EXP-0

Name: Dhruv Goel

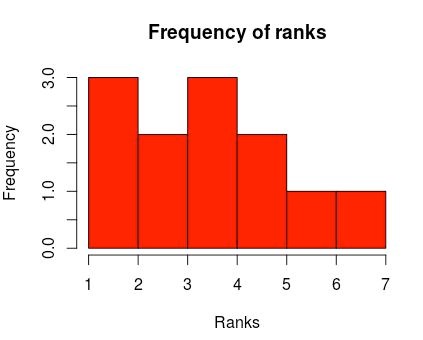
Reg No:16BCE0686

1. Frequency Histogram:

> arr1=c(1,2,3,4,5,3,4,2,5,6,7,4)

> hist(arr1, main='Frequency of ranks', xlab='Ranks', ylab='Frequency',col='red')

Output:



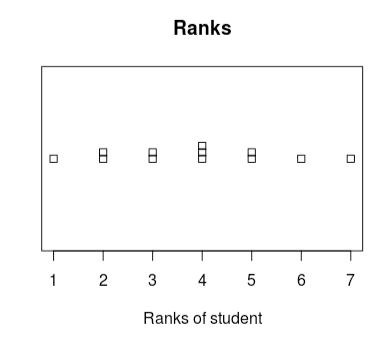
1. Strip Chart:

>stripchart(arr1,method="stack",

+ main='Ranks',

+ xlab='Ranks of student')

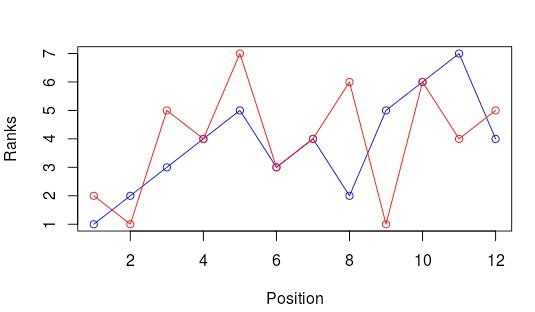
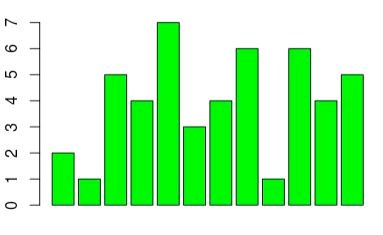
Output:



1. Line Graph:

> plot(arr1,type = "o",col="blue",ylab = "Ranks",xlab = "Position")

Output:



4.Bar Chart

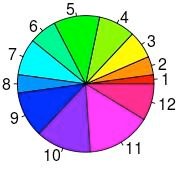
> barplot(arr,col = 'green')

Output:

5.Pie Chart:

> pie(arr1,col = rainbow(length(arr)))

Output:



6.

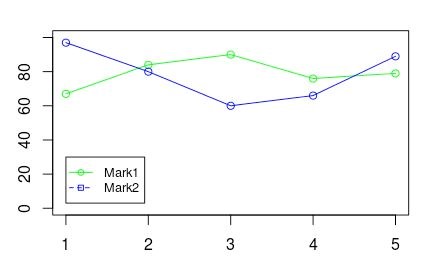
> plot(marksA,type = 'o',col='green',ylim = range(0,100) xlab='Test' ylab='Mark')

> lines(marksB,type = 'o',col='blue')

> legend(1,30,c("Marks of A","Marks of B"),cex =

0.8,col=c("green","blue"),pch=21:22, lty=1:2);

Output:



EXP-2

Data Visualisation

Name: Dhruv Goel

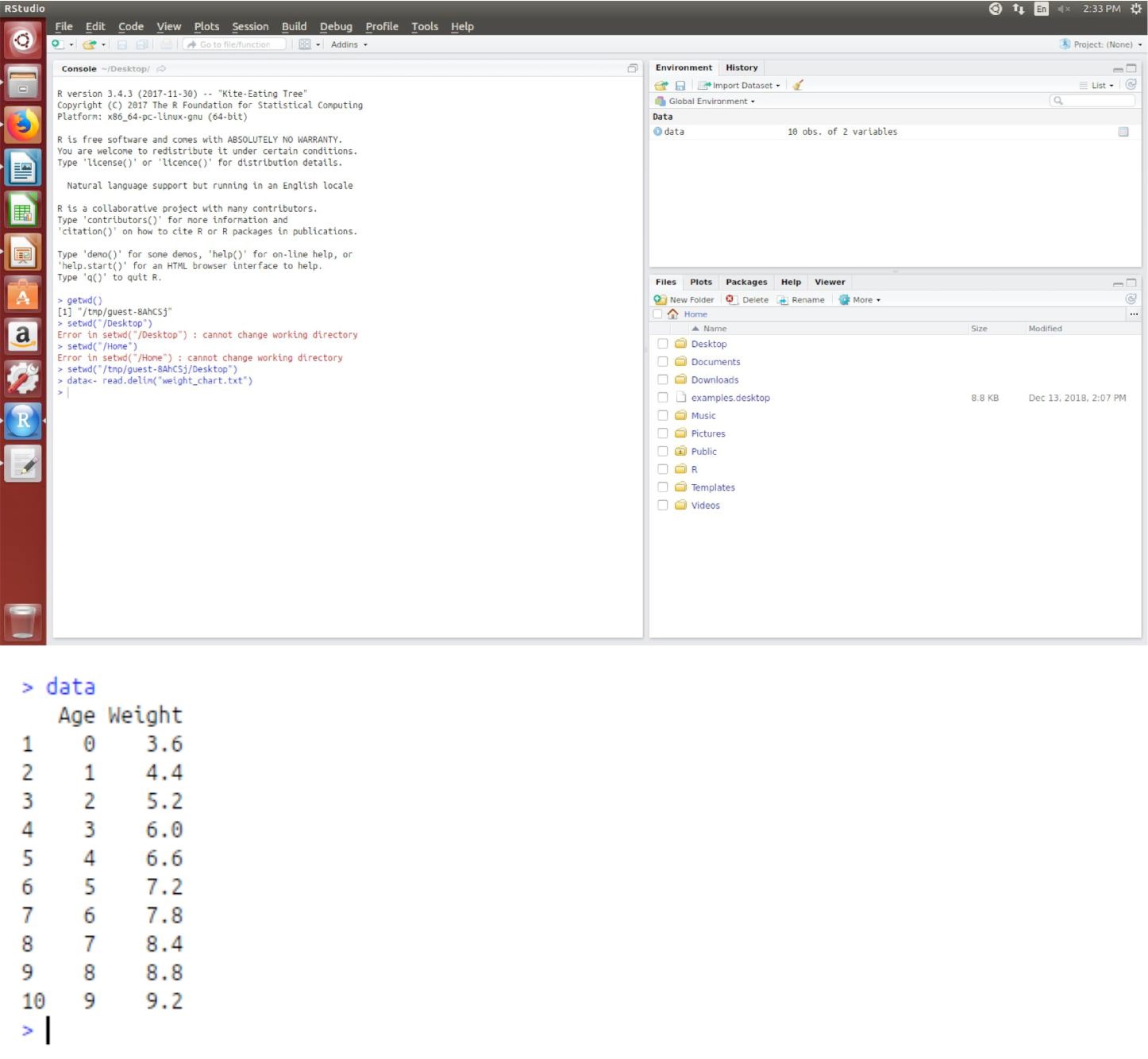
Reg.No: 16BCE0686

1. Weight Chart

Step 1: Read the file

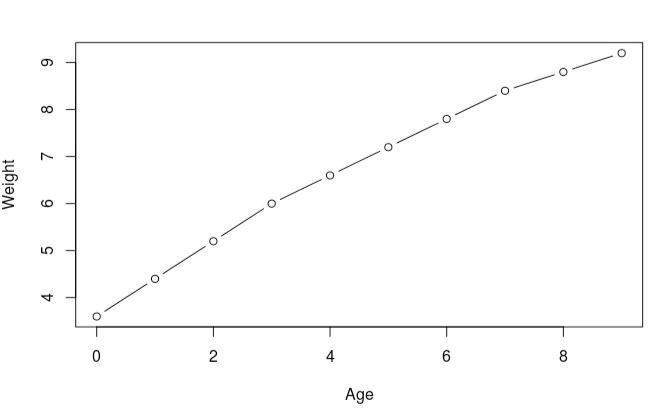
> setwd("/tmp/guest-8AhCSj/Desktop")

> data<- read.delim("weight\_chart.txt")



# Step 2: Plot with type 'b'

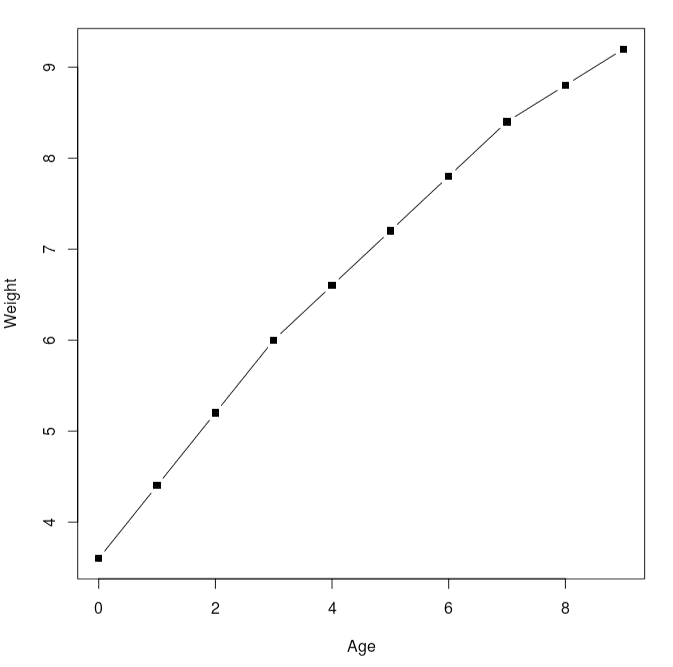
> plot(data,type="b")



Change the point character to be a filled square (pch=15)

> par(pch=15)

