Results Obtained

Google Colab Link (For proper graphs and conclusions)

https://colab.research.google.com/drive/1Y6AHLMTIP2CI7vsttPB7CpL2NfGaYuN9#scrollTo=dQaM5IvU0k8e

Support Vector Machines, Naïve Bayes algorithm and Random Forest Algorithms are used to test and fit the data. Support Vector Machines and naïve bayes algorithms achieved much higher accuracy than random forest for our classification of training and test split.

```
The accuracy of test data set in SVM  77.27272727272727 %
The accuracy of test data set in Bayes algorithm is   77.27272727272727 %
The accuracy of test data set in random forest method   is   74.67532467532467 %
```

All the functions of data like mean values above 75% ,50%,maximum values ,minimun value etc

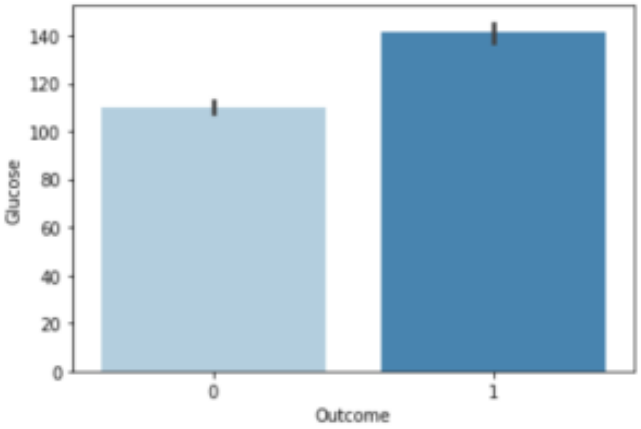
	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
count	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000
mean	3.845052	120.894531	69.105469	20.536458	79.799479	31.992578	0.471876	33.240885	0.348958
std	3.369578	31.972618	19.355807	15.952218	115.244002	7.884160	0.331329	11.760232	0.476951
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.078000	21.000000	0.000000
25%	1.000000	99.000000	62.000000	0.000000	0.000000	27.300000	0.243750	24.000000	0.000000
50%	3.000000	117.000000	72.000000	23.000000	30.500000	32.000000	0.372500	29.000000	0.000000
75%	6.000000	140.250000	80.000000	32.000000	127.250000	36.600000	0.626250	41.000000	1.000000
max	17.000000	199.000000	122.000000	99.000000	846.000000	67.100000	2.420000	81.000000	1.000000

Bar Graphs

Bar Graph of Each Column with respect to the outcome

```
[ ]  sns.barplot(x='Outcome',y='Glucose',data=diabetes_dataset,palette="Blues")
```

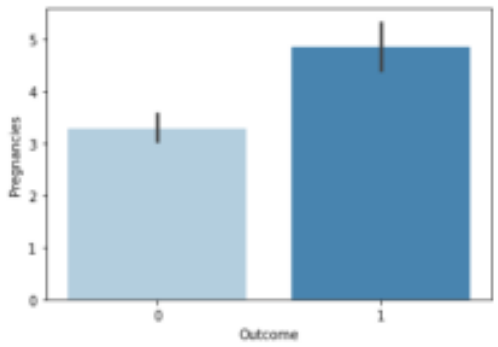
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f74cc1492d0>
```



The glucose and outcome relation could be seen as glucose increases the outcome of being diabetic also increases

```
[ ]  sns.barplot(x='Outcome',y='Pregnancies',data=diabetes_dataset,palette="Blues")
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f74cc0b8e50>
```

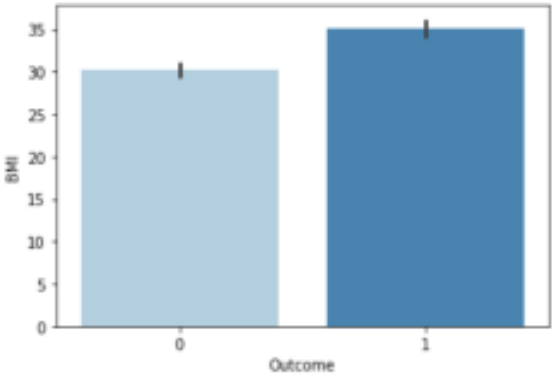


The glucose and outcome relation could be seen as insulin increases the outcome of being diabetic also increases

Relation between Outcome and BMI

