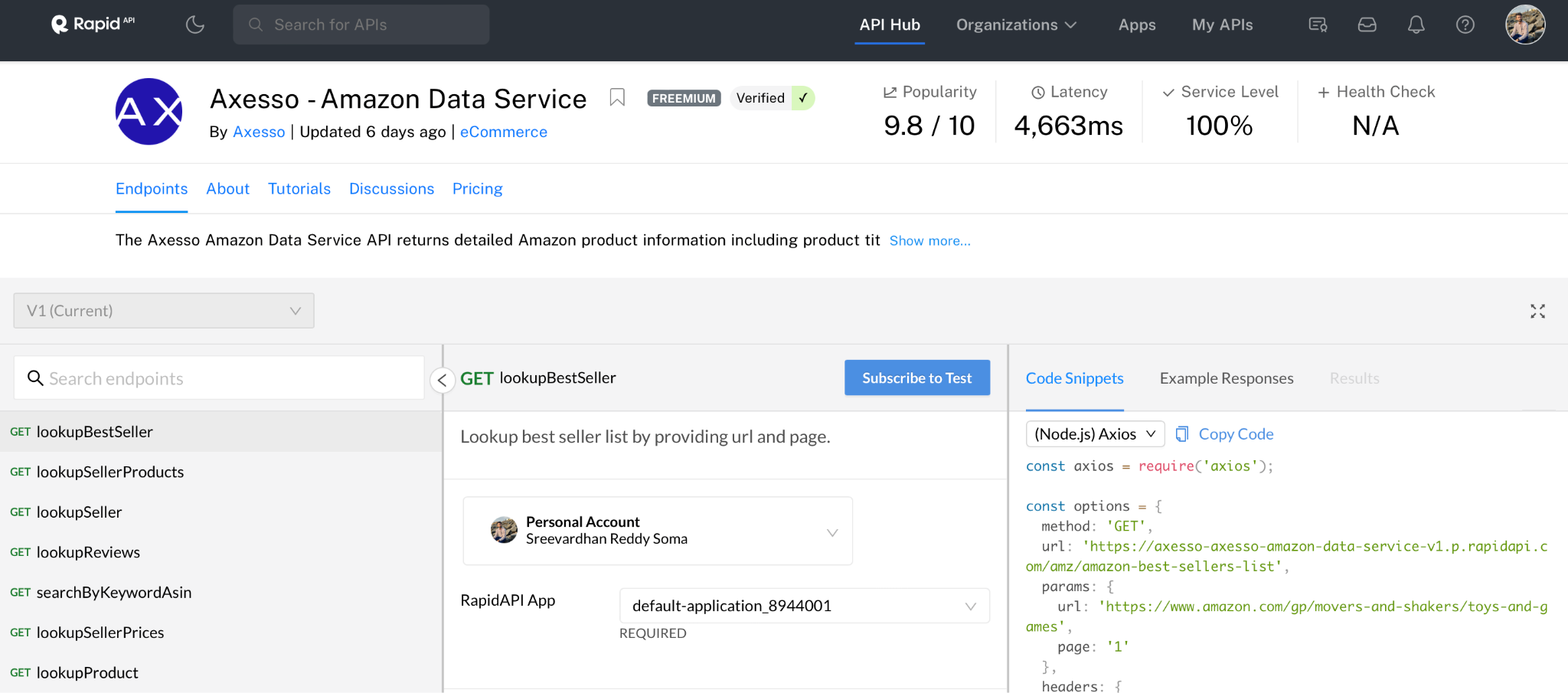
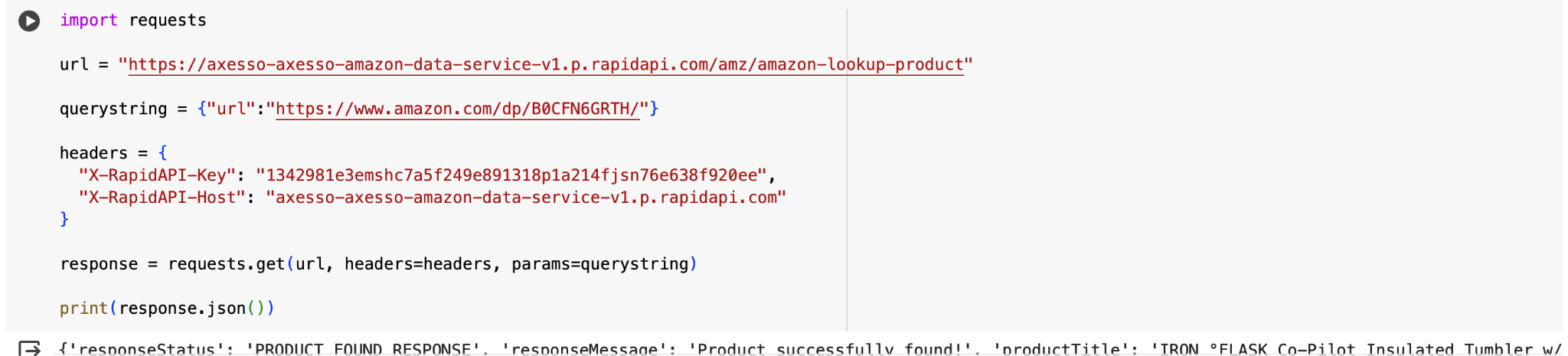
**Assignment-3&4**

