Ans 1)

Number of Times Pregnant : Not Normally Distributed

Plasma Glucose Concentration : Normally Distributed

Diastolic Blood pressure : Normally Distributed

Triceps skin fold thickness : Not Normally Distributed

2-Hour serum insulin : Not Normally Distributed

BMI : Normally Distributed

Diabetes pedigree function : Not Normally Distributed

Age : Not Normally Distributed

From the distribution of the plots, we can come to a conclusion that majority of the data is not distributed in a normal form. Apart from Plasm glucose, BMI and diastolic blood pressure, the rest of the data isn’t normally distributed as seen by the histograms and bar plots.

Ans 2) After comparing and correlating all the attributes with the class variable, Plasma Glucose Concentration has the highest correlation with the class variable.

Ans 3) The attributes having the highest mutual correlation are attribute1 and attribute 8 which are “No. of times pregnant” and “age”. The max correlation is 0.54