**Principles of Data Science Assignment 2**

Name: Lohith Reddy Poreddy

Student ID: 16337814

Results:

1. Mean values comparison for sample and population data.

Chart, bar chart

Description automatically generated

The results from the above bar plot clearly show that the mean of the sample and population glucose feature is almost the same.

Highest value comparison of sample and population data

Chart, bar chart

Description automatically generated

From the plot above the highest value for the sample is around 175 - 200 and for the population it is almost 200. It might be almost close for this sample to reach the population results, but that can’t be true with every sample as the samples change the values might change too.

1. Comparison of 98th percentile value in sample and population dataset.

Chart, bar chart

Description automatically generated

The results from the bar plot above show that 98th percentile of both the sample and population data doesn’t have much of a difference.

1. Bootstrap comparison results for mean, std, percentile of 500 samples and population data

Chart, histogram

Description automatically generated

Chart, histogram

Description automatically generated

Chart, bar chart

Description automatically generated

As we can see from the results that the mean, standard deviation and percentile of sample data and the population data almost close or same. From this we can conclude that for this data set working with the sample data and the whole data would yield in almost similar results.