Sukumar Bodapati -16326105

Results:

Task 1:

Data:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1

```
[ ]
Glucose_mean_s=df1['Glucose'].mean()

print("Glucose mean sample : "+str(Glucose_mean_s))
```

```
Glucose mean sample : 113.92
```

```
[ ]

Glucose_max_s=df1['Glucose'].max()
print("Glucose max sample : "+str(Glucose_max_s))
```

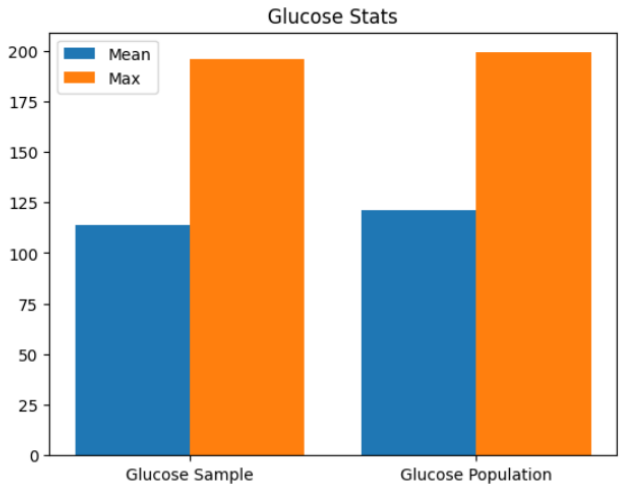
```
Glucose max sample : 196
```

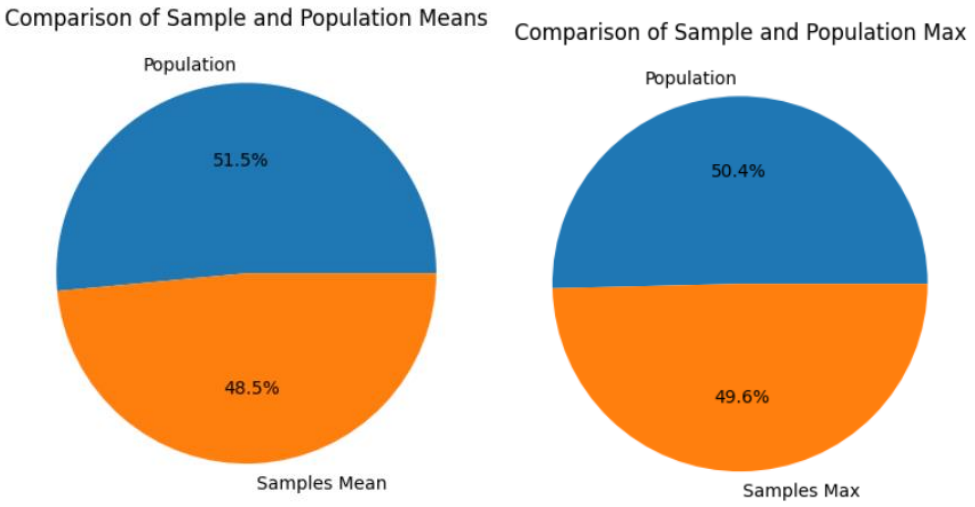
```
[ ]

Glucose_mean_p=df['Glucose'].mean()

Glucose_max_p=df['Glucose'].max()
print("Glucose mean population : "+str(Glucose_mean_p))
print("Glucose max population : "+str(Glucose_max_p))
```

```
Glucose mean population : 120.89453125
Glucose max population : 199
```



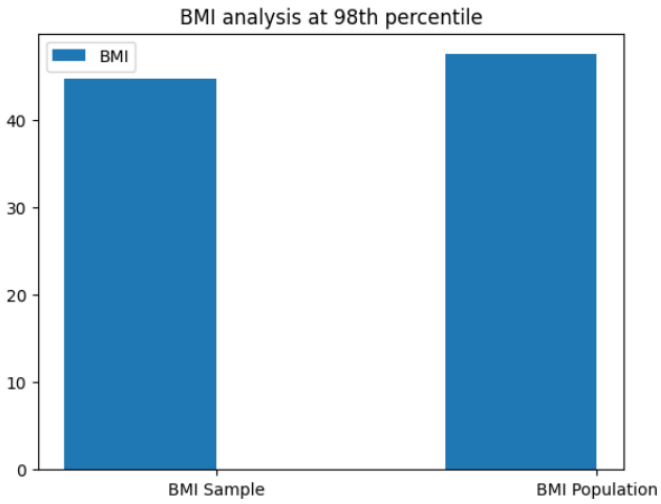


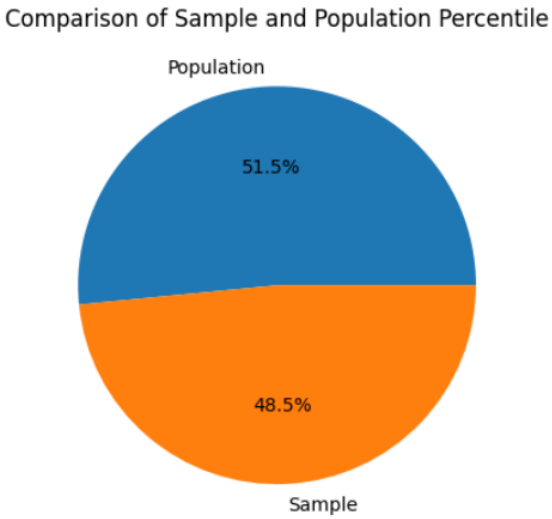
By the graphs the values of mean and max value of glucose in the sample and population are very closer.

