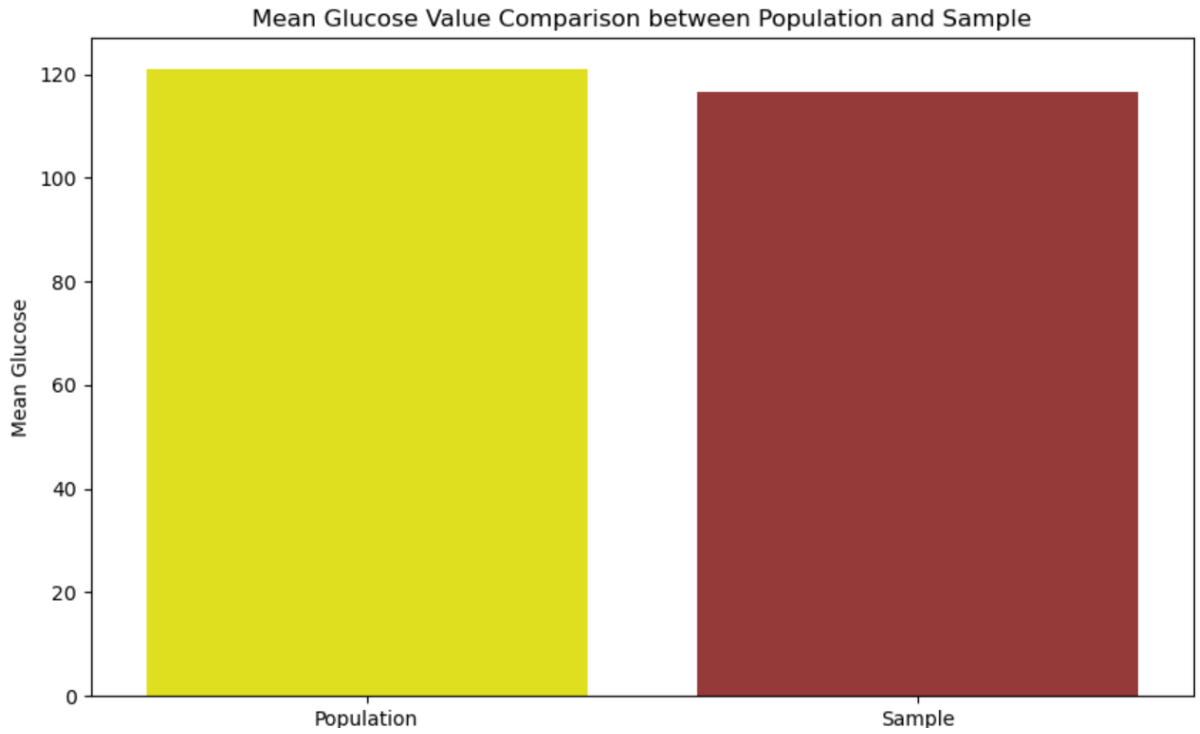
Assignment 3 and 4

Name : Bolisetty Siddharth

Id:16354137

2 a)