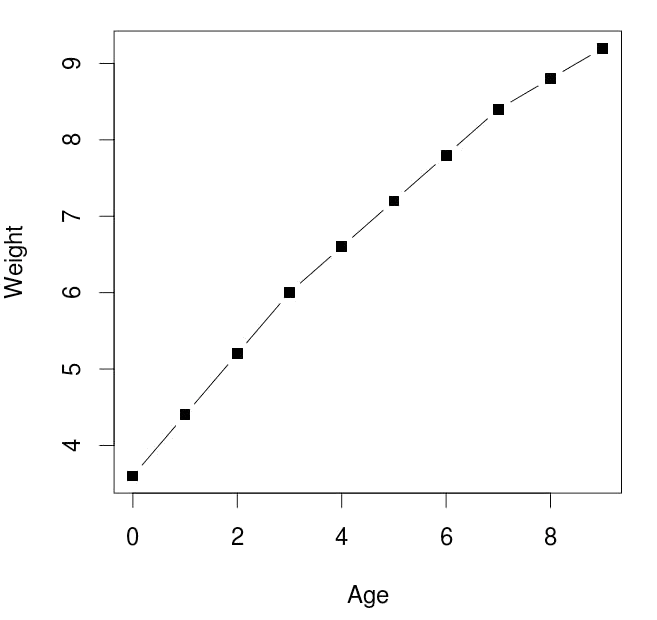
> plot(data,type="b")



Change the plot point size to be 1.5x normal size (cex=1.5)

> par(cex=1.5)

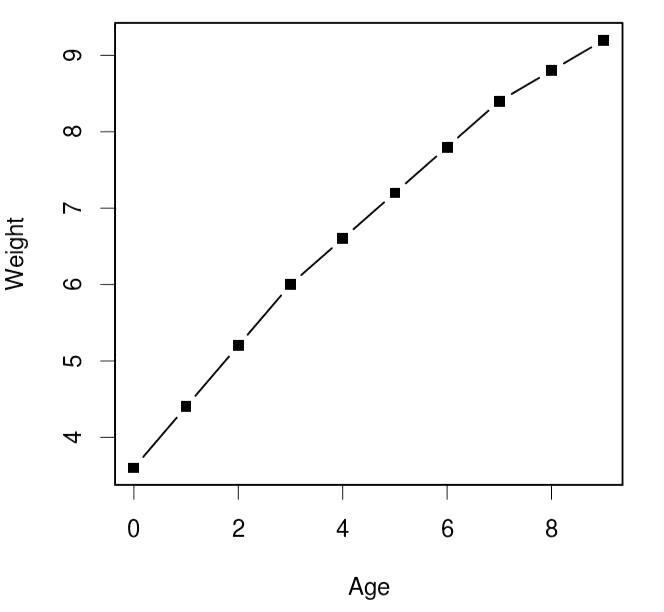
> plot(data,type="b")



Change the line thickness to be twice the default size (lwd=2)

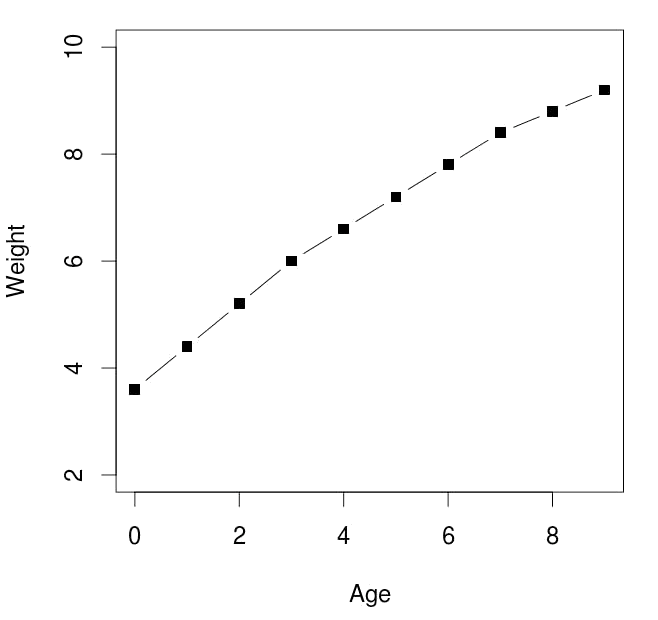
> par(lwd=2)

> plot(data,type="b")



Change the y-axis to scale between 2 and 10kg (ylim=c(2,10))

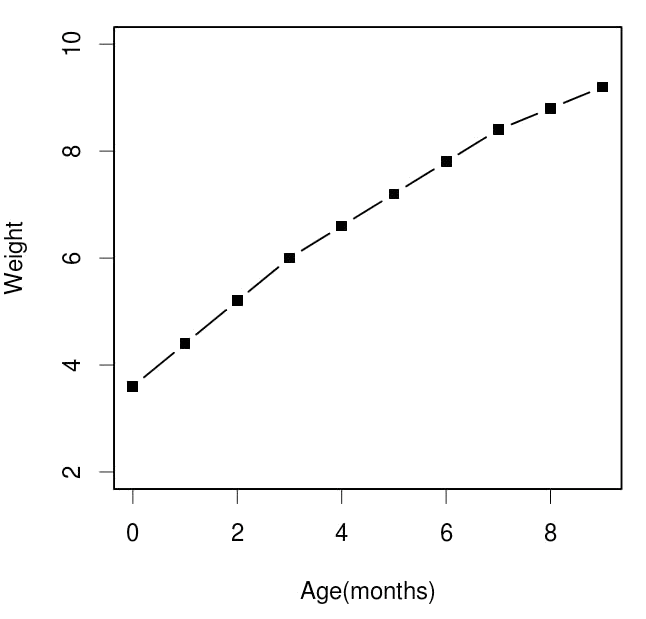
> plot(data,type="b",ylim=c(2,10))



Change the x-axis title to be Age (months) (xlab=”Age (months)”)

