KLASIFIKASI PENYAKIT DIABETES DENGAN MENGGUNAKAN ALGORITHMA NAIVE BAYES

	preg	plas	pres	skin	insu	mass	pedi	age	class
0	6.0	148.0	72.0	35.0	0.0	33.6	0.627	50.0	b'tested_positive'
1	1.0	85.0	66.0	29.0	0.0	26.6	0.351	31.0	b'tested_negative'
2	8.0	183.0	64.0	0.0	0.0	23.3	0.672	32.0	b'tested_positive'
3	1.0	89.0	66.0	23.0	94.0	28.1	0.167	21.0	b'tested_negative'
4	0.0	137.0	40.0	35.0	168.0	43.1	2.288	33.0	b'tested_positive'
...	...	...	...	...	...	...	...	...	...
763	10.0	101.0	76.0	48.0	180.0	32.9	0.171	63.0	b'tested_negative'
764	2.0	122.0	70.0	27.0	0.0	36.8	0.340	27.0	b'tested_negative'
765	5.0	121.0	72.0	23.0	112.0	26.2	0.245	30.0	b'tested_negative'
766	1.0	126.0	60.0	0.0	0.0	30.1	0.349	47.0	b'tested_positive'
767	1.0	93.0	70.0	31.0	0.0	30.4	0.315	23.0	b'tested_negative'

768 rows × 9 columns

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6.0	148.0	72.0	35.0	0.0	33.6	0.627	50.0	b'tested_positive'
1	1.0	85.0	66.0	29.0	0.0	26.6	0.351	31.0	b'tested_negative'
2	8.0	183.0	64.0	0.0	0.0	23.3	0.672	32.0	b'tested_positive'
3	1.0	89.0	66.0	23.0	94.0	28.1	0.167	21.0	b'tested_negative'
4	0.0	137.0	40.0	35.0	168.0	43.1	2.288	33.0	b'tested_positive'
...	...	...	...	...	...	...	...	...	...
763	10.0	101.0	76.0	48.0	180.0	32.9	0.171	63.0	b'tested_negative'
764	2.0	122.0	70.0	27.0	0.0	36.8	0.340	27.0	b'tested_negative'
765	5.0	121.0	72.0	23.0	112.0	26.2	0.245	30.0	b'tested_negative'
766	1.0	126.0	60.0	0.0	0.0	30.1	0.349	47.0	b'tested_positive'
767	1.0	93.0	70.0	31.0	0.0	30.4	0.315	23.0	b'tested_negative'

768 rows × 9 columns

EXPLORATORY DANA ANALYSIS

DATA PREPROCESSING

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6.0	148.0	72.0	35.0	0.0	33.6	0.627	50.0	1
1	1.0	85.0	66.0	29.0	0.0	26.6	0.351	31.0	0
2	8.0	183.0	64.0	0.0	0.0	23.3	0.672	32.0	1
3	1.0	89.0	66.0	23.0	94.0	28.1	0.167	21.0	0
4	0.0	137.0	40.0	35.0	168.0	43.1	2.288	33.0	1
...	...	...	...	...	...	...	...	...	...
763	10.0	101.0	76.0	48.0	180.0	32.9	0.171	63.0	0
764	2.0	122.0	70.0	27.0	0.0	36.8	0.340	27.0	0
765	5.0	121.0	72.0	23.0	112.0	26.2	0.245	30.0	0
766	1.0	126.0	60.0	0.0	0.0	30.1	0.349	47.0	1
767	1.0	93.0	70.0	31.0	0.0	30.4	0.315	23.0	0

