



	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin
0	6	148	72	35	0
1	1	85	66	29	0
2	8	183	64	0	0
3	1	89	66	23	94
4	0	137	40	35	168

[+ Code](#)[+ Text](#)

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin
0	6	148	72	35	0
1	1	85	66	29	0
2	8	183	64	0	0
3	1	89	66	23	94
4	0	137	40	35	168





θ

Pregnancies	0.0
Glucose	0.0
BloodPressure	0.0
SkinThickness	0.0
Insulin	0.0
BMI	0.0
DiabetesPedigreeFunction	0.0
Age	0.0
Outcome	0.0

dtype: float64



Data after processing:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	
0	6	148	72	35	0	3
1	1	85	66	29	0	2
2	8	183	64	0	0	2
3	1	89	66	23	94	2
4	0	137	40	35	168	4

	DiabetesPedigreeFunction	Age	Outcome
0	0.627	50	1
1	0.351	31	0
2	0.672	32	1
3	0.167	21	0
4	2.288	33	1

Missing values after processing:

Pregnancies	0
Glucose	0
BloodPressure	0
SkinThickness	0
Insulin	0
BMI	0
DiabetesPedigreeFunction	0



Summary Statistics:

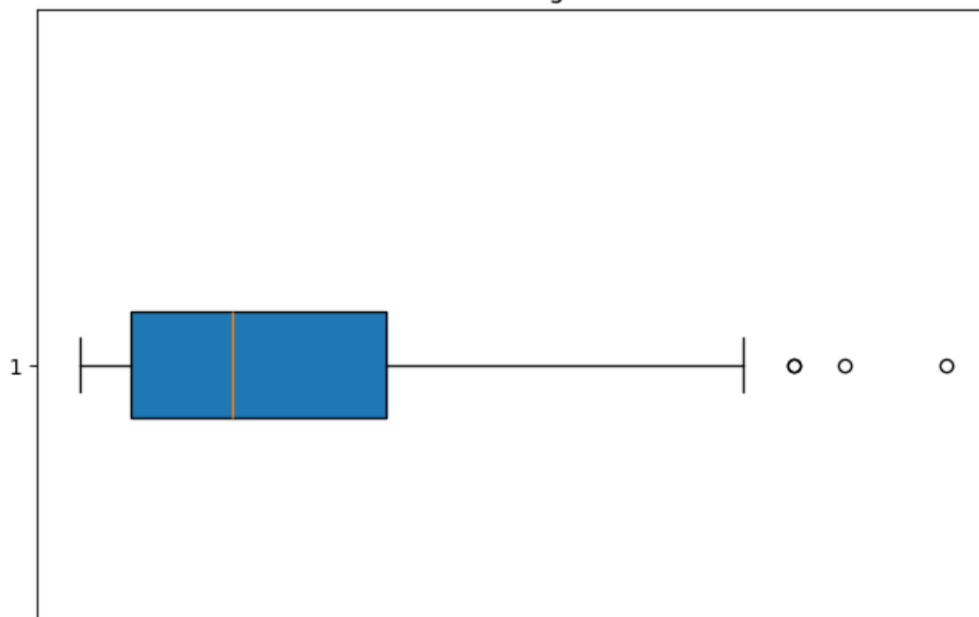
	Pregnancies	Glucose	BloodPressure	SkinThickness	
count	768.000000	768.000000	768.000000	768.000000	768.000000
mean	3.845052	120.894531	69.105469	20.536458	79.933641
std	3.369578	31.972618	19.355807	15.952218	115.980106
min	0.000000	0.000000	0.000000	0.000000	0.000000
25%	1.000000	99.000000	62.000000	0.000000	66.000000
50%	3.000000	117.000000	72.000000	23.000000	36.000000
75%	6.000000	140.250000	80.000000	32.000000	127.000000
max	17.000000	199.000000	122.000000	99.000000	846.000000