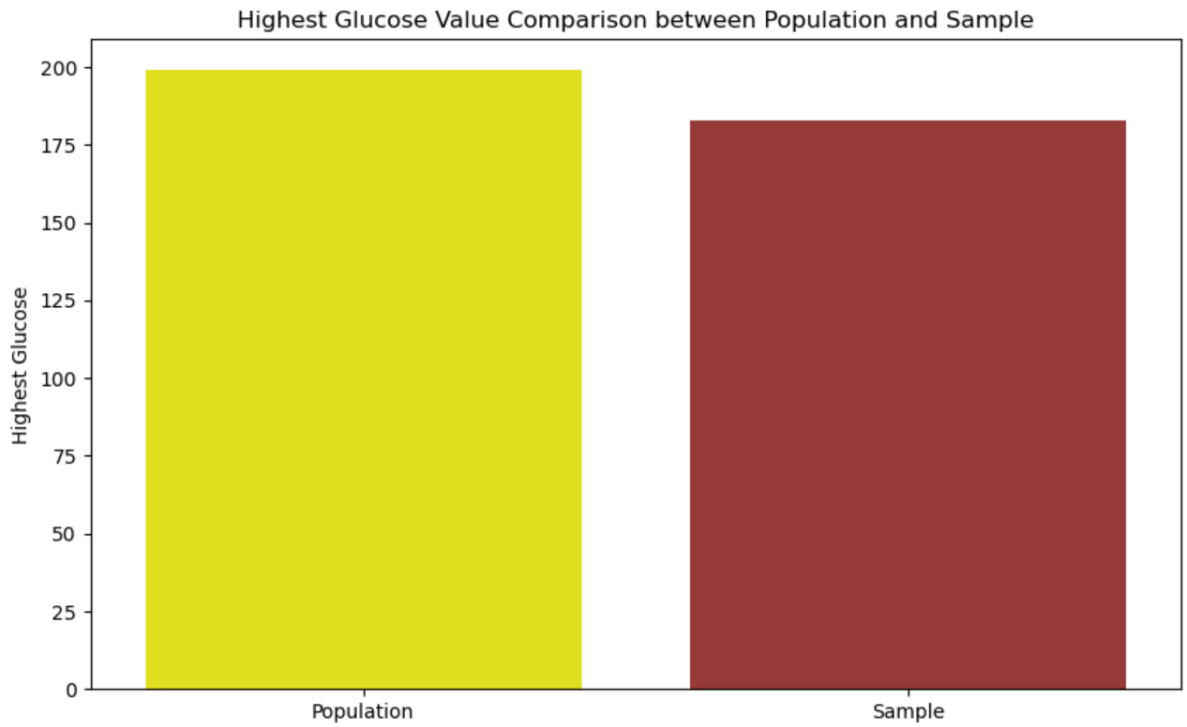


Mean Glucose Value Comparison:

Population Mean Glucose: This represents the average glucose level across all 768 patients in the dataset.

Sample Mean Glucose: This represents the average glucose level in the randomly sampled subset of 25 patients from the population.

The bar plot illustrates the comparison between the mean glucose values of the population and the sample. From the plot, it seems that the sample mean glucose value is higher than the population mean glucose value. This could indicate that the sample might have been randomly drawn from a subset of the population with higher glucose levels, or it could just be due to random chance.



Highest Glucose Value Comparison:

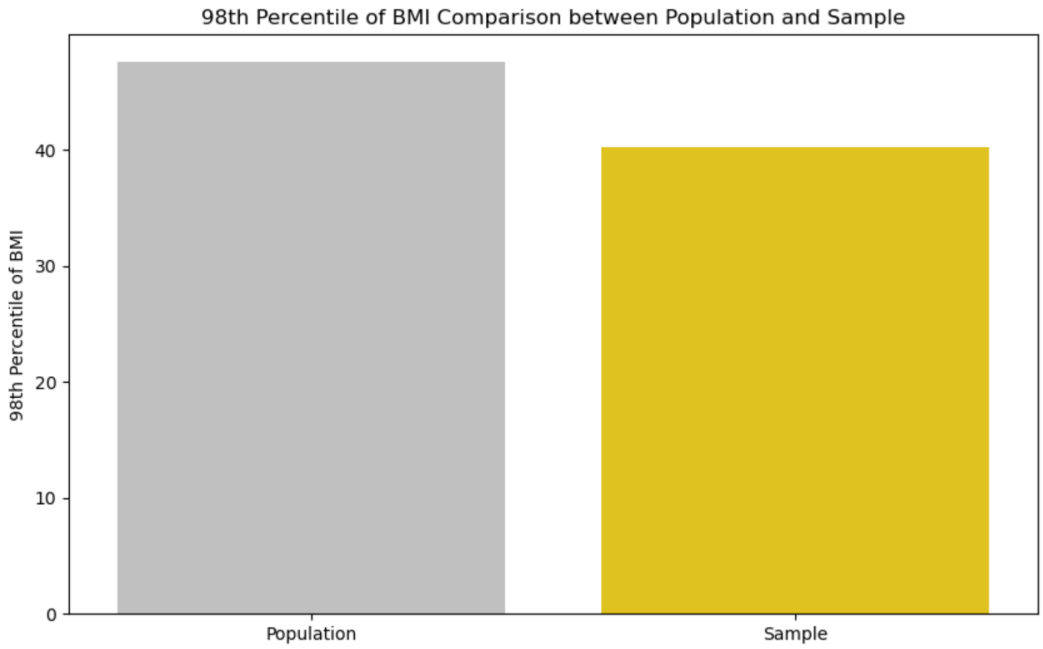
Population Maximum Glucose: This represents the highest recorded glucose level among all 768 patients in the dataset.