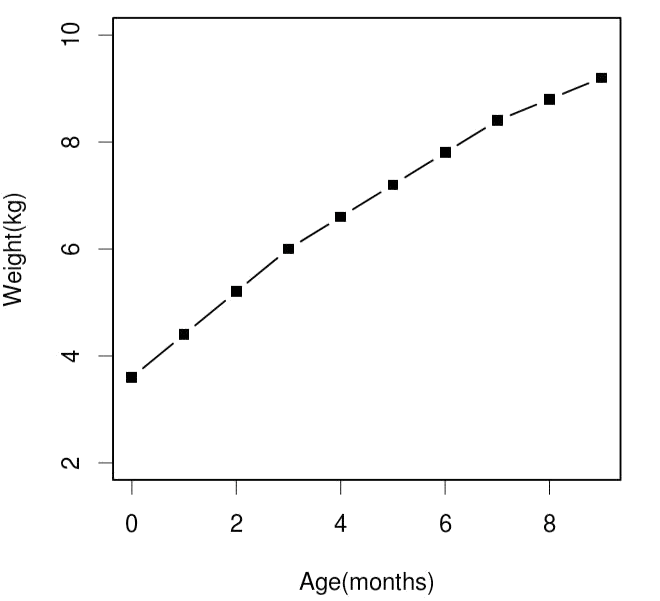
> par(pch=15,cex=1.5,lwd=2)

> plot(data,type="b",ylim=c(2,10),xlab="Age(months)")



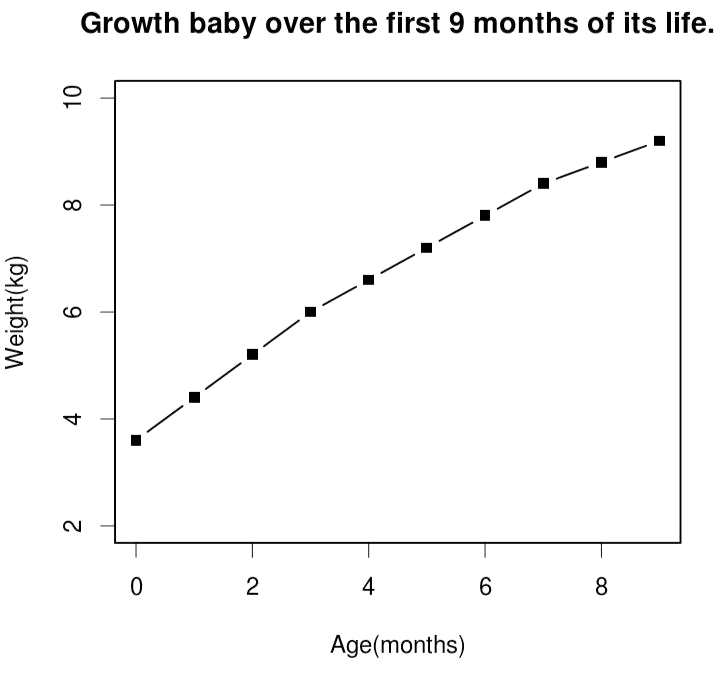
# Change the y-axis title to be Weight (kg) (ylab=”Weight (kg)”)

> plot(data,type="b",ylim=c(2,10),xlab="Age(months)",ylab="Weight(kg)")



# Add a suitable title to the top of the plot (main=”Some title”)

> plot(data,type="b",ylim=c(2,10),xlab="Age(months)",ylab="Weight(kg)",main="Growth baby over the first 9 months of its life.")



2.

The file feature\_counts.txt contains a summary of the number of features of different types in the mouse GRCm38 genome. Load this into a data frame using read.delim and then plot it as a barplot. Note that the data you need to barplot is contained within the Count column of the data frame, so you pass only that single column as the data for the plot.

> data=read.delim("feature\_counts.txt")