```
[ ]  sns.barplot(x='Outcome',y='BMI',data=diabetes_dataset,palette="Blues")
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f74dcf6df10>
```

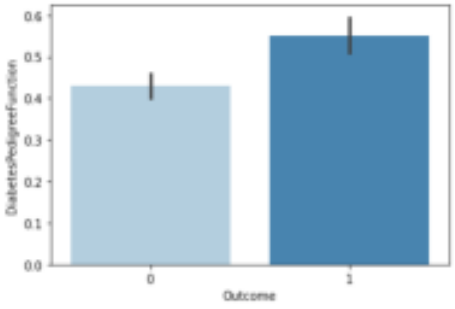


The BMI and outcome relation could be seen as glucose increases the outcome of being diabetic also increases(less difference as compared to others)

Relation between Diabetes Pedigree Function and Outcome

```
[ ]  sns.barplot(x='Outcome',y='DiabetesPedigreeFunction',data=diabetes_dataset,palette="Blues")
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fdc5d4fbe90>
```



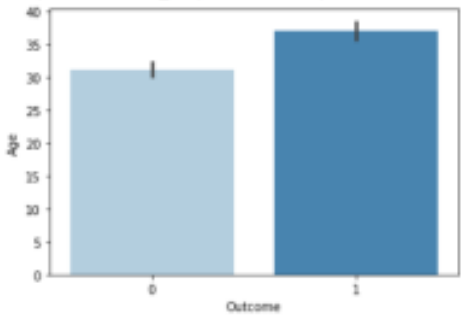
The Diabetes Pedeigree function and outcome relation could be seen as glucose increases the outcome of being diabetic also increases

```
[ ]  Relation between Age and Outcome
```

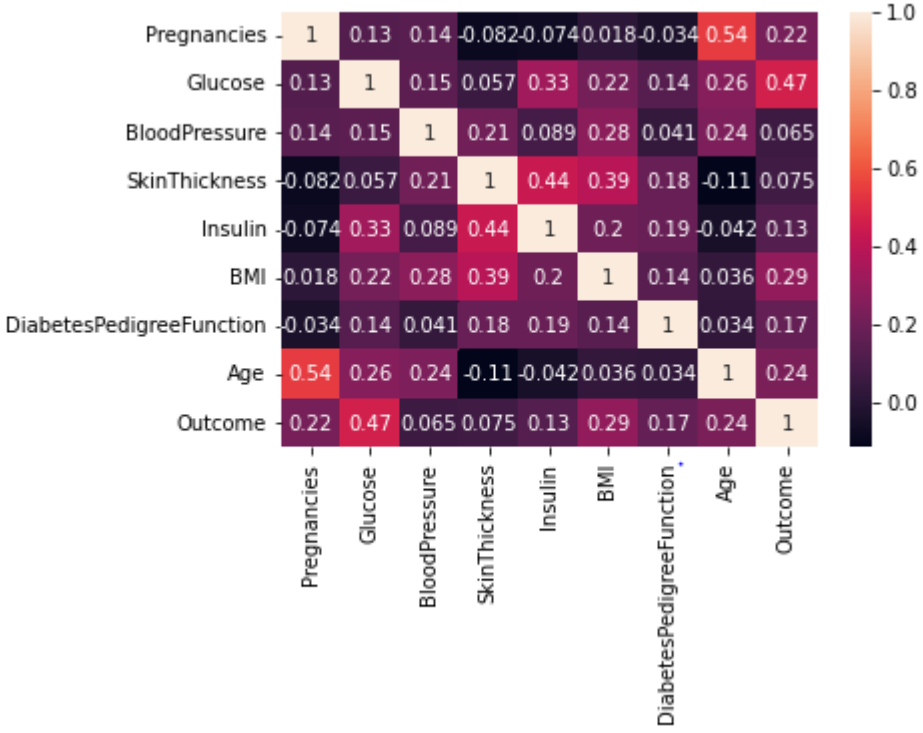
+ Code + Text

```
[ ]  sns.barplot(x='Outcome',y='Age',data=diabetes_dataset,palette="Blues")
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7fdc05651710>
```



K-map(Correlation between various attributes)



As we can see from the k-map that Age, BMI, Insulin and glucose are the ones which are having a high correlation with the outcome.

After finding out the mean values we input the values to find out the outcome whether we get diabetic or not. We got the result as non-diabetic.

Then we tried for values above than 75% and we got the predication as diabetic. This was also predicted by the describe() function of the data

Then we used same values for variables such as insulin, glucose, bmi and age as predicted by our heatmap and all other values at mean values. We got the output as diabetic. So our prediction from heat map is also true.

Mean values usually end up being non-diabetic but as we increase values slightly towards 75 per cent we go up on the outcome of being diabetic.

Also insulin, glucose, bmi and age highly impact the outcome of being diabetic as analysed.

For graphs and diagrams, please refer to the colab link