'data1.txt' and draw a histogram.

**Q1-2 (20%)**. In the 'data2.txt', there are two sets (rows) of measurement about anxiety scale for each subject. The first row represents the measured scale of 26 persons with a diagnosis of mental disorder and the second row represent the measured scale of 21 healthy persons. Construct a Box and Whisker plots using the data. Note that you should produce a side-by-side box plot that allow us to compare them directly.

**Q1-3 (10%)**. In Q1-2, state your conclusions briefly.

**Question 2 (60%)**
Let look at the other uploaded data set ('CardiacData.xls'), which was collected for measuring risk of heart attack. In the data, the first 21 columns (from A to U) include measurements based on the prior measurements and during the test. This is followed by specific outcomes, namely four types of cardiac events during the ensuing 12 months in columns V-Y. The next six columns (from Z to AE) list the history of medical events of the patient. The final column, AF, is representing whether the patient suffered a new cardiac event during the ensuing 12 months, i.e., whether any of the events in Columns V-Y occurred.

**Q2-1 (10%)**. Draw a histogram for the first column (column 'A'); "Basal Heart Rate".

**Q2-2 (20%)**. Draw a box plot of Baseline Cardiac Ejection Fraction (column 'O') for two groups based on the column 'T': in the first group (group 1), heart wall motion anomaly was observed, in the other group (group 2), no heart wall motion anomaly was observed.

**Q2-3 (10%)**. For each group in Q2-2, figure out calculated quartiles (Q1 and Q3), mean, median, maximum, and minimum.

**Q2-4 (20%)**. Create a cross table between two columns: Any New Cardiac Event (column 'AF') vs. Stress ECG Positive (column 'U'). Also, draw a bar chart of new cardiac events as a function of whether the stress ECG was positive.