**Principles of Data Science Assignment 2**

SAI SARANYA BYRAPUNENI

16338876

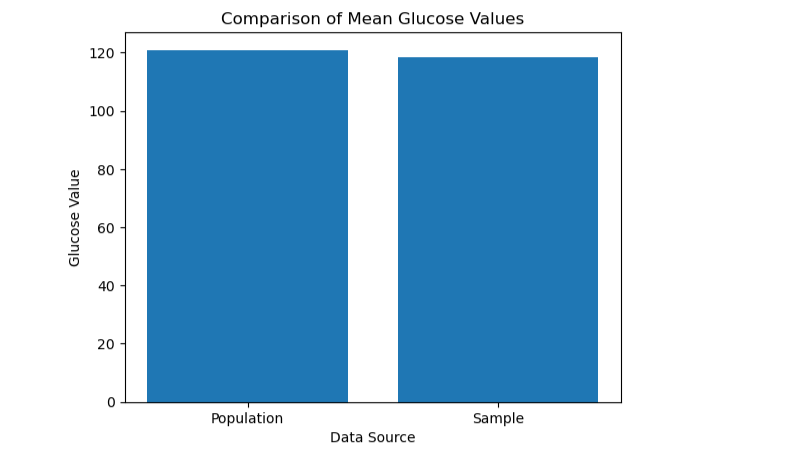
a) set a seed (to ensure work reproducibility) and take a random sample of 25 observations and  
find the mean Glucose and highest Glucose values of this sample and compare these statistics  
with the population statistics of the same variable. You should use charts for this comparison.

From the sample dataset,the mean value of glucose is : 118.32

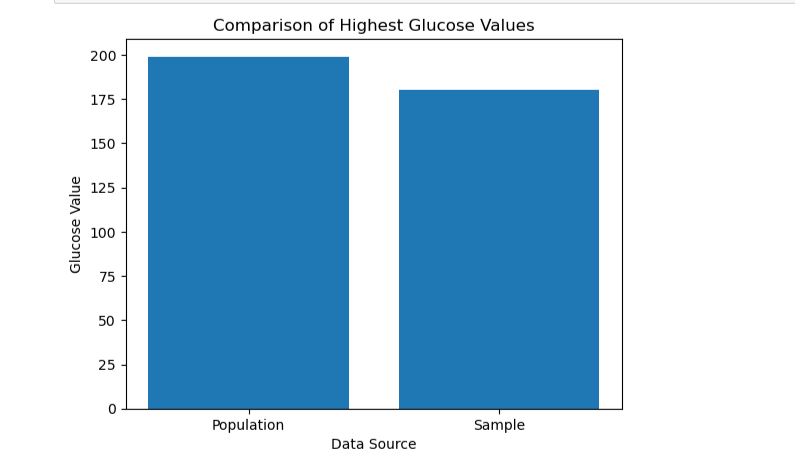
From the sample dataset,the heighest value of glucose is : 180

In entire the dataset,the mean value for glucose values is : 120.89453125

In entire the dataset,the maximum value for glucose values is : 199

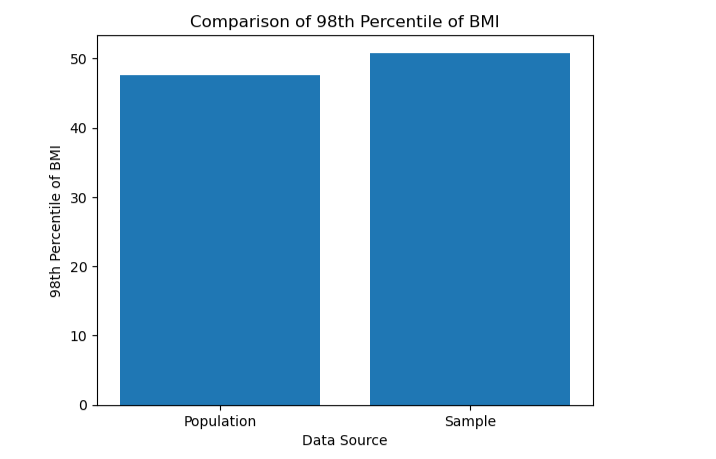


The above bar graph shows the comparison of mean of glucose for sample and population