**Name: Sreevardhan Reddy Soma** **ID: 16352646**

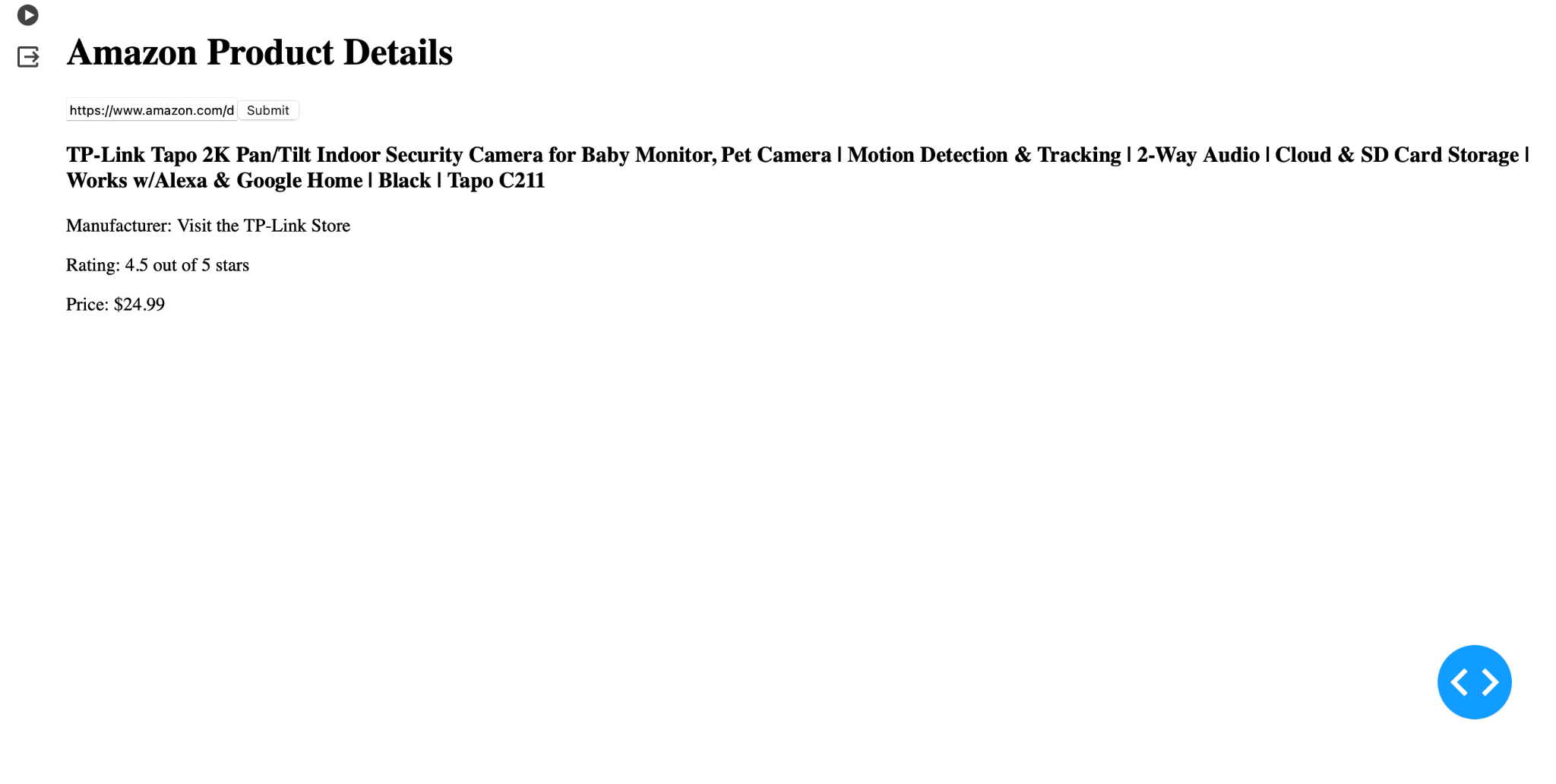
1. **Use any open-source API to access some data in Jason format and then parse the Jason data and display it as some kind of dashboard**
2. The API used is Axesso which is a real-time API which contains the details of the products currently listed in Amazon website.