> data

Feature Count

1 Messenger RNA 79049

2Coding Sequences 50770

1. Transfer RNAs 26248
2. CpG islands 13840 6 Pseudogenes 5195
3. Micro-RNAs 1638
4. Small nucleolar RNAs 1602

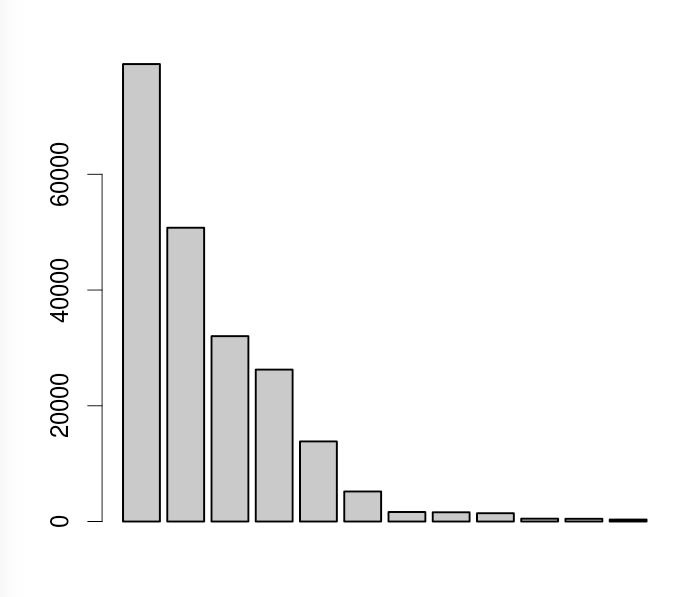
9Small nuclear RNAs 1431

1. Miscellaneous RNA 491
2. Immunoglobulin Segments 474
3. Ribosomal RNAs 341

>

> data1=data$Count

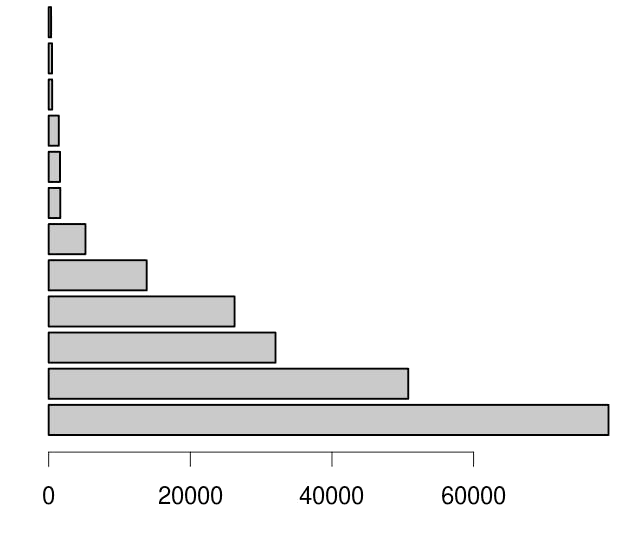
> barplot(data1)



Once you have the basic plot make the following changes:

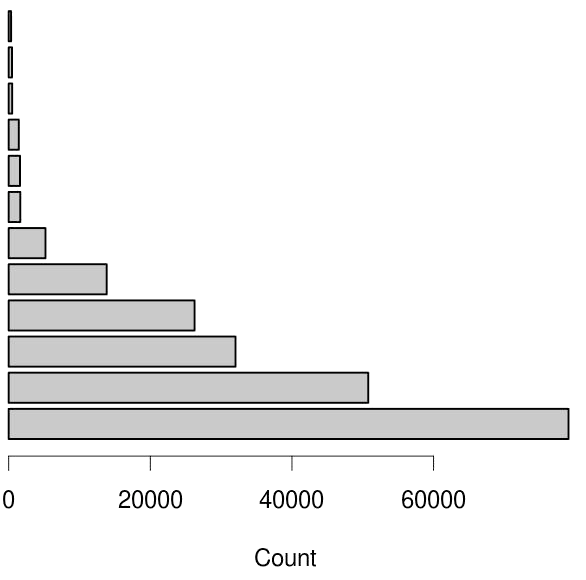
The bars should be horizontal rather than vertical (horiz=TRUE).

> barplot(data1,horiz=TRUE)



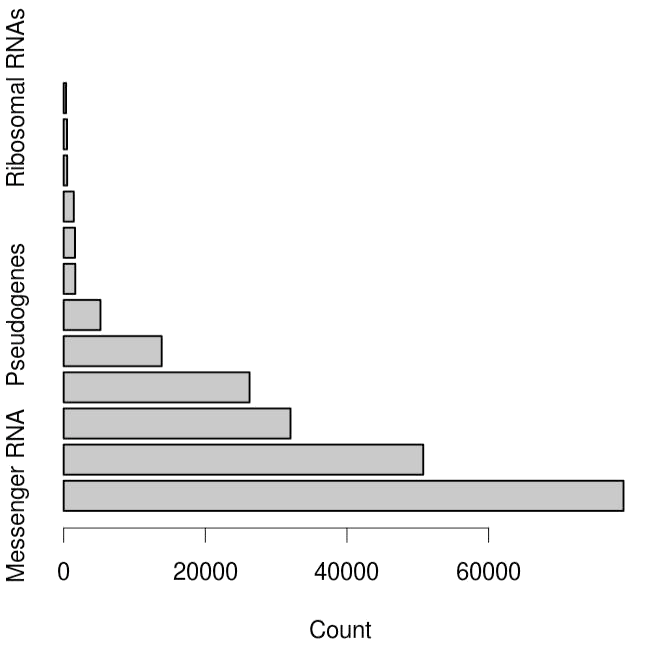
The count axis should be labelled (xlab=”A title”)

> barplot(data1,horiz=TRUE,xlab = "Count")



# The feature names should be added to the y axis. (set names.arg to the Feature column of the data frame)

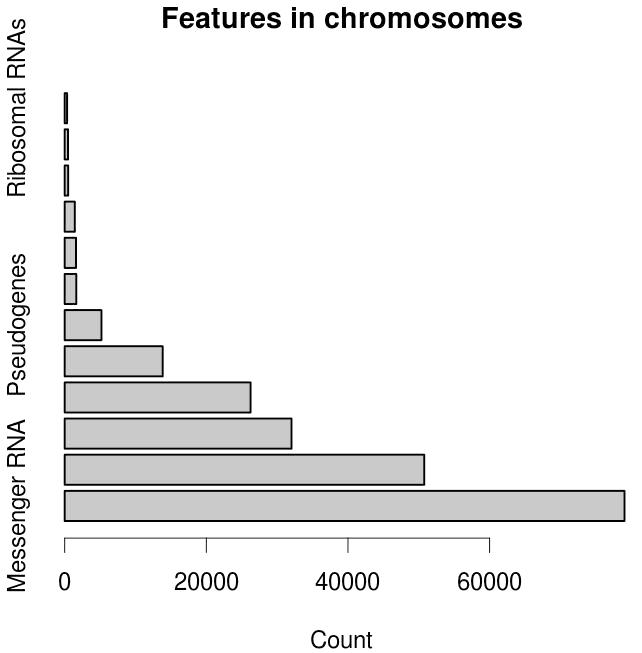
> barplot(data1,horiz=TRUE,xlab = "Count",names.arg=data$Feature)



