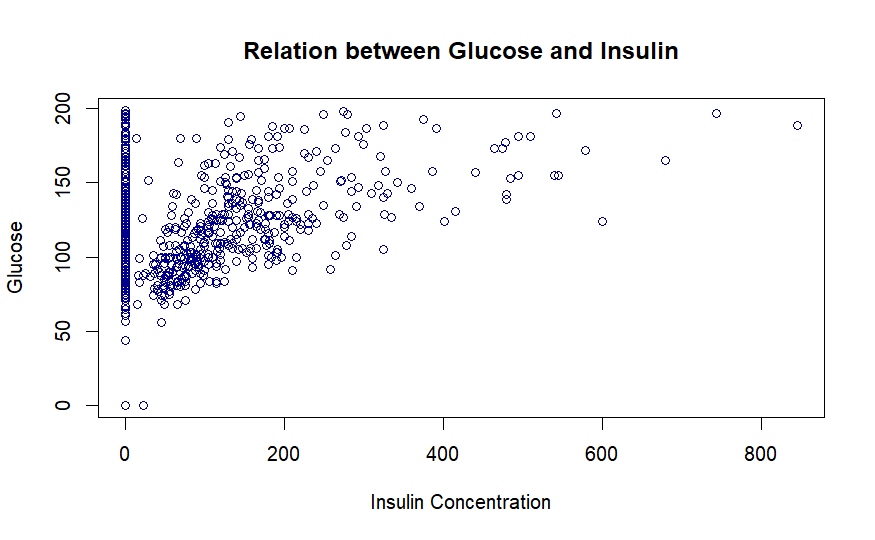
> plot 1

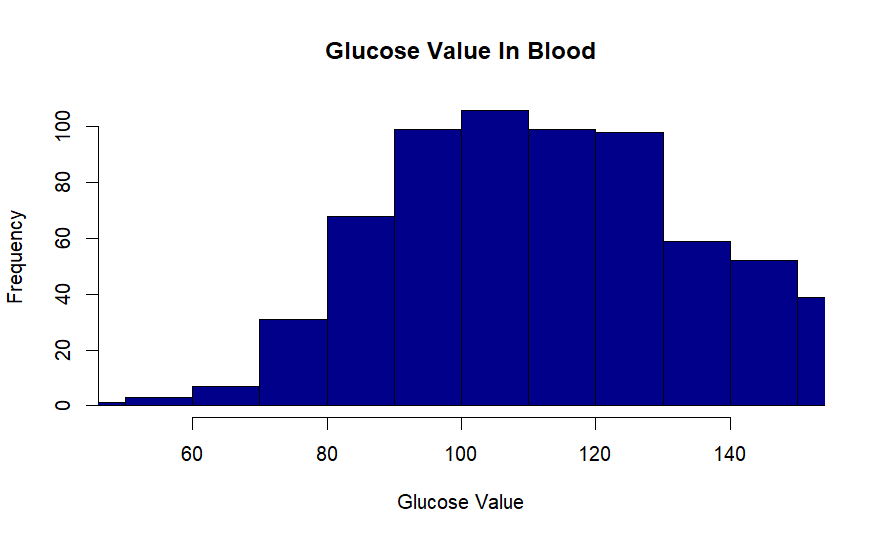
plot(x = data$Insulin,y = data$Glucose,xlab = "Insulin Concentration",ylab = "Glucose",main = "Relation between Glucose and Insulin", col = "darkblue")



Observation 1: The glucose data shows an insulin saturation level of 0. then assumes to gradually rise as the concentration of insulin does.

>plot 2

hist(x = data$Glucose ,main= "Glucose Value In Blood",xlab = "Glucose Value",xlim = c(50,150),breaks = 20,col = "darkblue",border = "black")



Observation 2: The graph starts out rising till the blood glucose level reaches 100 before gradually declining as the glucose level rises. (Acts as normal distribution graph)

> Plot 3

boxplot(data$BloodPressure, data = data, notch = FALSE, varwidth = FALSE, xlab = "Blood Pressure",ylab = "Frequency", main ="Blood Pressure Measurements")