Inorder to generate the data, we need to request it in colab using API keys



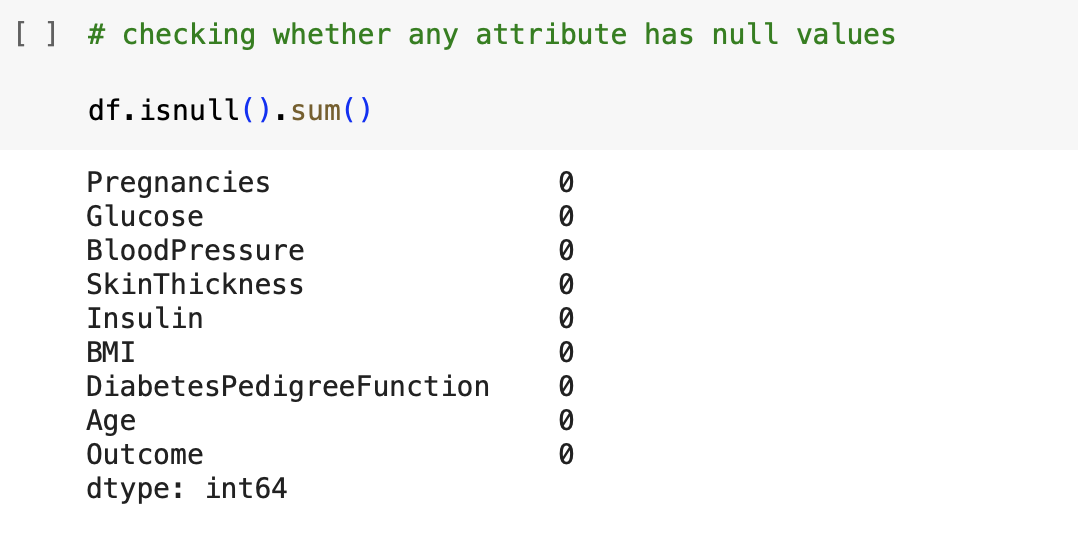
1. **Creating the dashboard:** The dashboard here takes the link of the Amazon product as input and as soon as we click submit, all the details about the product are displayed in the dashboard. The dashboard looks as follows.