	BMI	DiabetesPedigreeFunction	Age	Outcome
count	768.000000	768.000000	768.000000	768.000000
mean	31.992578	0.471876	33.240885	0.348982
std	7.884160	0.331329	11.760232	0.476953
min	0.000000	0.078000	21.000000	0.000000
25%	27.300000	0.243750	24.000000	0.000000
50%	32.000000	0.372500	29.000000	0.000000
75%	36.600000	0.626250	41.000000	1.000000
max	67.100000	2.420000	81.000000	1.000000

✓
2s



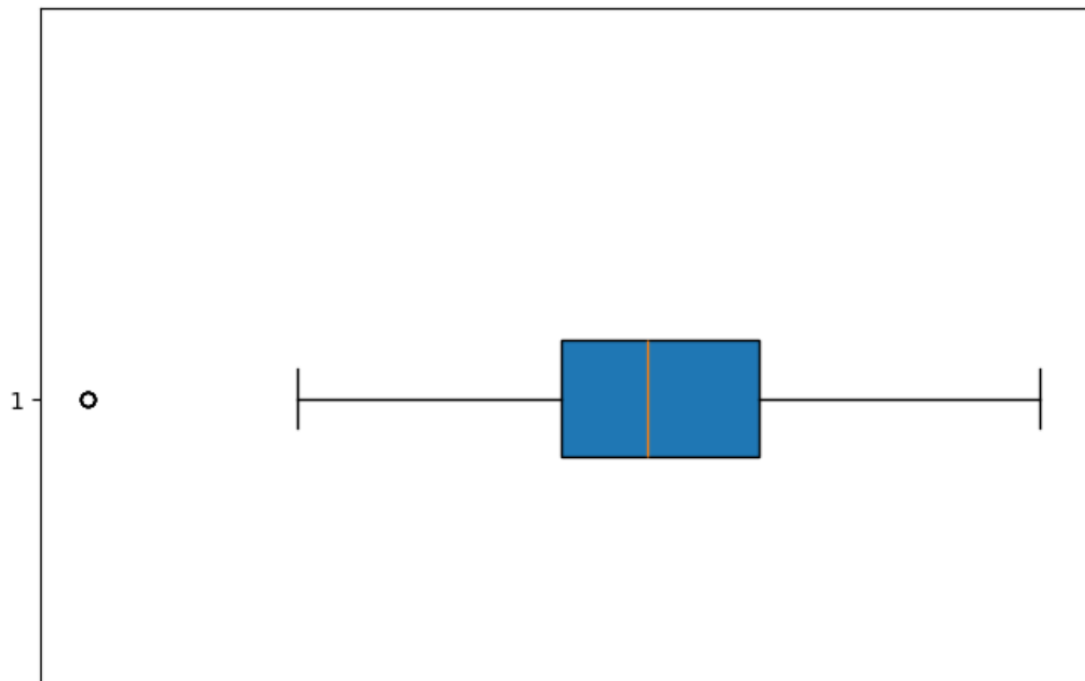
