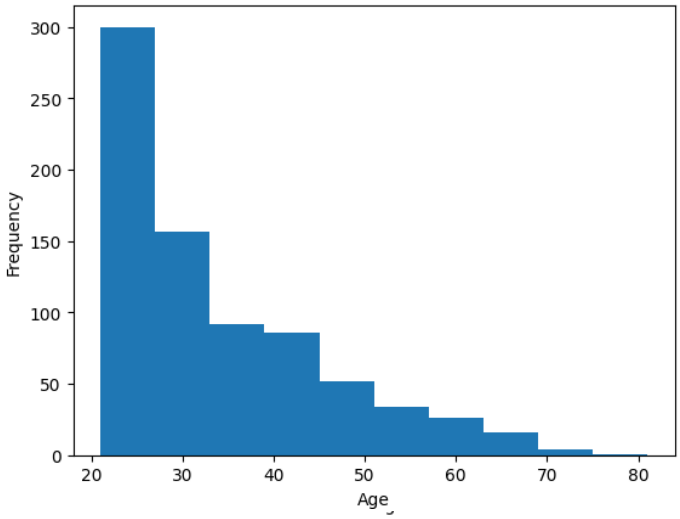
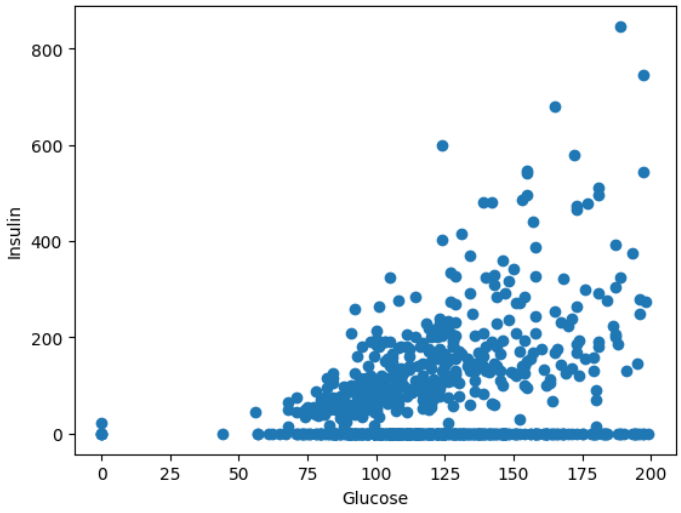
SUPERVISED LEARNING PROJECT

1) Importing the necessary libraries:

```
import numpy as np
import pandas as pd
from sklearn.preprocessing import StandardScaler
from sklearn.model_selection import train_test_split
from sklearn import svm
from sklearn.metrics import accuracy_score
```

2) Loading the dataset using pandas
3) Used Matplotlib to make a couple of viz to detect outliers





4) Removed the outliers from Glucose and Insulin
5) Pre-processed the data using Scaler Standard to fit and standardize the dataframe

6) Split the data into Train and Test
7) Used Linear model using SVM Classifier
8) Checked for accuracy of the model
9) Made a predictive system

```
input_data = (5,166,72,19,175,25.8,0.587,51)

# changing the input_data to numpy array
input_data_as_numpy_array = np.asarray(input_data)

# reshape the array as we are predicting for one instance
input_data_reshaped = input_data_as_numpy_array.reshape(1,-1)

# standardize the input data
std_data = scaler.transform(input_data_reshaped)
print(std_data)

prediction = classifier.predict(std_data)
print(prediction)

if (prediction[0] == 0):
  print('The person is not diabetic')
else:
  print('The person is diabetic')
```

```
[[ 0.53028362  1.44231151  0.11350595 -0.96491164  0.21252316 -1.01086809
   0.18323037  1.99672344]]
[1]
The person is diabetic
```