Chart, box and whisker chart

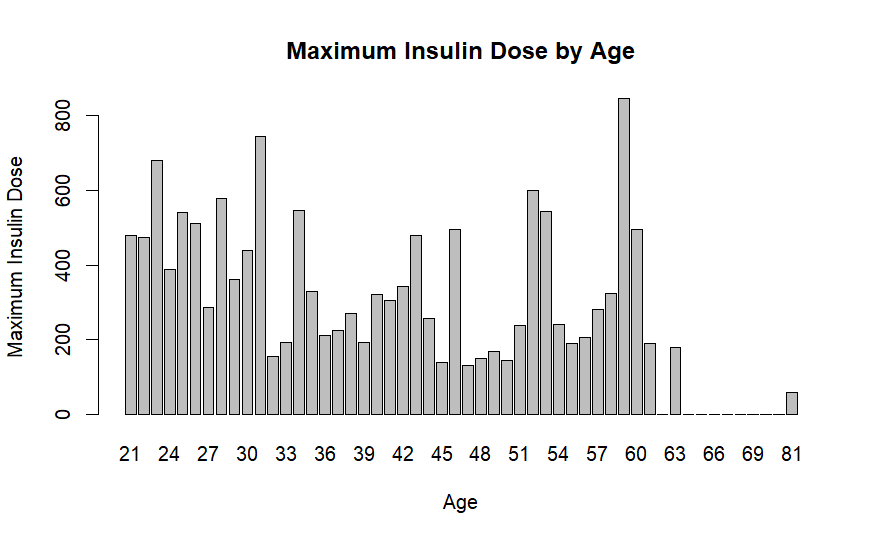
Description automatically generated

Observation 3: The plots reveal 3 outliers below the normally distributed graph at 40 frequencies and 4 outliers above the blood pressure at 100 frequencies. (Seems normally distributed)

>plot 4

max\_insulin <- aggregate(data$Insulin, by = list(data$Age), FUN = max)

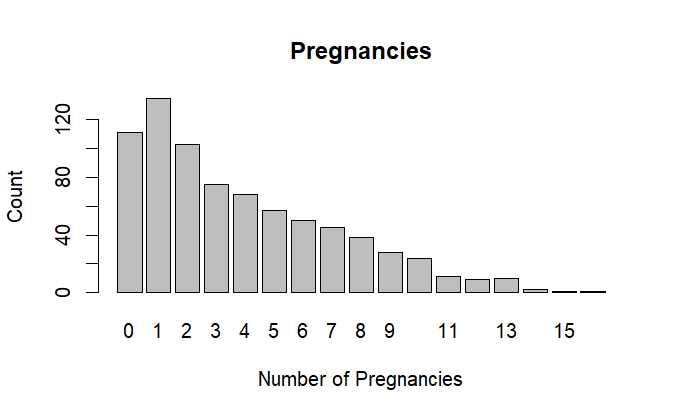
barplot(max\_insulin$x, names.arg = max\_insulin$Group.1,xlab = "Age",ylab = "Maximum Insulin Dose",main = "Maximum Insulin Dose by Age")



Observation 4: In general, the graph indicates a consistent frequency of maximum insulin dose by age, with the highest maximum insulin dose being shown at about age 59 - 60.

>Plot 5

pregnancies\_table <- table(data$Pregnancies)barplot(pregnancies\_table, main="Pregnancies", xlab="Number of Pregnancies", ylab="Count")



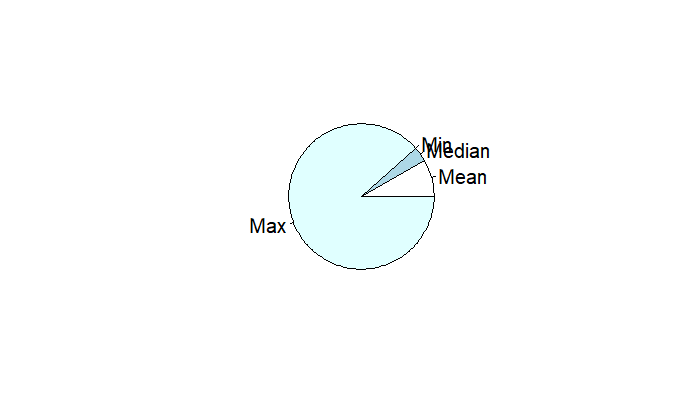
Observation 5: The higher the count, the fewer the pregnancies. Right skewness is displayed on the graph.

>Plot 6

insulin\_summary <- c(mean = mean(data$Insulin), median = median(data$Insulin), min = min(data$Insulin), max = max(data$Insulin))

# Mean = 79.8 # Median = 30.5 # Min = 0 # Max = 846

labels <- c(paste("Mean:", insulin\_summary[1]),paste("Median:", insulin\_summary[2]),paste("Man:", insulin\_summary[3]),paste("Max:", insulin\_summary[4]))

pie(insulin\_summary, labels)

Observation 6: The Pie chart represents some Statistics.