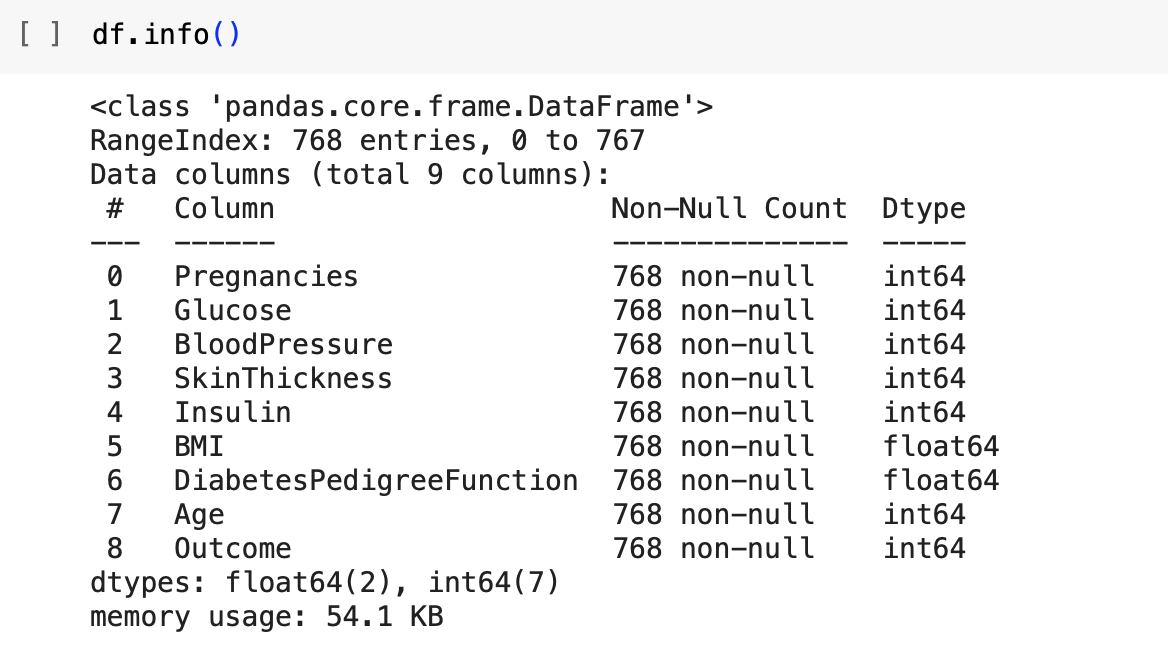
2.

