plot cindy

**載入**

library(readr)  
library(ggplot2)

## Warning: 套件 'ggplot2' 是用 R 版本 4.1.2 來建造的

library(dplyr)

##   
## 載入套件：'dplyr'

## 下列物件被遮斷自 'package:stats':  
##   
## filter, lag

## 下列物件被遮斷自 'package:base':  
##   
## intersect, setdiff, setequal, union

library(lattice)

## Warning: 套件 'lattice' 是用 R 版本 4.1.2 來建造的

**讀檔案**

data<-read\_csv("C:\\Users\\Student\\Desktop\\專題\\outlier\_fill.csv")

## Rows: 76722 Columns: 142

## -- Column specification --------------------------------------------------------  
## Delimiter: ","  
## chr (2): Name, Label  
## dbl (140): Sex, Country, Eye\_R\_S\_W, Eye\_R\_B\_W, Eye\_L\_S\_W, Eye\_L\_B\_W, Eye\_R\_H...

##   
## i Use `spec()` to retrieve the full column specification for this data.  
## i Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

face<-data

**確認各類別**

ftable(face$Label)

## boss doctor entertainer ordinary\_people politician sport  
##   
## 3267 565 13697 56966 1783 444

\*\*將actor.singer.model改為entertainer\*

face$Label<-gsub("actor","entertainer",face$Label)  
face$Label<-gsub("singer","entertainer",face$Label)  
face$Label<-gsub("model","entertainer",face$Label)

**確認各類別**

ftable(face$Label)

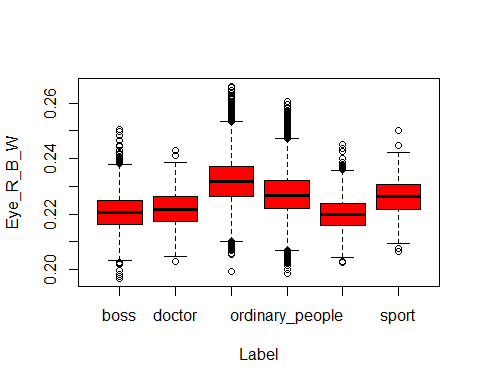
## boss doctor entertainer ordinary\_people politician sport  
##   
## 3267 565 13697 56966 1783 444

**做左右眼寬度的平均**

face$Eye\_LR\_W<-((face$Eye\_R\_B\_W)+(face$Eye\_L\_B\_W))/2

**箱形圖**

boxplot(formula = Eye\_LR\_W ~ Label, # Y ~ X (代表X和Y軸要放的數值)   
 data = face, # 資料  
 xlab = "Label", # X軸名稱  
 ylab = "Eye\_R\_B\_W", # Y軸名稱  
 col = "red") #顏色



**看四分位距**

quantile(face$Eye\_LR\_W)

## 0% 25% 50% 75% 100%   
## 0.1967531 0.2218827 0.2269688 0.2326958 0.2661090

**把眼寬分類**

face$new\_width[face$Eye\_LR\_W>0.2363401] <-"高於75%"

## Warning: Unknown or uninitialised column: `new\_width`.

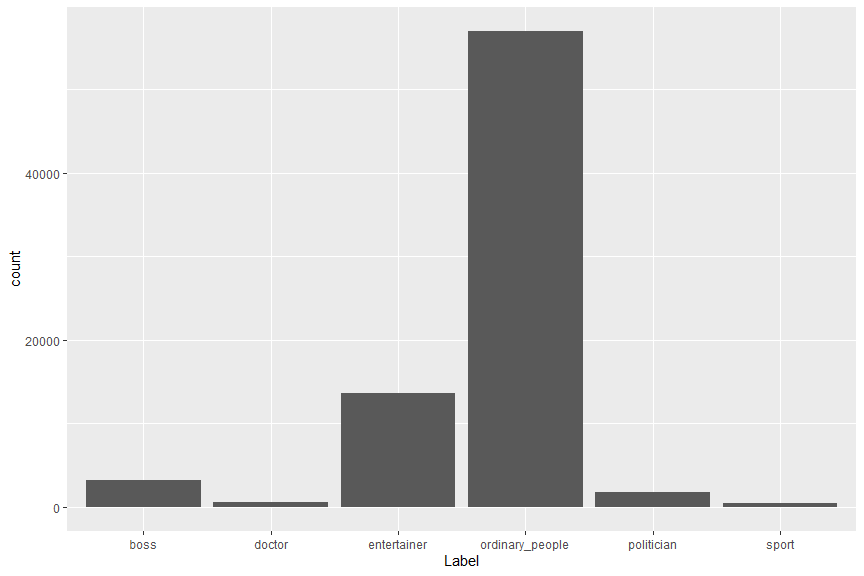
face$new\_width[face$Eye\_LR\_W<=0.2363401 &face$Eye\_LR\_W>=0.2227888] <-"介於25%-75%"  
face$new\_width[face$Eye\_LR\_W<0.2227888] <-"低於25%"  
ftable(face$new\_width)