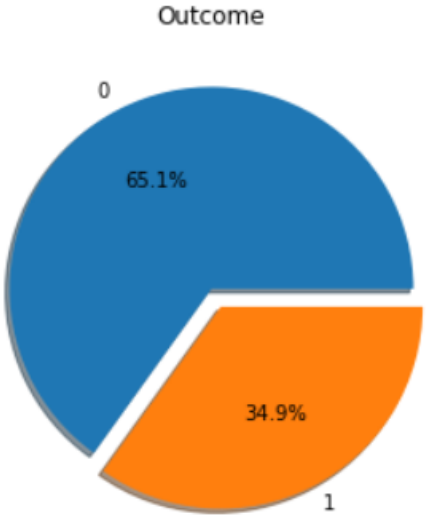
768 rows × 9 columns

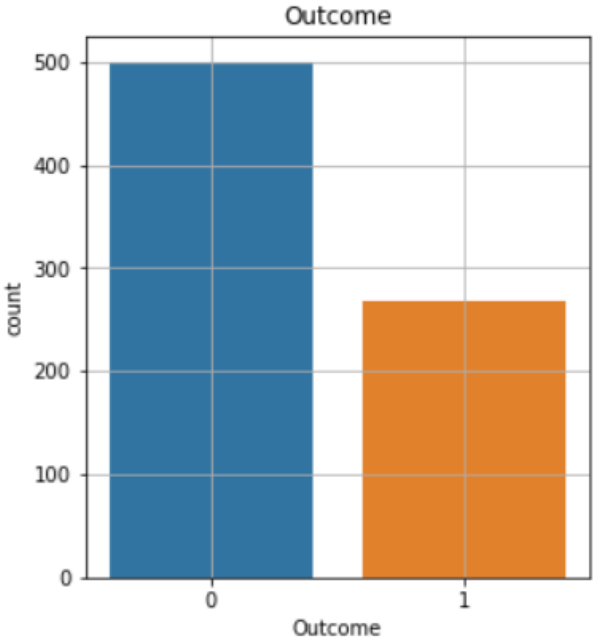
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 768 entries, 0 to 767
Data columns (total 9 columns):
 #   Column                    Non-Null Count  Dtype  
---  ------                    --------------  -----  
 0   Pregnancies               768 non-null    float64
 1   Glucose                   768 non-null    float64