1. **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.**

Checking for null values

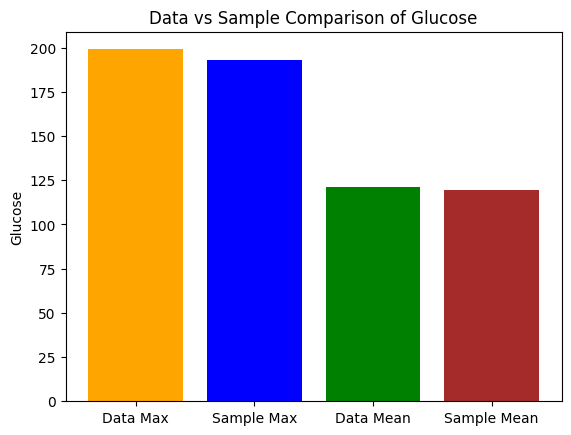


The below function gives all the information about dataset





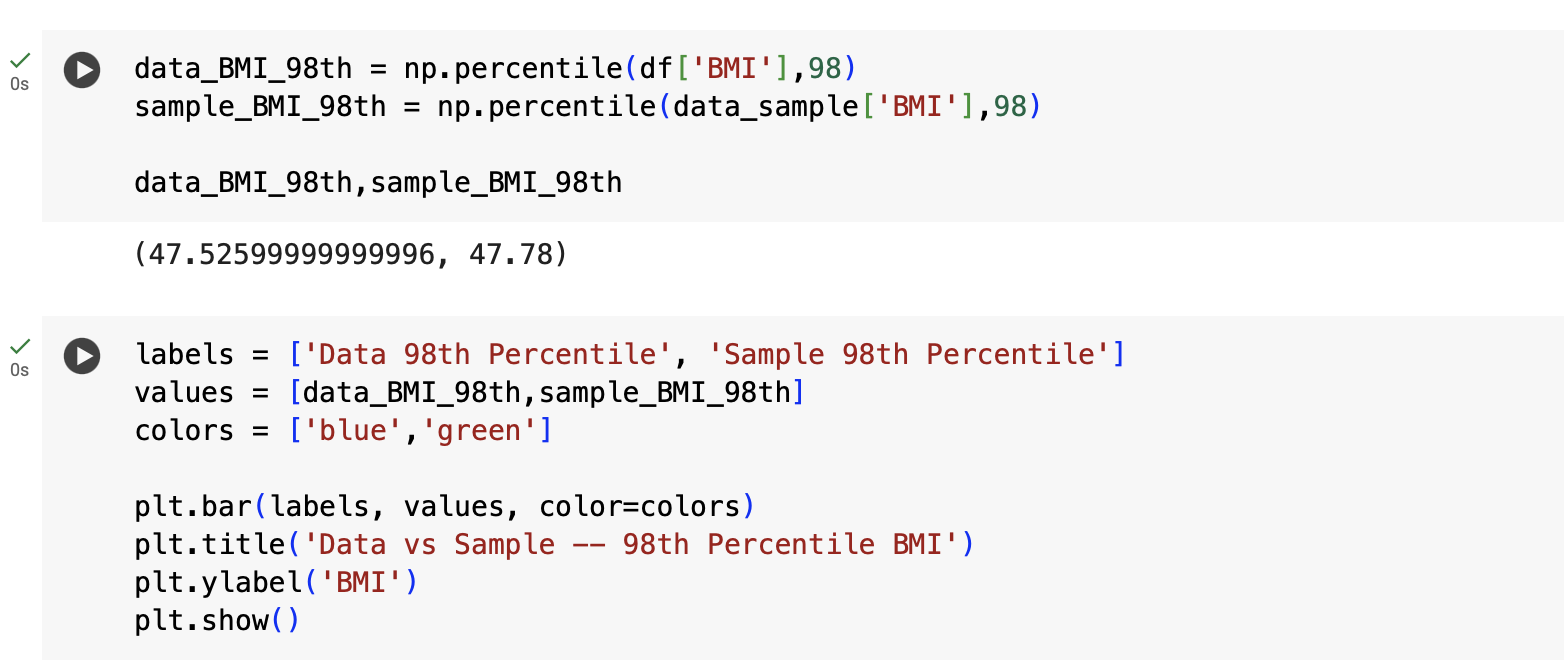
**Result:**



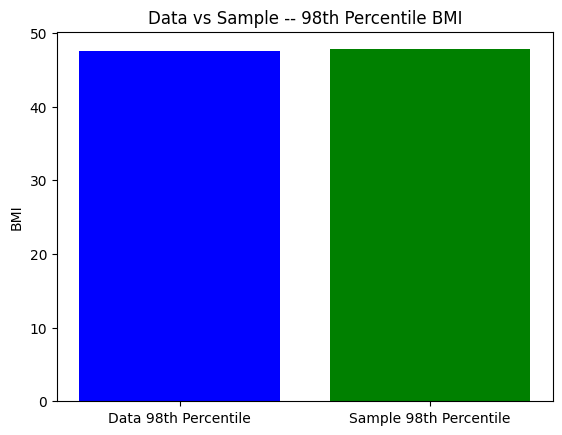
**Analysis:**

From the above analysis, data mean(120.89) is slightly higher than sample mean(119.24) for the glucose. This depicts that the average glucose value in the population dataset is higher than that of the glucose values in the sample data which we took randomly. Coming to maximum glucose value in the population and sample, it is 199 for the population and 193 for the sample. Even in this case, the same thing is happening as above where the maximum glucose levels in the population is higher than in the sample dataset. This comparison is behaving almost similarly in both the cases.