## 介於25%-75% 低於25% 高於75%  
##   
## 43759 22335 10628

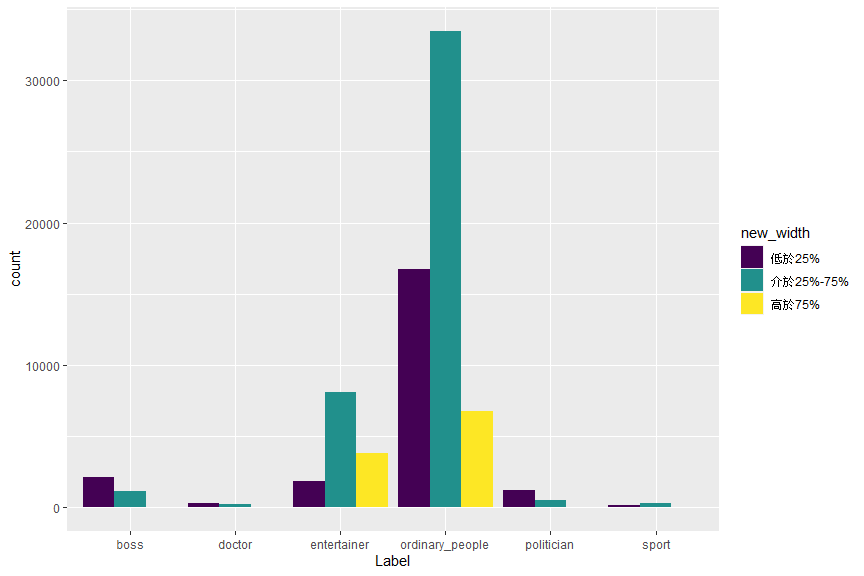
face$new\_width %>% ordered(levels=c("低於25%","介於25%-75%","高於75%")) -> face$new\_width

**長條分布圖**

d<-ggplot(face,aes(Label))  
d+geom\_bar()

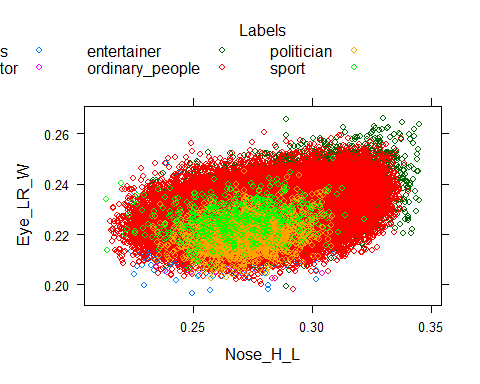


s<-ggplot(face,aes(Label,fill = new\_width))  
s+geom\_bar(position = "dodge")



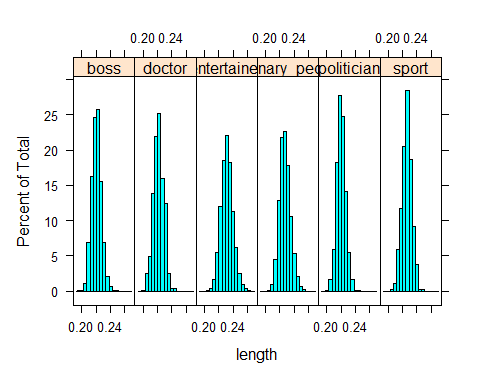
**散佈圖**

xyplot(x=Eye\_LR\_W~Nose\_H\_L, # Eye\_LR\_W放在Y軸，Nose\_H\_L放在X軸  
 data=face,   
 group = Label, # 根據Label，把資料點用顏色區分開來   
   