Sample Maximum Glucose: This represents the highest recorded glucose level in the randomly sampled subset of 25 patients from the population.

The bar plot compares the highest glucose values between the population and the sample. In this case, it seems that the highest glucose value in the sample is lower than the highest glucose value in the population. Again, this could be due to random chance or the specific subset of patients included in the sample.

Overall, these visualizations provide a comparison between the population and the sample in terms of mean and highest glucose values, offering insights into the characteristics of the sampled subset compared to the entire population.

2 b)



The graph compares the 98th percentile of BMI (Body Mass Index) between the population and a randomly sampled subset. Here are the findings based on the provided code:

Population 98th Percentile of BMI: This value represents the BMI at the 98th percentile across all 768 patients in the dataset.

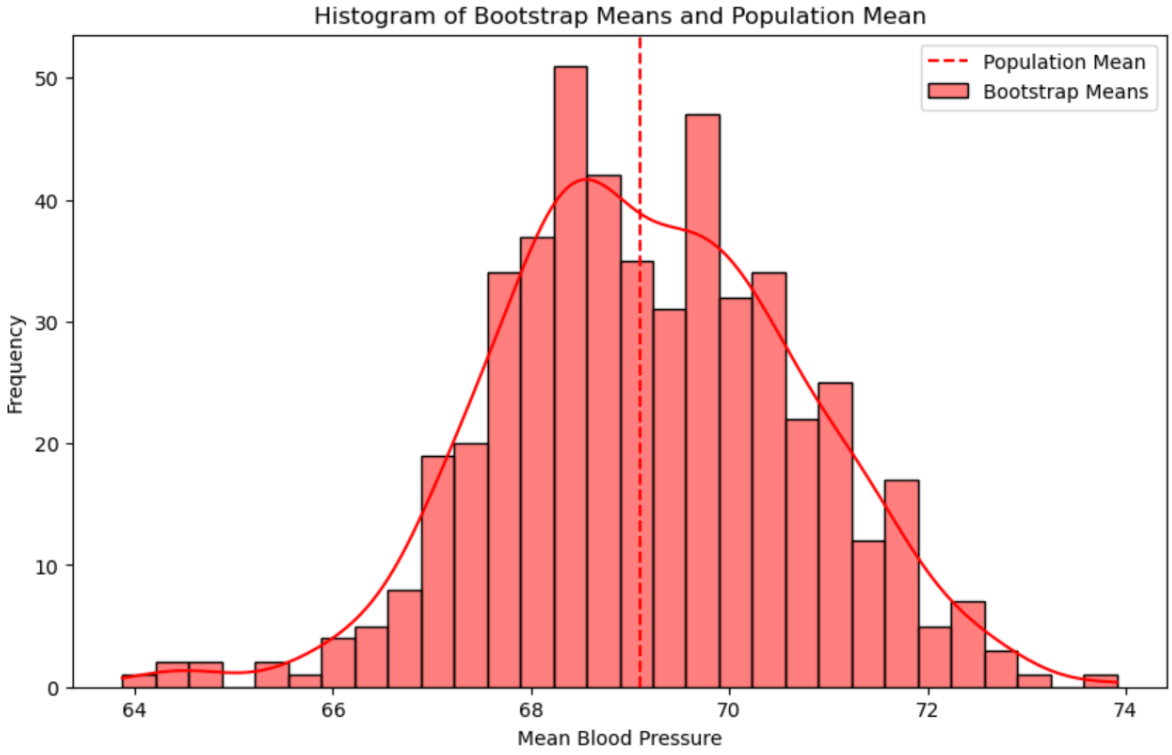
Sample 98th Percentile of BMI: This value represents the BMI at the 98th percentile in the randomly sampled subset of 25 patients from the population.

