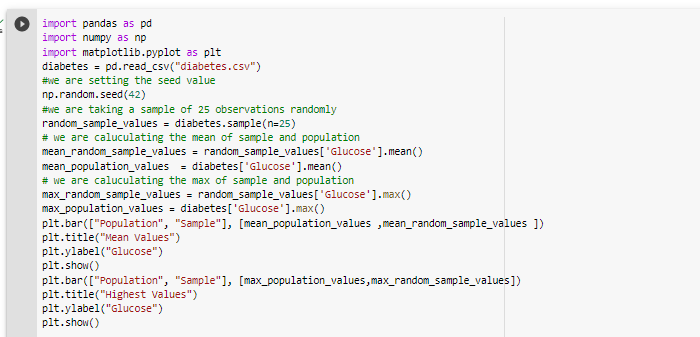
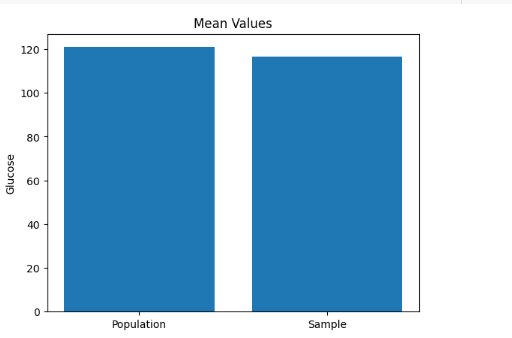
**PDS ASSIGNMENT-2 REPORT**

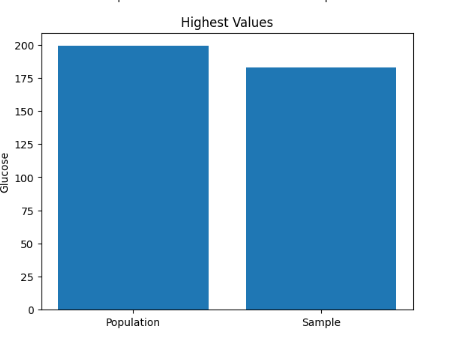
**SAI PRIYATHAM GUMMADI**

A) A sample of 25 observations was randomly chosen from the dataset and seed value is set so as to ensure the reproducibility of the data, and the mean and maximum values for both the sample and the population were computed.



Created a bar graph to compare the max and mean of sample and population values. The graph demonstrates that the maximum glucose value for the sample is lower than the maximum glucose value for the population, the mean glucose value for the sample is lower than the mean glucose value for the population.

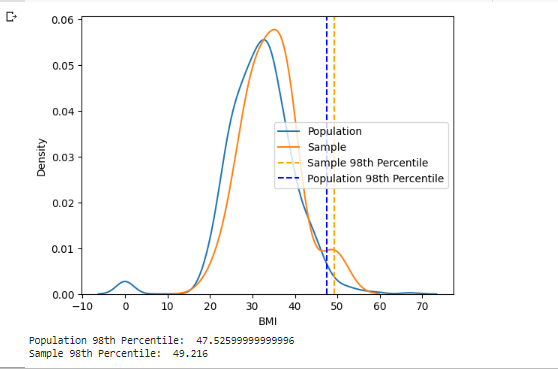




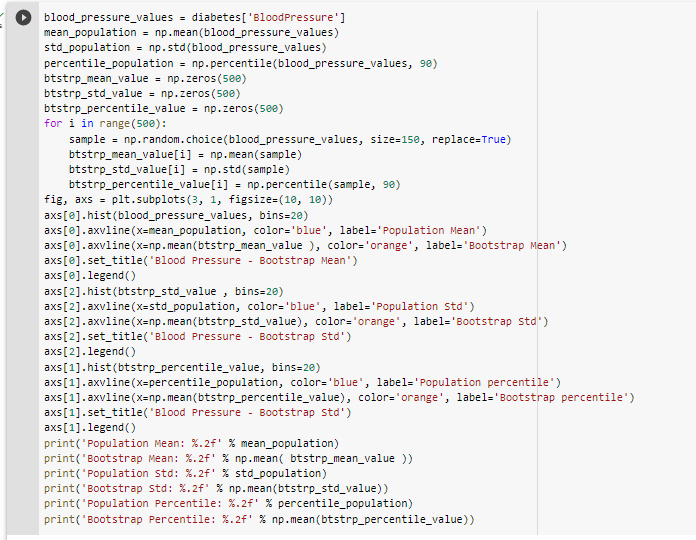
B) Compared the BMI values of a sample of 25 observations to the BMI values of the total population using a kernel density estimate (KDE). Additionally, the 98th percentile values for the sample and population are determined and shown on the plot as vertical dashed lines.



The KDE plot demonstrates that the sample BMI data have a bigger peak and longer tails than the population BMI values. The 98th percentile of the BMI values for the sample is also greater than the 98th percentile of the BMI values for the population.



C) Here, we are generating 500 samples using bootstrap and then we are caluculating the mean,stanadard deviation and 90th percentile for bootstrap values and population values.Histograms are used to represent the distribution of the blood pressure data and the bootstrap statistics generated from the data.



In histogram the population's blood pressure measurements are indicated a blue vertical line and 500 bootstrap samples as an orange line. The histogram demonstrates that the bootstrap mean values are grouped around the population mean and that the distribution's shape resembles that of the population as a whole.

The bootstrap sample means are a decent approximation of the population and the bootstrap values such as mean, standard deviation and percentiles are near to the population values