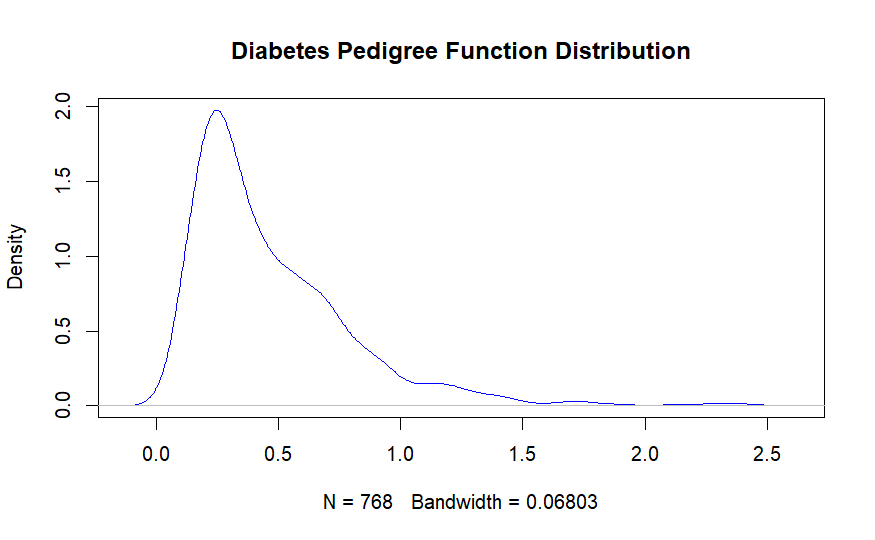
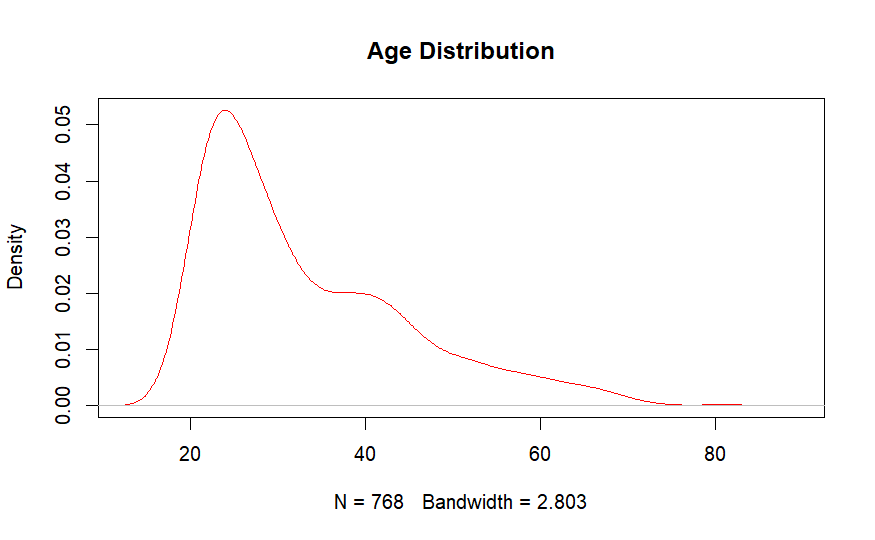
( # Mean = 79.8 # Median = 30.5 # Min = 0 # Max = 846).

>Plot 7

# Create the density plot for Diabetes Pedigree Function

plot(density(data$DiabetesPedigreeFunction), main = "Diabetes Pedigree Function Distribution" , col = "blue")

# Create the density plot for Age

plot(density(data$Age), main = "Age Distribution", col ="red")

Observation 7: Right skewness can be seen in both graphs. Additionally, the density (bandwidth) of the age distribution is larger than that of the diabetes pedigree function distribution.

>Plot 8

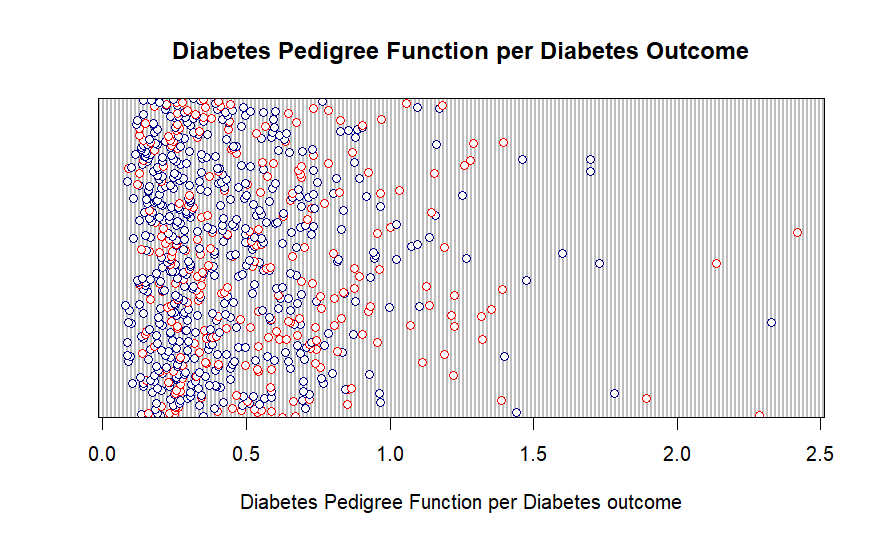
colors <- ifelse(data$Outcome == 1, "red", "darkblue")