 2   BloodPressure             768 non-null    float64
 3   SkinThickness             768 non-null    float64
 4   Insulin                   768 non-null    float64
 5   BMI                       768 non-null    float64
 6   DiabetesPedigreeFunction  768 non-null    float64
 7   Age                       768 non-null    float64
 8   Outcome                   768 non-null    int64  
dtypes: float64(8), int64(1)
memory usage: 54.1 KB
```

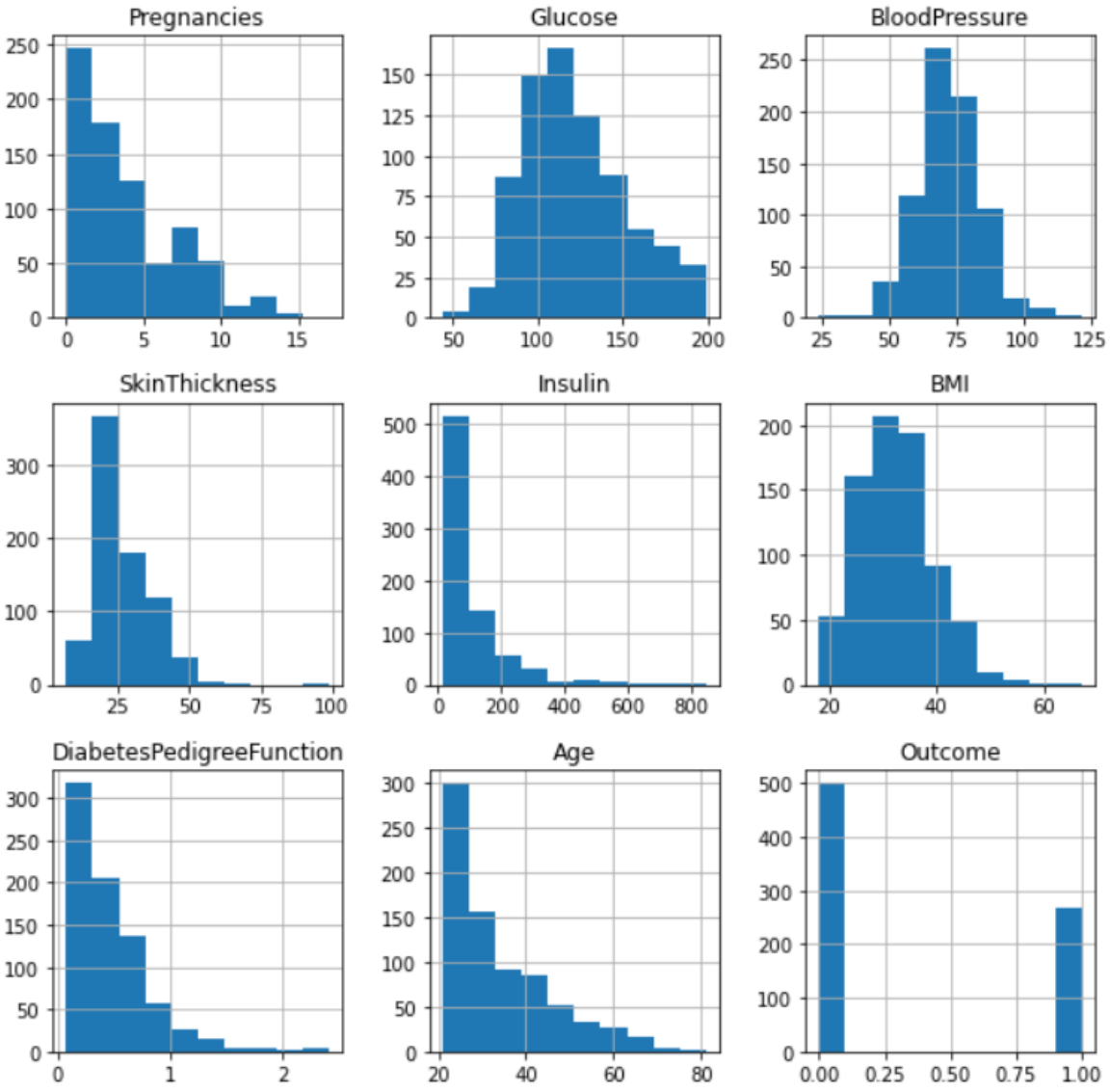
	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
count	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000	768.000000
mean	3.845052	120.894531	69.105469	20.536458	79.799479	31.992578	0.471876	33.240885	0.348958
std	3.369578	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

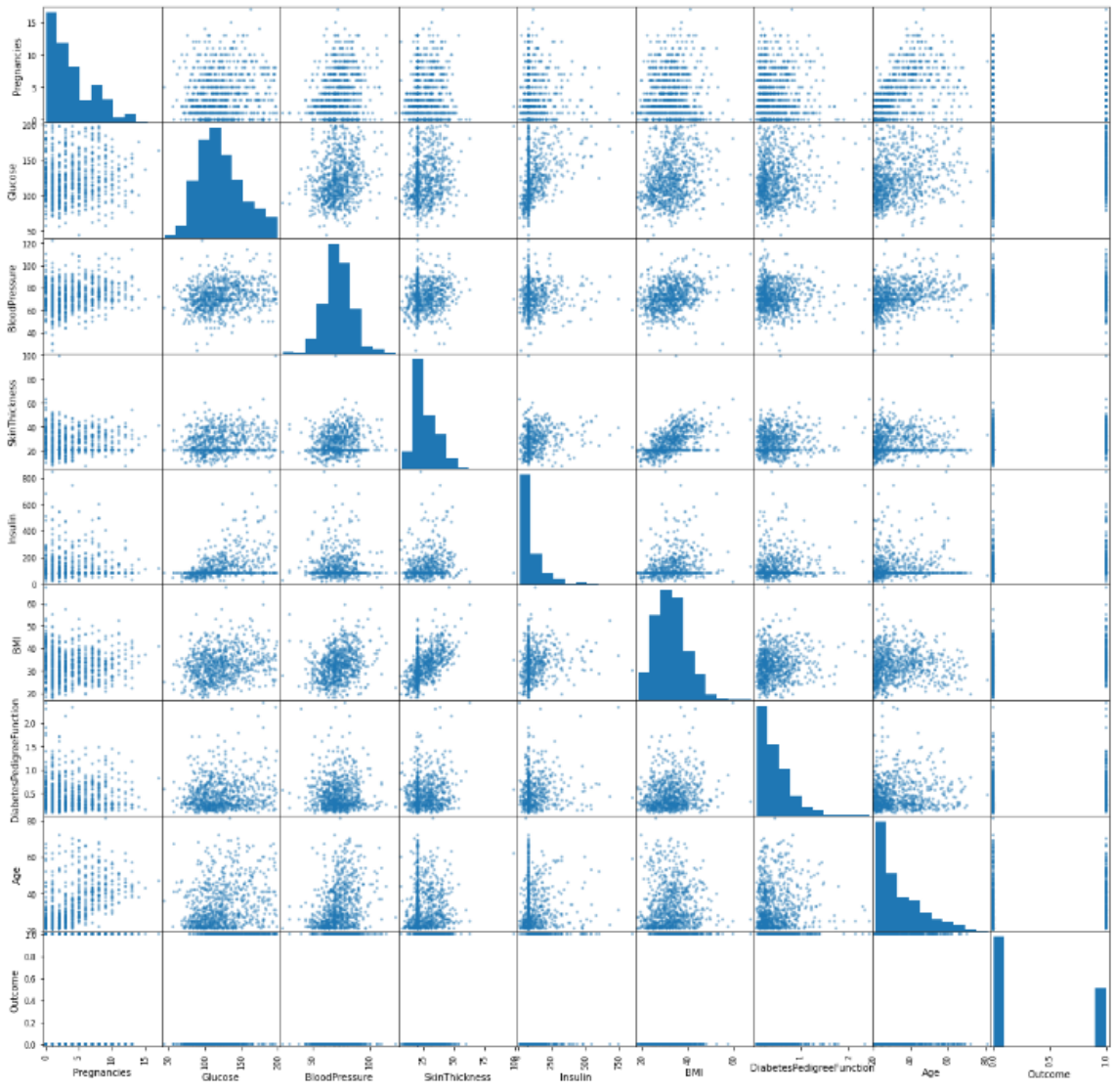
DATA VISUALIZATION