Box Plot of Pregnancies



✓ Connected to Python 3 Google Compute Engine backend

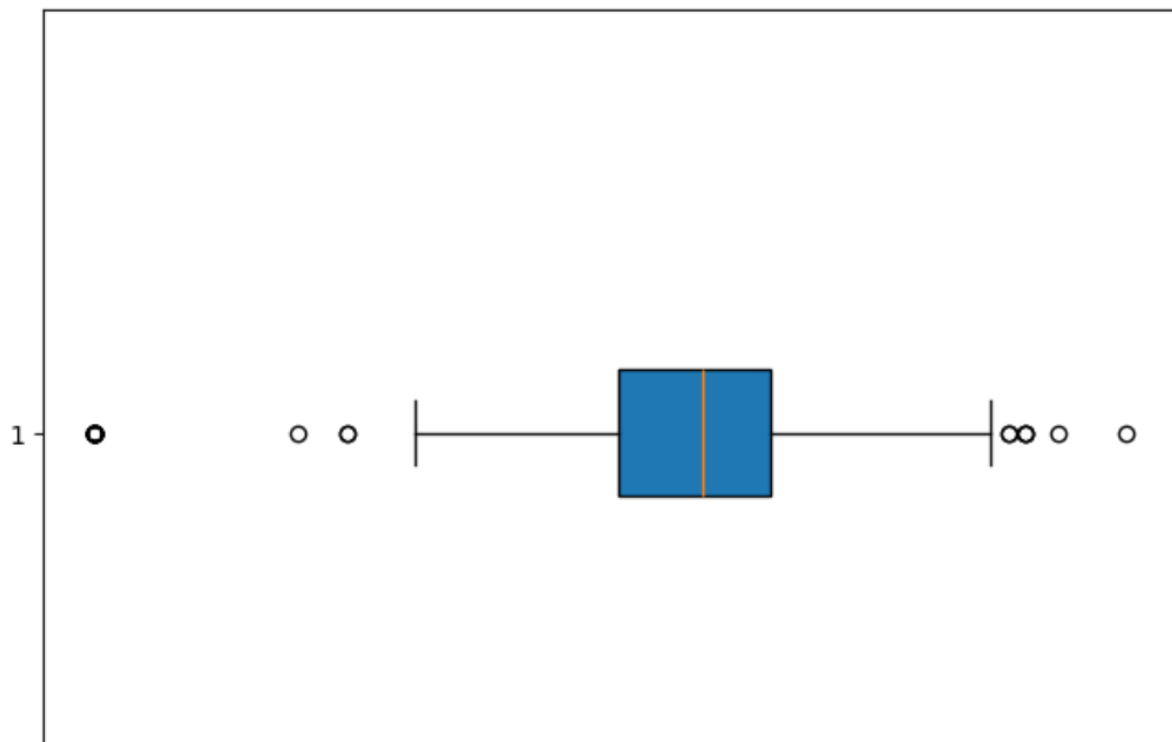


Box Plot of Glucose

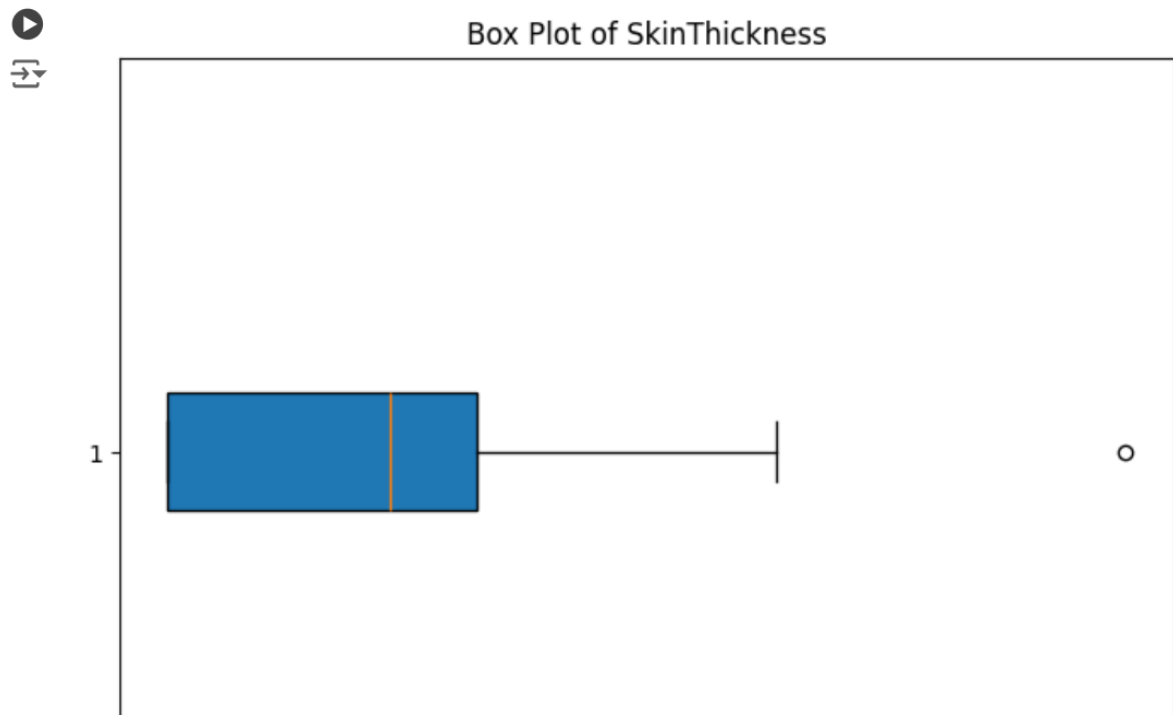


✓ Connected to Python 3 Google Compute Engine backend