# The plot should be given a suitable title (main=”Some title”)

> barplot(data1,horiz=TRUE,xlab =

"Count",names.arg=data$Feature,main="Features in chromosomes")

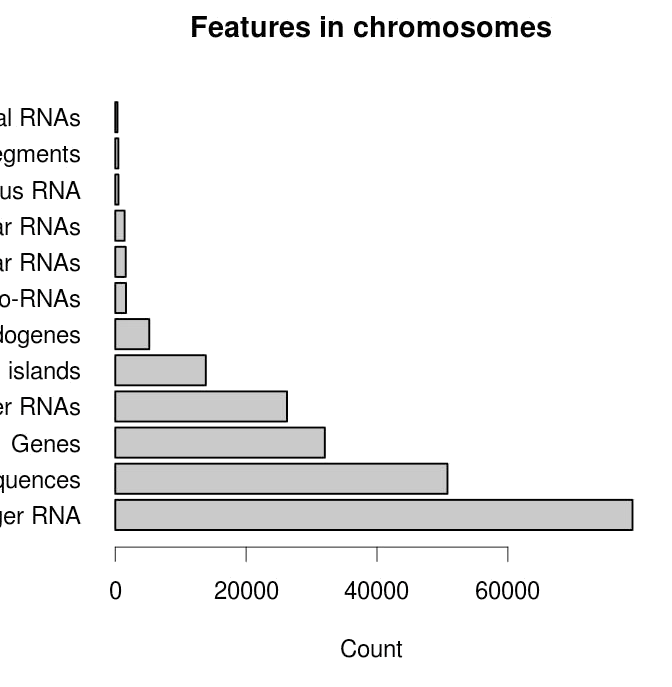


The text labels should all be horizontal (las=1) Note that you can pass this parameter either via par, or as an additional option to barplot.

> par(las=1)

> barplot(data1,horiz=TRUE,xlab =

"Count",names.arg=data$Feature,main="Features in chromosomes")



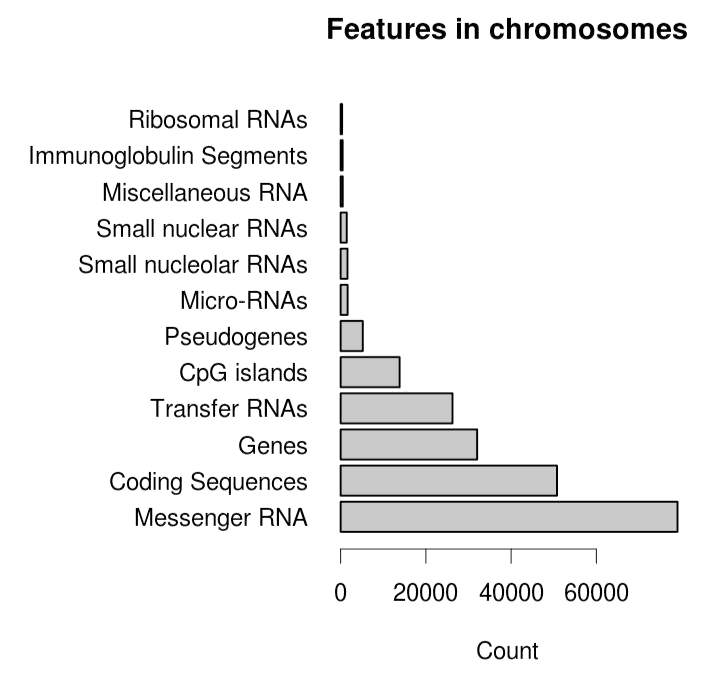
The margins should be adjusted to accommodate the labels (par mar parameter). You need to supply a 4 element vector for the bottom,left,top and right margin values. Look at the value of par()$mar to see what the default values are so you know where to start.

Note that you will have to redraw the barplot after making the changes with par.

> par(mar=c(5.1,13.1,4.1,2.1))

> barplot(data1,horiz=TRUE,xlab =

"Count",names.arg=data$Feature,main="Features in chromosomes")



3.

The file male\_female\_counts.txt contains a time series split into male and female count values.

• Plot this as a barplot

> data=read.delim("male\_female\_counts.txt")

> data

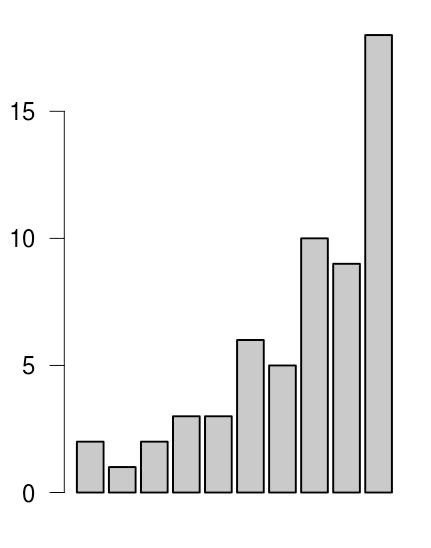
Sample Count

|  |  |
| --- | --- |
| 1 D1 Male | 2 |
| 2 D1 Female | 1 |
| 3 D2 Male | 2 |
| 4 D2 Female | 3 |