Task 2:

```
[ ]  print("BMI percentile of sample : "+str(sample_bmi_perc))
     print("BMI percentile of population : "+str(population_bmi_perc))
```

```
BMI percentile of sample : 44.676
BMI percentile of population : 47.52599999999996
```





At the 98TH percentile, both the sample data and population are quite closer.

Task 3:

```
[ ]  bp_mean_s=sample_data['BloodPressure'].mean()
     bp_mean_p=df['BloodPressure'].mean()
     print("Blood Pressure mean of sample data : "+str(bp_mean_s))
     print("Blood Pressure mean of population : "+str(bp_mean_p))
```

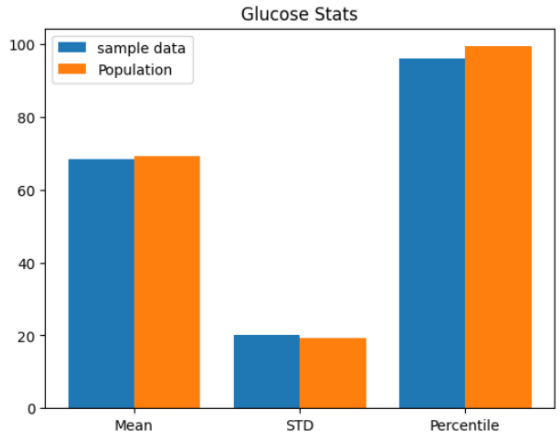
```
Blood Pressure mean of sample data : 68.348
Blood Pressure mean of population : 69.10546875
```

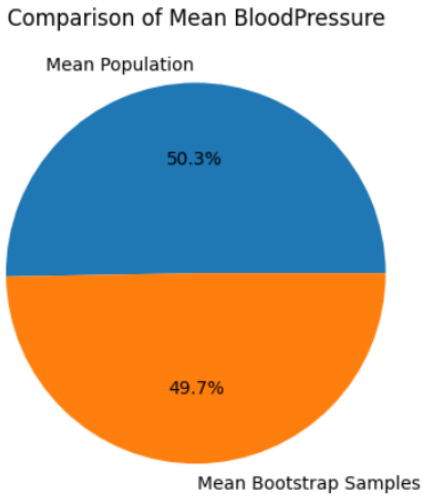
```
[ ]
     import statistics as st
     bp_sd_s=st.stdev(sample_data['BloodPressure'])
     bp_sd_p=st.stdev(df['BloodPressure'])
     print("Blood Pressure standard deviation of sample data : "+str(bp_sd_s))
     print("Blood Pressure standard deviation of population : "+str(bp_sd_p))
```

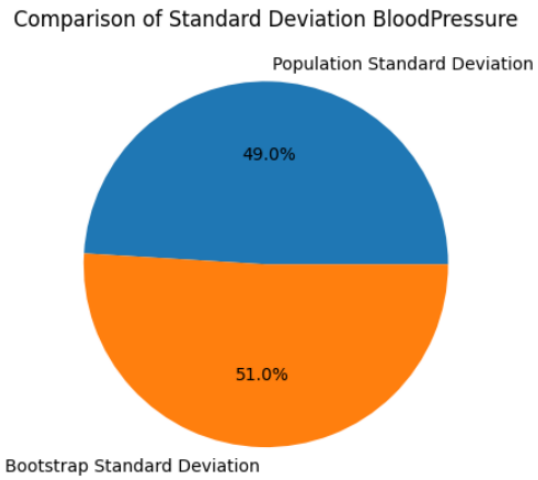
```
Blood Pressure standard deviation of sample data : 20.10729974427228
Blood Pressure standard deviation of population : 19.355807170644777
```

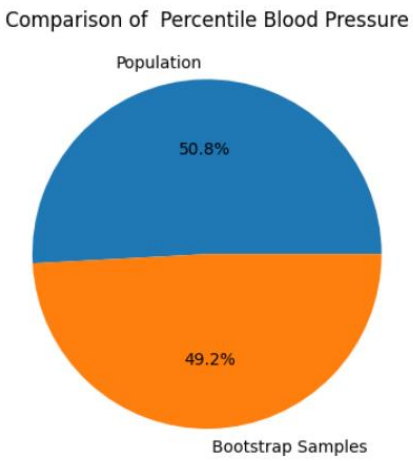
```
[ ]  bp_perc_s=np.percentile(sample_data['BloodPressure'], 98)
     bp_perc_p=np.percentile(df['BloodPressure'],98)
     print("Blood Pressure mean of sample data : "+str(bp_perc_s))
     print("Blood Pressure mean of population : "+str(bp_perc_p))
```

```
Blood Pressure mean of sample data : 96.0
Blood Pressure mean of population : 99.31999999999994
```









The mean population is 69. 105.The population standard deviation is 19.355807170644777, which shows that the statistic of interest has a large range of frequencies in the population. The population percentile of 99.31999999999994 shows that the population may contain some extreme values or outliers.

The mean of the bootstrap sample is 68.34. The standard deviation of the bootstrap sample is 19.35, which is slightly less than the population standard deviation. This implies that the variable's spread in the sample may be slightly lower than in the population. The population percentile is 99.31, but the bootstrap sample percentile is 96.