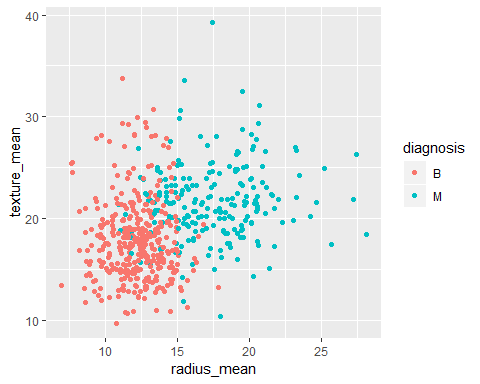
R Notebook

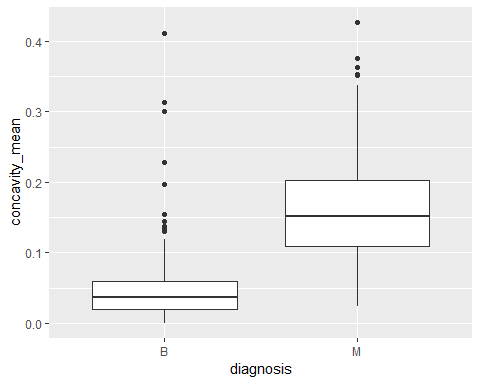
#Install these packages from the menu Tools>Install packages.# To load the packages use the following library function  
library(ggplot2)  
library(readxl)  
Breast\_cancer\_dataset<-read\_excel("breast\_cancer.xlsx")  
summary.data.frame(Breast\_cancer\_dataset)

## id diagnosis radius\_mean texture\_mean   
## Min. : 8670 Length:569 Min. : 6.981 Min. : 9.71   
## 1st Qu.: 869218 Class :character 1st Qu.:11.700 1st Qu.:16.17   
## Median : 906024 Mode :character Median :13.370 Median :18.84   
## Mean : 30371831 Mean :14.127 Mean :19.29   
## 3rd Qu.: 8813129 3rd Qu.:15.780 3rd Qu.:21.80   
## Max. :911320502 Max. :28.110 Max. :39.28   
## perimeter\_mean area\_mean smoothness\_mean compactness\_mean   
## Min. : 43.79 Min. : 143.5 Min. :0.05263 Min. :0.01938   
## 1st Qu.: 75.17 1st Qu.: 420.3 1st Qu.:0.08637 1st Qu.:0.06492   
## Median : 86.24 Median : 551.1 Median :0.09587 Median :0.09263   
## Mean : 91.97 Mean : 654.9 Mean :0.09636 Mean :0.10434   
## 3rd Qu.:104.10 3rd Qu.: 782.7 3rd Qu.:0.10530 3rd Qu.:0.13040   
## Max. :188.50 Max. :2501.0 Max. :0.16340 Max. :0.34540   
## concavity\_mean concave points\_mean symmetry\_mean   
## Min. :0.00000 Min. :0.00000 Min. :0.1060   
## 1st Qu.:0.02956 1st Qu.:0.02031 1st Qu.:0.1619   
## Median :0.06154 Median :0.03350 Median :0.1792   
## Mean :0.08880 Mean :0.04892 Mean :0.1812   
## 3rd Qu.:0.13070 3rd Qu.:0.07400 3rd Qu.:0.1957   
## Max. :0.42680 Max. :0.20120 Max. :0.3040   
## fractal\_dimension\_mean radius\_se texture\_se perimeter\_se   
## Min. :0.04996 Min. :0.1115 Min. :0.3602 Min. : 0.757   
## 1st Qu.:0.05770 1st Qu.:0.2324 1st Qu.:0.8339 1st Qu.: 1.606   
## Median :0.06154 Median :0.3242 Median :1.1080 Median : 2.287   
## Mean :0.06280 Mean :0.4052 Mean :1.2169 Mean : 2.866   
## 3rd Qu.:0.06612 3rd Qu.:0.4789 3rd Qu.:1.4740 3rd Qu.: 3.357   
## Max. :0.09744 Max. :2.8730 Max. :4.8850 Max. :21.980   
## area\_se smoothness\_se compactness\_se concavity\_se   
## Min. : 6.802 Min. :0.001713 Min. :0.002252 Min. :0.00000   
## 1st Qu.: 17.850 1st Qu.:0.005169 1st Qu.:0.013080 1st Qu.:0.01509   
## Median : 24.530 Median :0.006380 Median :0.020450 Median :0.02589   
## Mean : 40.337 Mean :0.007041 Mean :0.025478 Mean :0.03189   
## 3rd Qu.: 45.190 3rd Qu.:0.008146 3rd Qu.:0.032450 3rd Qu.:0.04205   
## Max. :542.200 Max. :0.031130 Max. :0.135400 Max. :0.39600   
## concave points\_se symmetry\_se fractal\_dimension\_se  
## Min. :0.000000 Min. :0.007882 Min. :0.0008948   
## 1st Qu.:0.007638 1st Qu.:0.015160 1st Qu.:0.0022480   
## Median :0.010930 Median :0.018730 Median :0.0031870   
## Mean :0.011796 Mean :0.020542 Mean :0.0037949   
## 3rd Qu.:0.014710 3rd Qu.:0.023480 3rd Qu.:0.0045580   
## Max. :0.052790 Max. :0.078950 Max. :0.0298400   
## radius\_worst texture\_worst perimeter\_worst area\_worst   
## Min. : 7.93 Min. :12.02 Min. : 50.41 Min. : 185.2   
## 1st Qu.:13.01 1st Qu.:21.08 1st Qu.: 84.11 1st Qu.: 515.3   
## Median :14.97 Median :25.41 Median : 97.66 Median : 686.5   
## Mean :16.27 Mean :25.68 Mean :107.26 Mean : 880.6   
## 3rd Qu.:18.79 3rd Qu.:29.72 3rd Qu.:125.40 3rd Qu.:1084.0   
## Max. :36.04 Max. :49.54 Max. :251.20 Max. :4254.0   
## smoothness\_worst compactness\_worst concavity\_worst concave points\_worst  
## Min. :0.07117 Min. :0.02729 Min. :0.0000 Min. :0.00000   
## 1st Qu.:0.11660 1st Qu.:0.14720 1st Qu.:0.1145 1st Qu.:0.06493   
## Median :0.13130 Median :0.21190 Median :0.2267 Median :0.09993   
## Mean :0.13237 Mean :0.25427 Mean :0.2722 Mean :0.11461   
## 3rd Qu.:0.14600 3rd Qu.:0.33910 3rd Qu.:0.3829 3rd Qu.:0.16140   
## Max. :0.22260 Max. :1.05800 Max. :1.2520 Max. :0.29100   
## symmetry\_worst fractal\_dimension\_worst  
## Min. :0.1565 Min. :0.05504   
## 1st Qu.:0.2504 1st Qu.:0.07146   
## Median :0.2822 Median :0.08004   
## Mean :0.2901 Mean :0.08395   
## 3rd Qu.:0.3179 3rd Qu.:0.09208   
## Max. :0.6638 Max. :0.20750