```
Negative (0):  500
Positive (1):  268
```

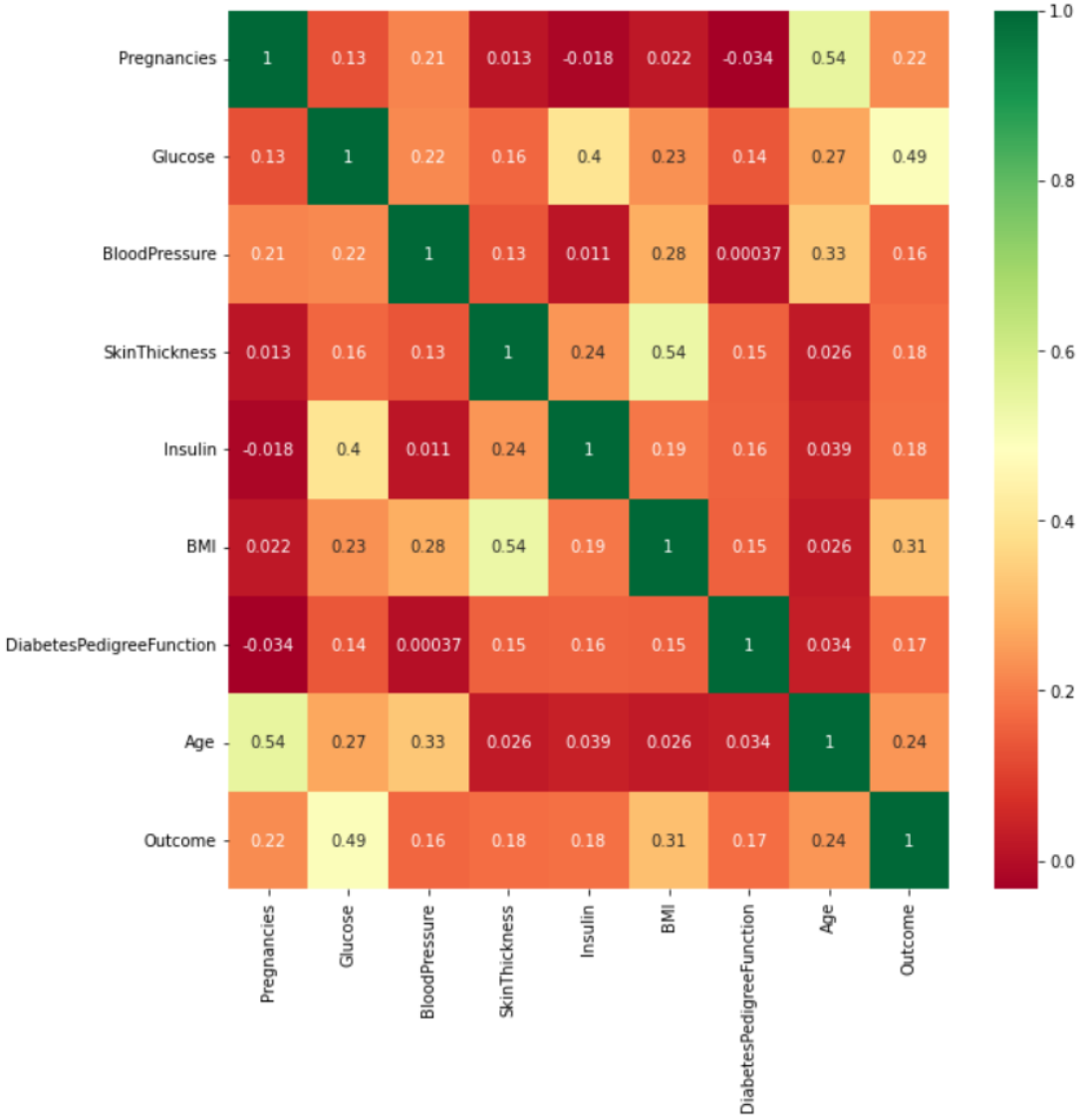





Pregnancies
Glucose
BloodPressure
SkinThickness
Insulin
BMI
DiabetesPedigreeFunction
Age
Outcome


Pregnancies
Glucose
BloodPressure
SkinThickness
Insulin
BMI
DiabetesPedigreeFunction
Age
Outcome

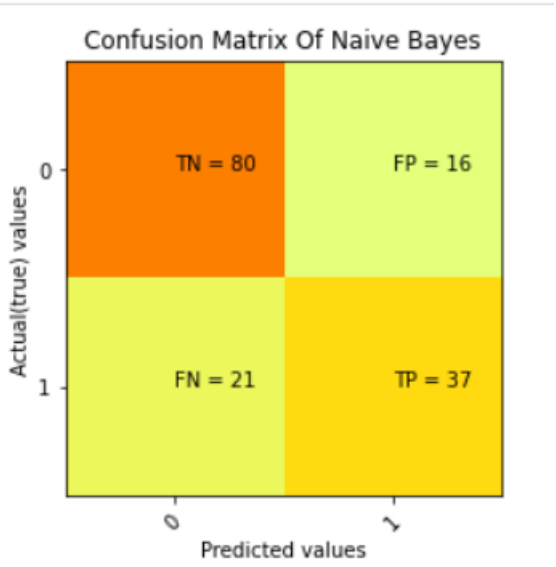

Pregnancies
Glucose
BloodPressure
SkinThickness
Insulin
BMI
DiabetesPedigreeFunction
Age



BUILD MODEL USING NAÏVE BAYES ALGORITHM

```
Train Accuracy of Naive Bayes 74.42996742671009
Accuracy (Test) score of Naive Bayes 75.97402597402598
Accuracy score of Naive Bayes 75.97402597402598
```

CONFUSION MATRIX



Classification Report of Naive Bayes:

	precision	recall	f1-score	support
0	0.7921	0.8333	0.8122	96
1	0.6981	0.6379	0.6667	58
accuracy			0.7597	154
macro avg	0.7451	0.7356	0.7394	154
weighted avg	0.7567	0.7597	0.7574	154