The bar plot compares these two values, providing insights into the extreme high BMI

'silver' represents the population and 'gold' represents the sample:

The bar plot shows that the 98th percentile of BMI in the sample is higher than that of the population. This suggests that among the 25 randomly sampled patients, the subset might have a higher prevalence of extreme BMI values compared to the entire population.

2 c)

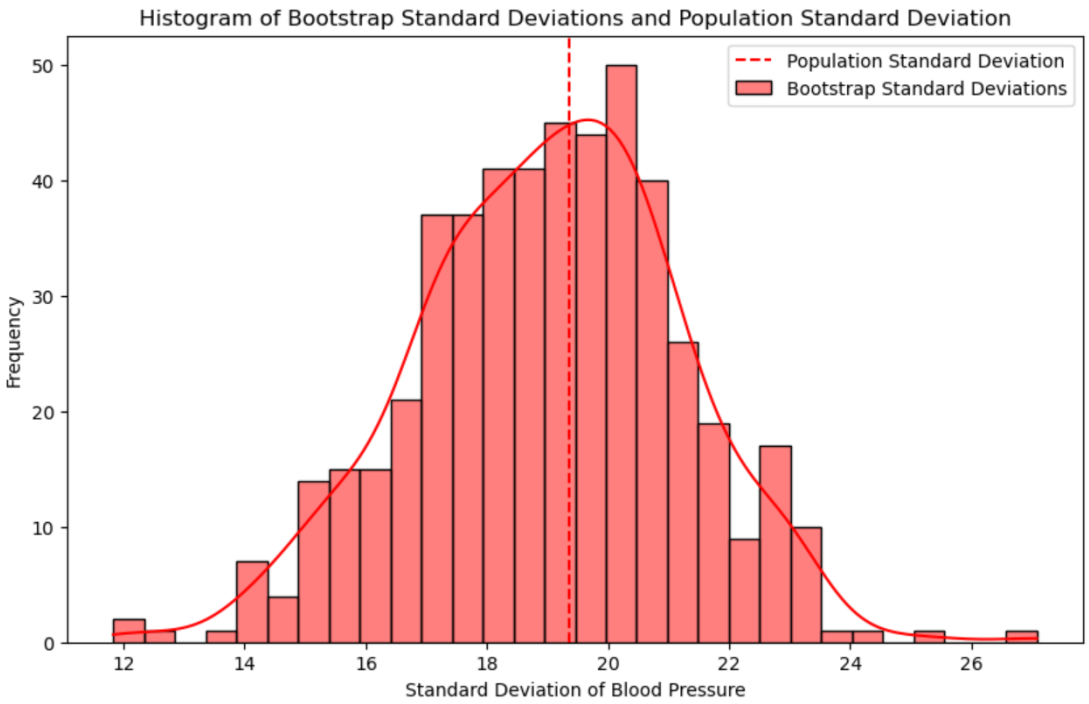


Histogram of Bootstrap Means and Population Mean:

The histogram compares the distribution of bootstrap means (estimated from 500 bootstrap samples) with the population mean of blood pressure.

The vertical dashed line represents the population mean of blood pressure.