 # auto.key參數，表示設定標籤與其他資訊  
 auto.key=list(space="top", # 位置在上方   
 columns=5, # 1x5的方式呈現標籤  
 title="Labels", # 標籤名稱  
 cex.title=1) # 標籤字體大小  
)



**直方圖(Histogram)的函式**

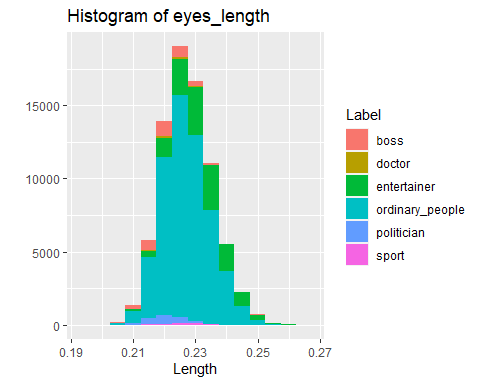
histogram(x= ~ Eye\_LR\_W | Label, # 根據類別(Label)的條件，繪製眼寬(Eye\_LR\_W)的直方圖  
 data=face,   
 xlab="length",   
 layout=c(6,1)) # 以6x1的方式呈現圖表



***使用ggplot2繪圖***

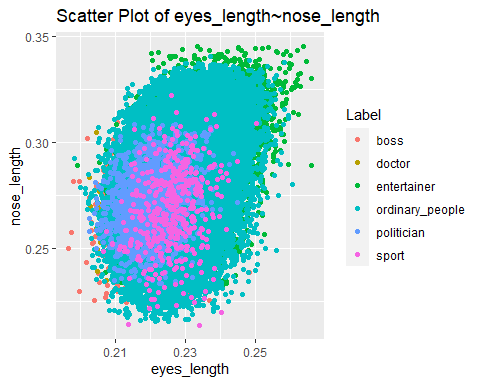
**ggplot2直方圖**

qplot(x=Eye\_LR\_W,   
 data=face,   
 geom="histogram", # 圖形=histogram  
 main = "Histogram of eyes\_length",   
 xlab="Length",   
 binwidth = 0.005, # 每0.005單位為一區隔  
 fill= Label # 以顏色標註Label，複合式的直方圖  
)



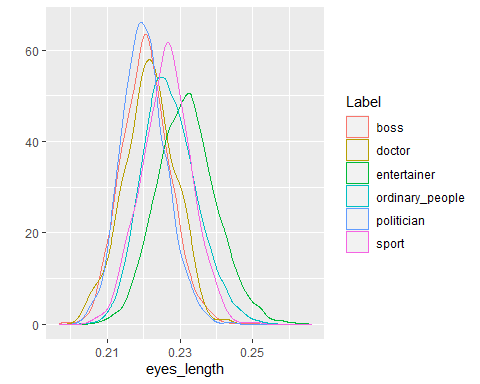
**散布圖(scatter plot)**

qplot(x=Eye\_LR\_W,   
 y=Nose\_H\_L,   
 data=face,   
 geom="point", # 圖形=scatter plot  
 main = "Scatter Plot of eyes\_length~nose\_length",   
 xlab="eyes\_length",   
 ylab="nose\_length",   
 color= Label # 以顏色標註類別，複合式的散布圖  
)



**機率密度圖(density plot)**

qplot(x=Eye\_LR\_W,   
 data=face,   
 geom="density", # 圖形=density  
 xlab="eyes\_length",   
 color= Label # 以顏色標註類別，複合式的機率密度圖  
)



**箱型圖(boxplot)**

qplot(x=Label,   
 y=Eye\_LR\_W,  
 data=face,   
 geom="boxplot", # 圖形=boxplot  
 xlab="Label",   
 color= Label, # 以顏色標註類別，複合式的合鬚圖  
 width = 500  
)

