# Report

## Q1)

- Part 1 - just copy the columns with a new name
- Part 2 - Use sklearn Standard Scaler
- PART 2 UPDATE. When using the standard scaler the test will throw an error so I switched to using the formula of:
    - X - Mean / Standard dev.
- Part 3 - Use sklearn Min Max Scaler
- Part 4 - use pandas built in Mean function but on the axis 1
- Part 5 - Save

## Q2)

- Part 1
    - To reduce the number of attributes we first have to perform a PCA
        - PCA is a long process involving calculating covariances between all values and then getting the eigenvectors associated with the covariance matrix.
    - Step-1: perform PCA
    - Step-2: check to see at which Principal Components sum to give us a variance 95% or greater
    - Step-3: Transform your data using the Principal Components
- Part 2
    - Use sklearns built in discritize with Bins to discritize the data using the equal width method
    - Put the new data into the Original Data with a simple for loop
- Part 3
    - Use sklearns built in discritize with Bins to discritize the data using the equal frequency method
    - Put the new data into the Original Data with a simple for loop