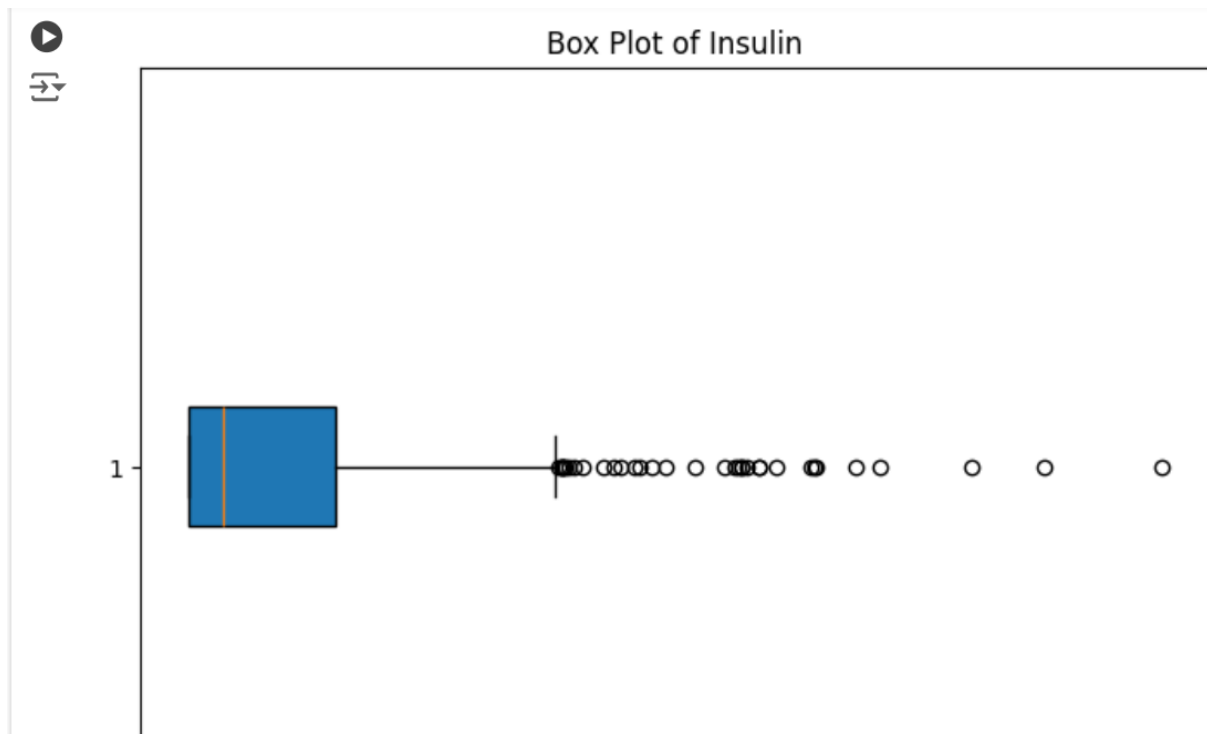
Box Plot of BloodPressure



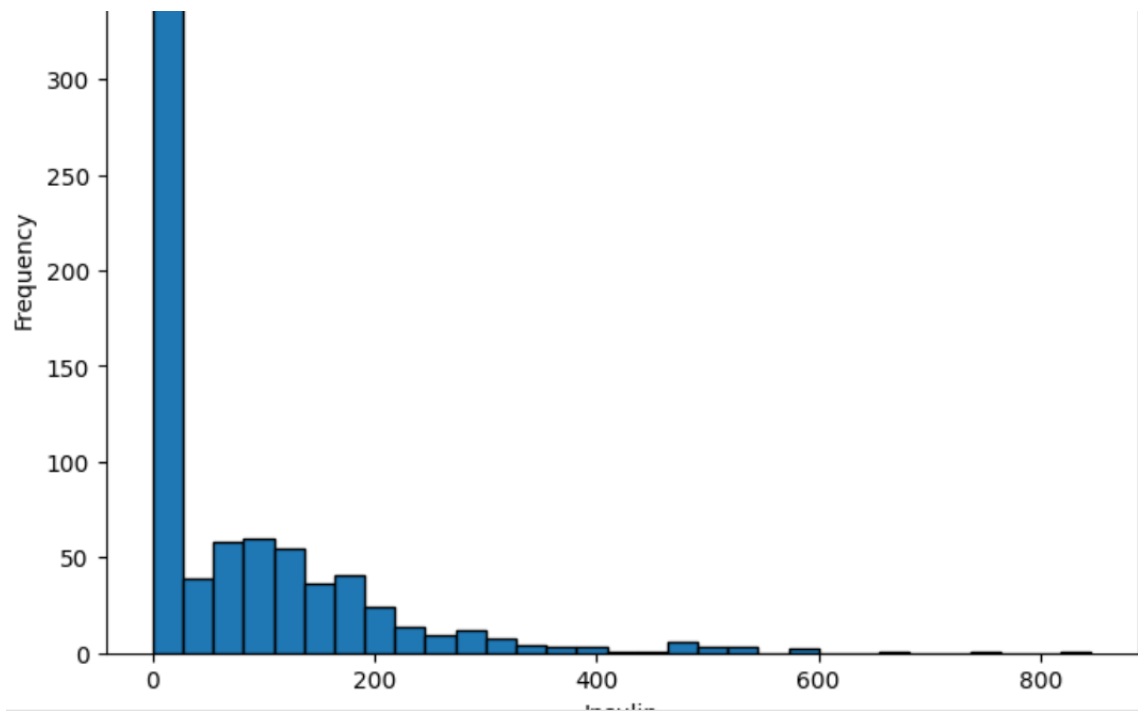
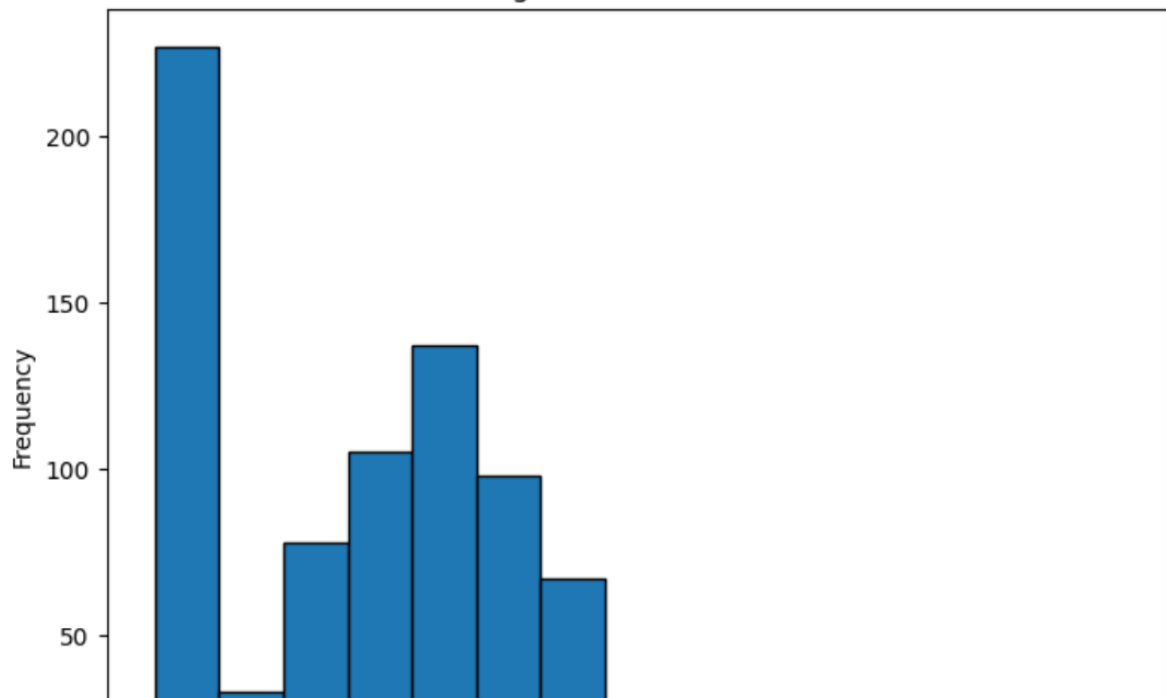
✓ Connected to Python 3 Google Compute Engine backend