From the histogram, it appears that the bootstrap means are centered around the population mean, suggesting that the bootstrap estimates are consistent with the population mean.

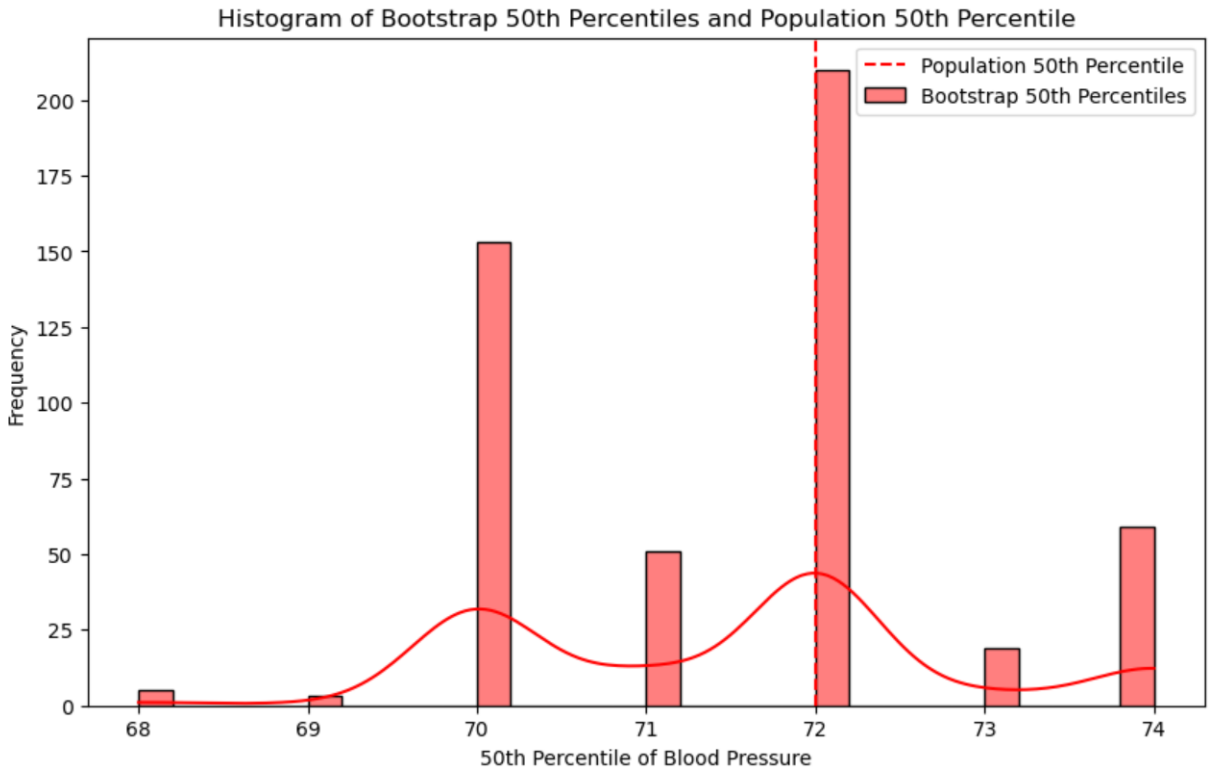


Histogram of Bootstrap Standard Deviations and Population Standard Deviation:

This histogram compares the distribution of bootstrap standard deviations (estimated from 500 bootstrap samples) with the population standard deviation of blood pressure.

The vertical dashed line represents the population standard deviation of blood pressure.

The histogram indicates how the variability of the bootstrap standard deviations compares to the population standard deviation.



Histogram of Bootstrap 50th Percentiles and Population 50th Percentile:

This histogram compares the distribution of bootstrap estimates for the 50th percentile (median) of blood pressure with the population 50th percentile.

The vertical dashed line represents the population 50th percentile of blood pressure.

The histogram provides insights into how the median estimate from bootstrap samples compares to the population median.