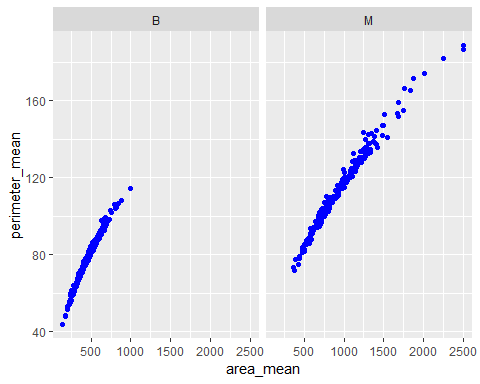
#This graph depicts a correlation of tumor mean radius and mean texture with the type of cancer diagnosis.  
  
ggplot(Breast\_cancer\_dataset, aes(x=radius\_mean,y=texture\_mean,colour=diagnosis )) +geom\_point()



#The boxplot portrays that patients with a malignant diagnosis have tumors with a slightly higher mean concavity, compared to patients with a benign diagnosis.  
  
ggplot(Breast\_cancer\_dataset, aes(diagnosis,concavity\_mean)) +geom\_boxplot()



#This graph portrays the distribution of tumor mean area and perimeter as a side-by-side comparison based on the type of cancer diagnosis.  
  
ggplot(Breast\_cancer\_dataset,aes(area\_mean,perimeter\_mean))+geom\_point(colour="blue")+facet\_wrap(~diagnosis)



#The frequency polygon shows that a patients diagnosed with malignant tumors have tumors with a mean perimeter ranging from 75-150. Additionally, most patients diagnosed with a benign tumor had a tumor with a mean perimeter ranging from 50-100.  
  
ggplot(Breast\_cancer\_dataset,aes(perimeter\_mean,colour=diagnosis))+geom\_freqpoly(binwidth=2.5)