✓ Connected to Python 3 Google Compute Engine backend

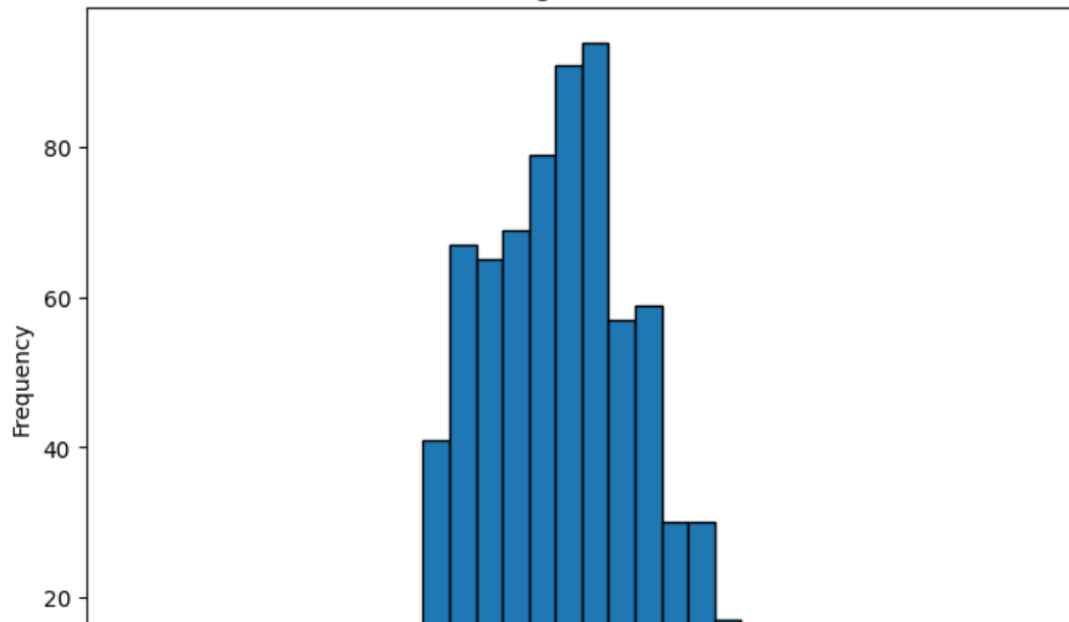


Histogram of SkinThickness

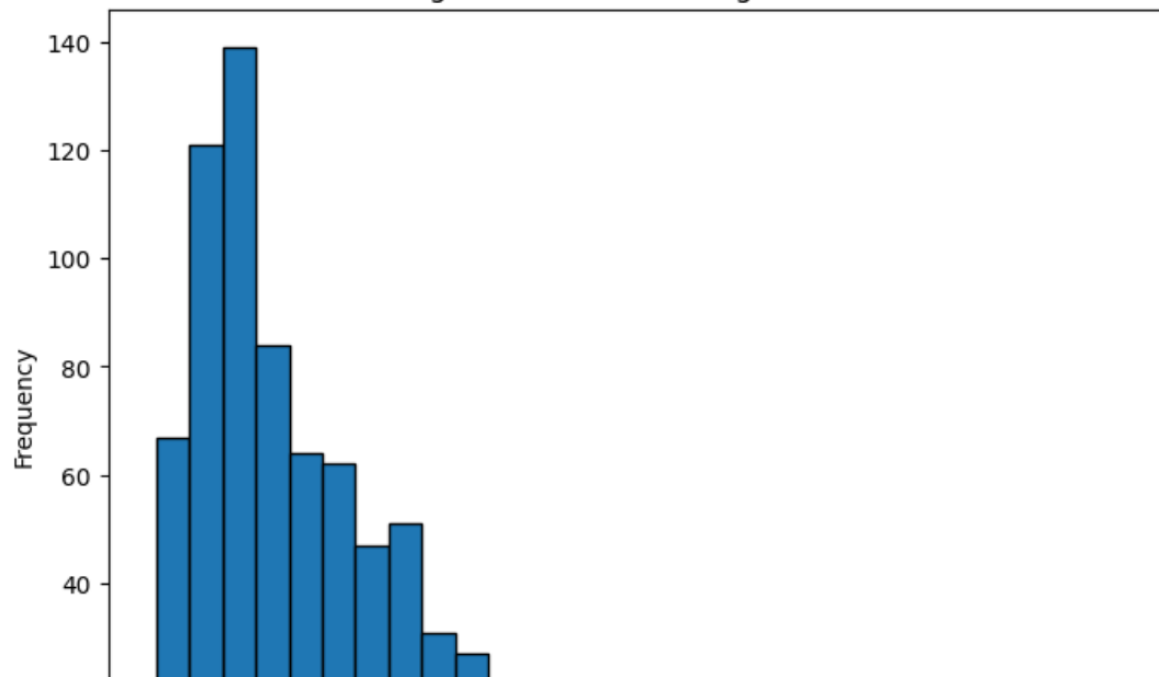




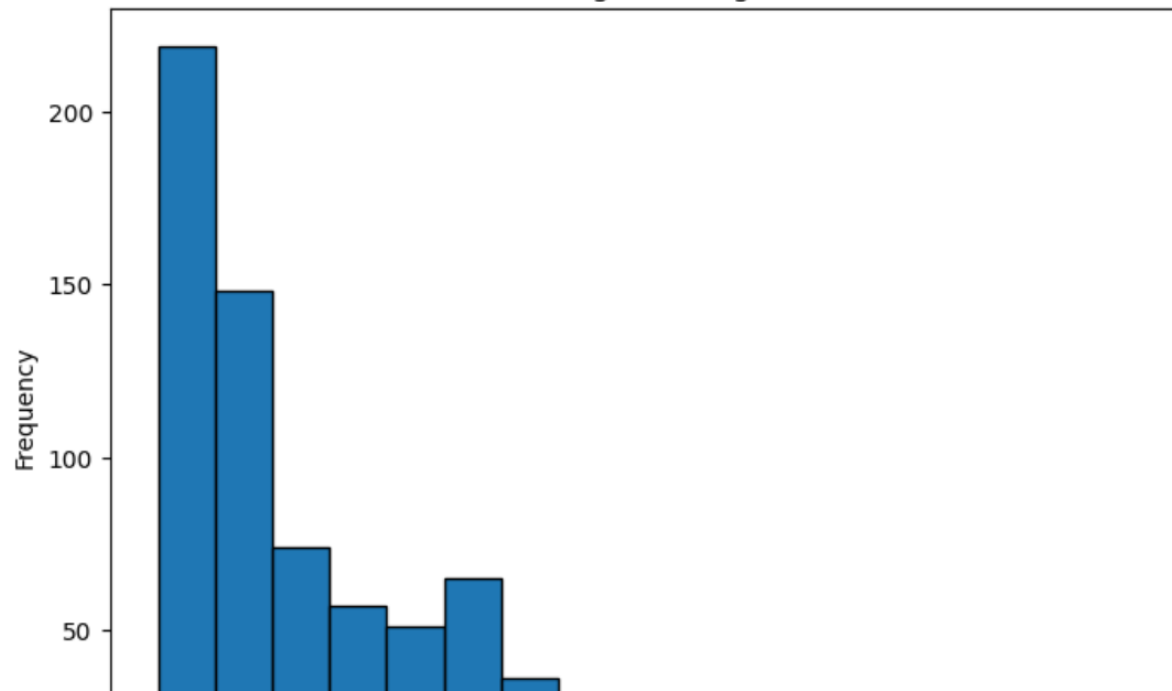
Histogram of BMI



Histogram of DiabetesPedigreeFunction

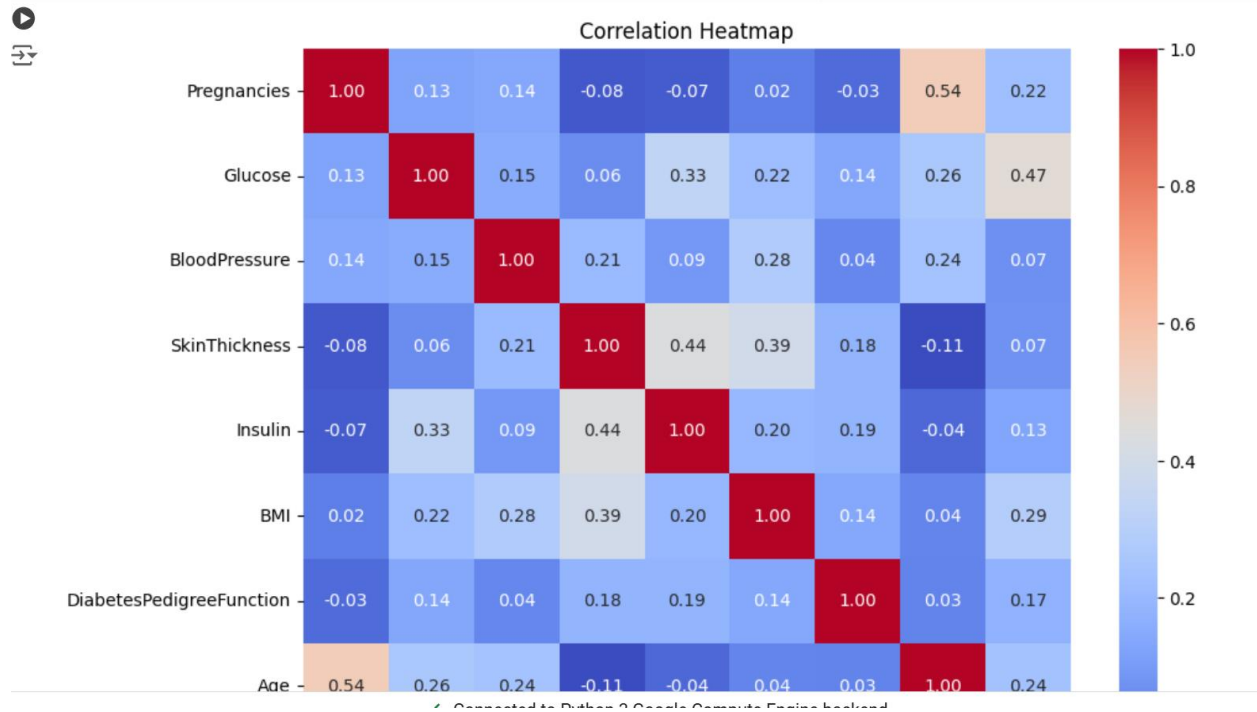


Histogram of Age



Histogram of Outcome





Test passed successfully!

Predictions: [0 1 0 1 1]
Actual labels: [0 0 0 0 0]
Test case passed successfully!

Accuracy of the KNN model on the test set: 67.39%

⇒ Accuracy for k=1: 68.70%
Accuracy for k=2: 72.61%
Accuracy for k=3: 67.39%
Accuracy for k=4: 72.17%
Accuracy for k=5: 68.70%
Accuracy for k=6: 70.00%
Accuracy for k=7: 69.13%
Accuracy for k=8: 71.30%
Accuracy for k=9: 70.00%
Accuracy for k=10: 71.74%
Accuracy for k=11: 73.04%
Accuracy for k=12: 73.91%
Accuracy for k=13: 74.35%
Accuracy for k=14: 72.61%
Accuracy for k=15: 73.91%
Accuracy for k=16: 73.48%
Accuracy for k=17: 72.61%
Accuracy for k=18: 71.30%
Accuracy for k=19: 72.17%
Accuracy for k=20: 70.87%

⇒ Shape of X_train: (538, 8)
Shape of X_test: (230, 8)
Shape of y_train: (538,)
Shape of y_test: (230,)

Accuracy of KNN with Different Values of k