5 D3 Male 3 6 D3 Female 6 7 D4 Male 5 8 D4 Female 10 9 D5 Male 9

10 D5 Female 18

> barplot(data$Count)

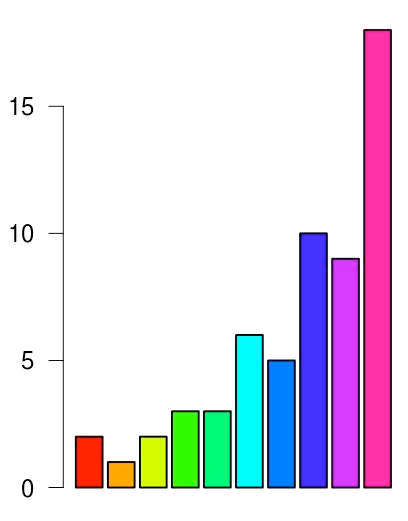


Make all bars different colours using the rainbow function

• rainbow takes a single argument, which is the number of colours to generate, eg

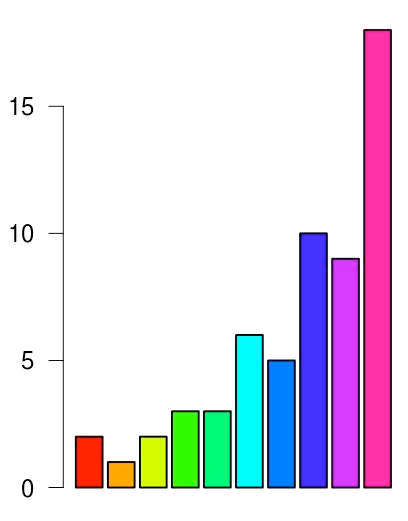
rainbow(10) Try making the vector of colours separately before passing it as the col argument to barplot (col=rainbow(10)).

> barplot(data$Count,col=rainbow(10))



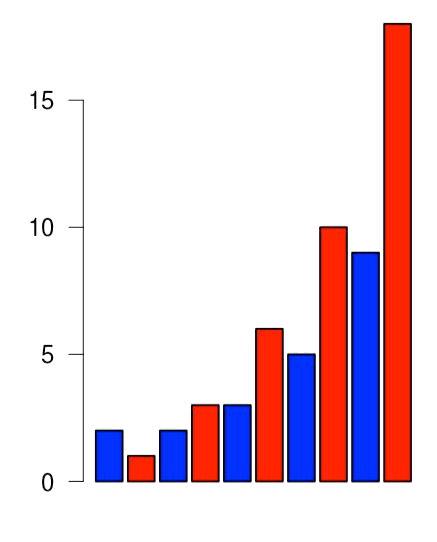
Rather than hard coding the number of colours, think how you could use nrow to automatically generate the correct number of colours for the size of dataset.

> barplot(data$Count,col=rainbow(nrow(data)))



Replot, and make the bars for the males a different colour to those for the females. In this case the male and female samples alternate so you can just pass a 2 colour vector to the col parameter to achieve this effect. (col=c(“blue2”,”red2”)) for example.

> barplot(data$Count,col=c("blue","red"))



EXP-3

Name: Dhruv Goel

Reg.No: 16BCE0686

Topic: ggplot

1.Importing library ggplot2 and getting the input files

> library(ggplot2)

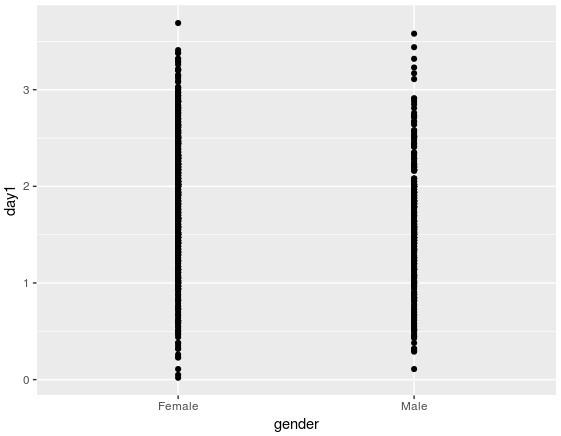
> setwd("/tmp/guest-LDdlZN/Desktop")

> data1=read.delim("Changing.csv")

> data2=read.delim("DownloadFestival.dat")

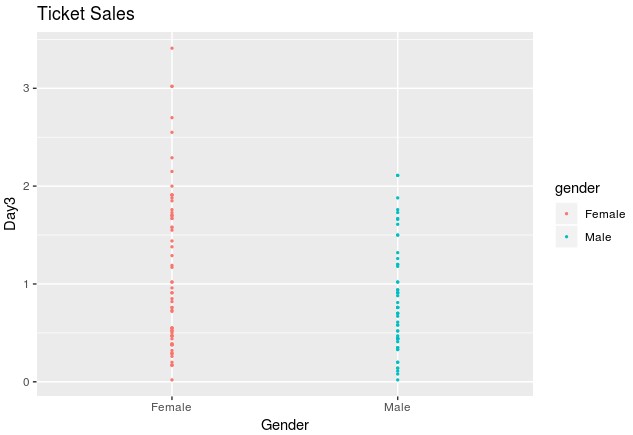
2.Scatter plot

> ggplot(data2, aes(x=gender,y=day1))+geom\_point()