**b. Find the 98th percentile of BMI of your sample and the population and compare the results using charts**



**Result:**



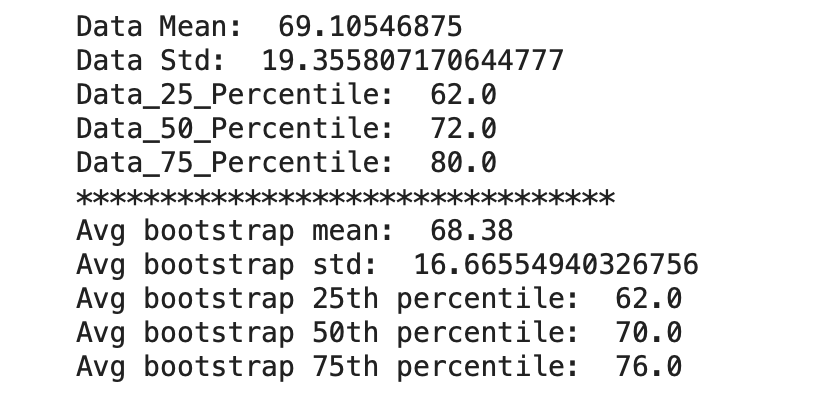
**Analysis:**

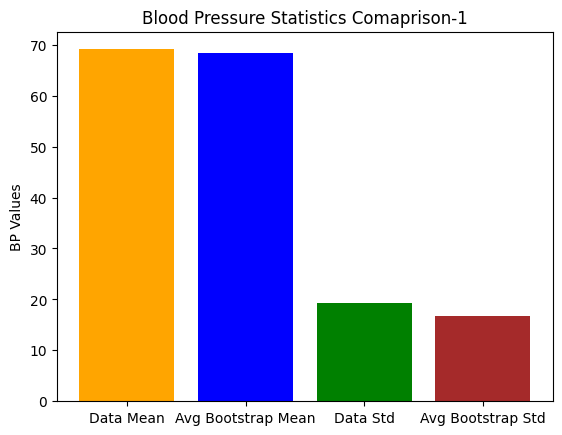
The 98th percentile of the population dataset is 47.525 whereas the obtained 98th percentile for the sample dataset is 47.78. Interestingly, here the percentile for the sample data is slightly higher than the population data but the difference between the two is very negligible.

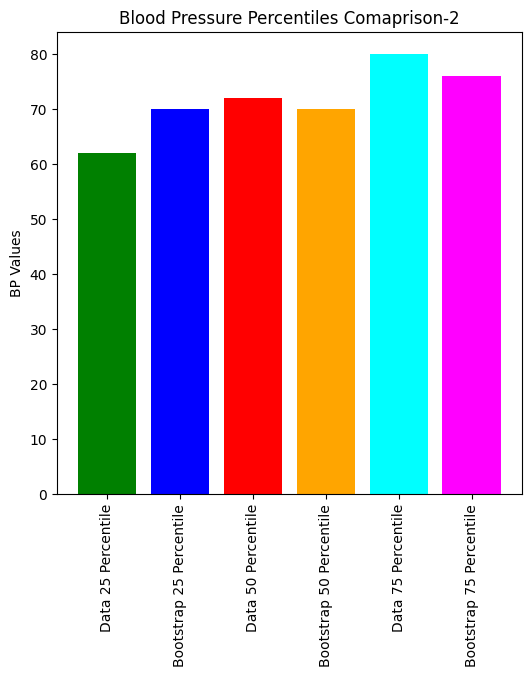
c) **Using bootstrap (replace= True), create 500 samples (of 150 observation each) from the population and find the average mean, standard deviation and percentile for BloodPressure and compare this with these statistics from the population for the same variable. Again, you should create charts for this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



**Result:**







**Analysis:**

Firstly, by observing the mean and standard deviation of the Blood Pressure feature for the population and bootstrap data, we can observe that both these factors are higher for population data than that of sample data which indicates that the blood pressure readings of bootstrap samples are consistent with the mean population. The percentiles of the sample and population are almost similar with very few negligible values and this indicates that the distribution of Blood pressure in the sample is performed well.