# Create a dot chart

dotchart(data$DiabetesPedigreeFunction,

xlab = "Diabetes Pedigree Function per Diabetes outcome",

main = "Diabetes Pedigree Function per Diabetes Outcome",

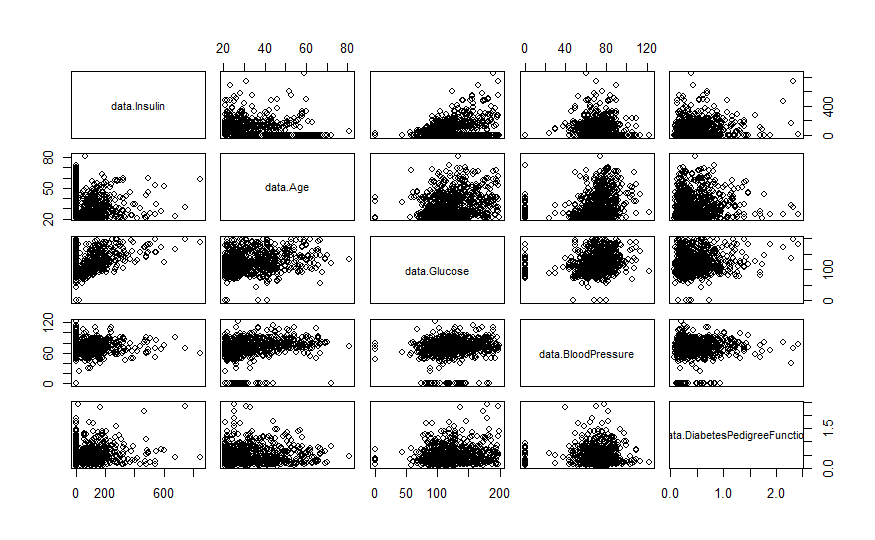
 col = colors)

Observation 8: The graph illustrates the relationship between diabetes outcome and diabetes pedigree function. ( Well Scattered)

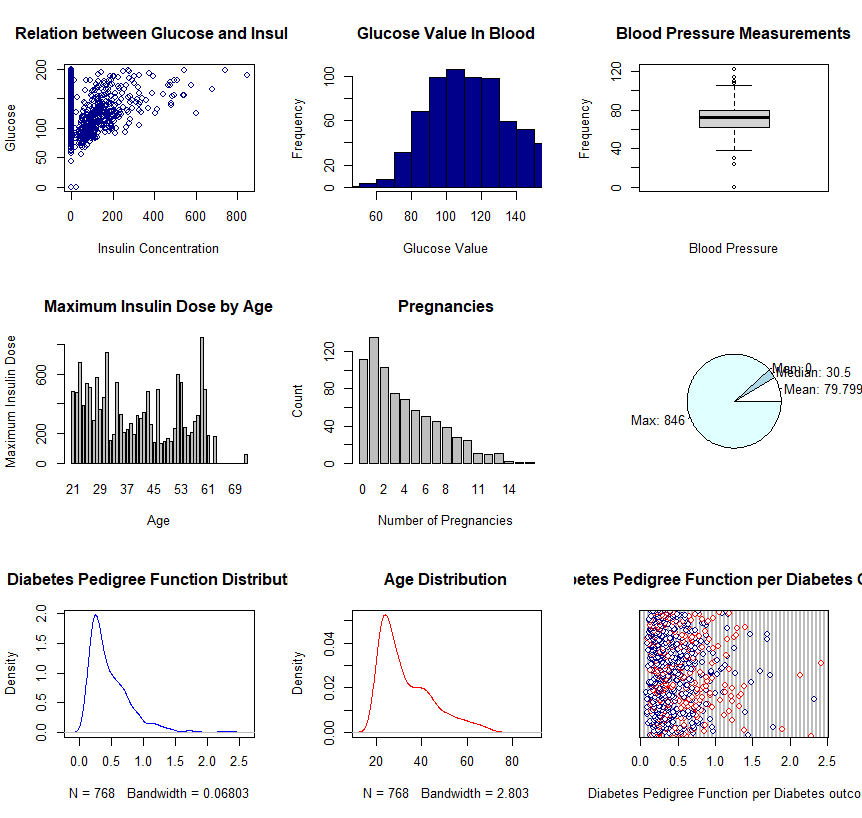
>Plot 9

sample\_data <- data.frame(data$Insulin, data$Age, data$Glucose, data$BloodPressure,data$DiabetesPedigreeFunction)

#create pairs plot

pairs( sample\_data )

Observation 9: The variables Insulin, Age, Glucose, Blood Pressure, and DiabetesPedigreeFunction are shown in a paired scatter plot. The figure that results displays the correlations between each variable and each other variable.

>Summary