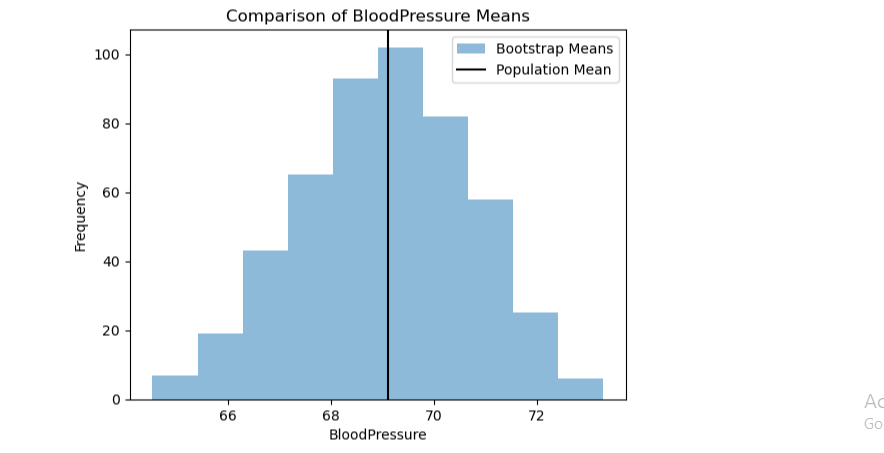
above bar graph shows the comparison of maximum of glucose for sample and population

b) Find the 98th percentile of BMI of your sample and the population and compare the results  
using charts. (5 points)

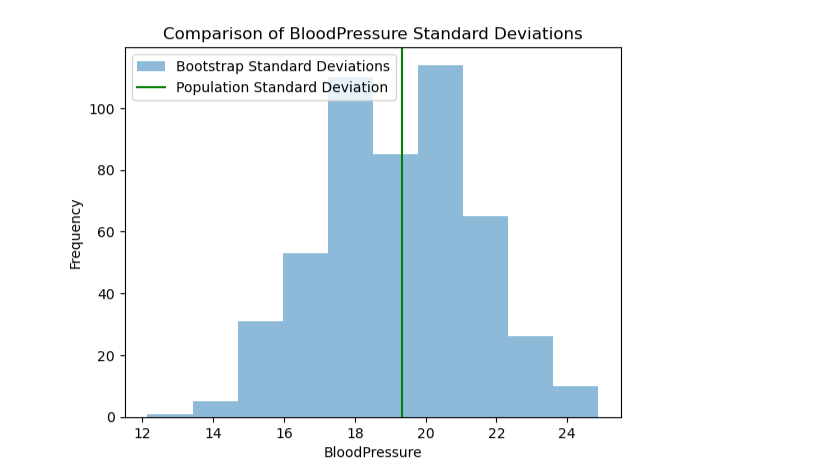


above bar graph shows the comparison of 98th percentile of BMI for sample and population

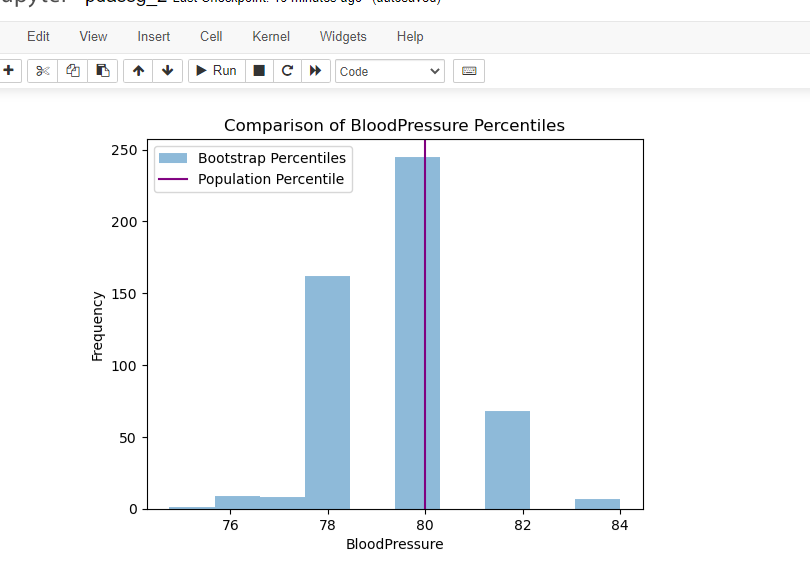
c) Using bootstrap (replace= True), create 500 samples (of 150 observation each) from the  
population and find the average mean, standard deviation and percentile for BloodPressure and  
compare this with these statistics from the population for the same variable. Again, you should  
create charts for this comparison. Report on your findings. (10 points)



Above bar graph shows the comparison of Blood pressure mean value



Above bar graph shows the comparison of Blood pressure standard deviation value



Above bar graph shows the comparison of Blood pressure percentile