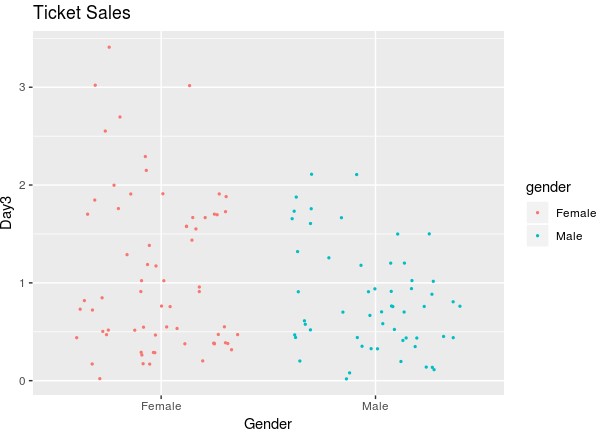


>

ggplot(data2, aes(x=gender,y=day3,col=gender))+geom\_point(size=.5)+xlab("Gender")+ylab("Day3") +ggtitle("Ticket Sales")

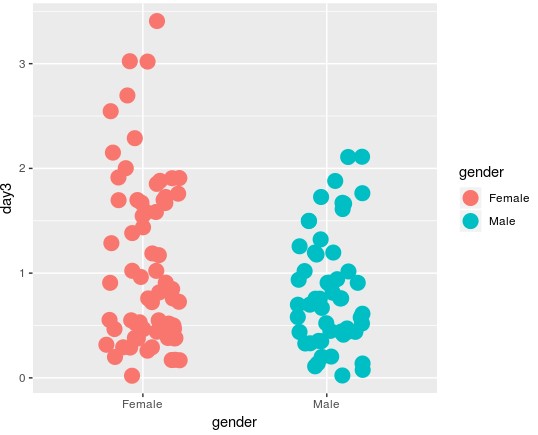
 > ggplot(data2,

aes(x=gender,y=day3,col=gender))+geom\_point(size=.5,position="jitter")+xlab("Gender")+ylab("Day3") +ggtitle("Ticket Sales")



3.StripChart

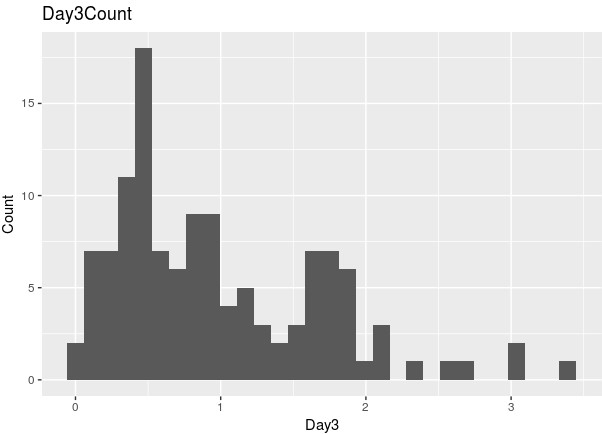
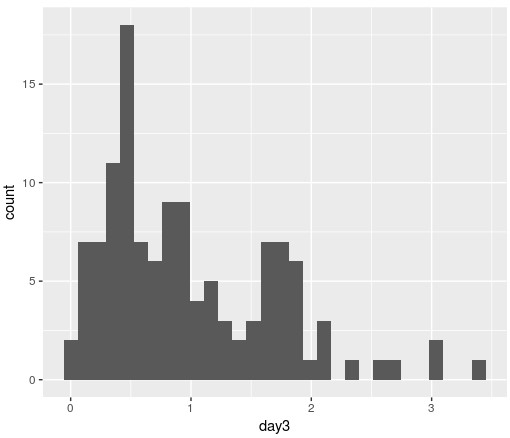
> ggplot(data2, aes(gender,day3,col=gender))+geom\_jitter(position = position\_jitter(width=0.2),size=5)



4.Histogram

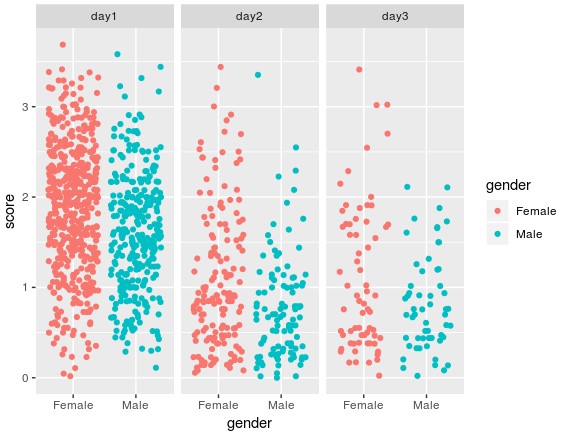
> ggplot(data2, aes(day3))+geom\_histogram()

>



ggplot(data2, aes(day3))+geom\_histogram()+labs(x="Day3",y="Count")+ggtitle("Day3Count")

> ggplot(festival.data.stack, aes(gender, score,colour=gender))+geom\_point(position="jitter") +facet\_grid(~day)



EXP-4

Name: Dhruv Goel

Reg.No: 16BCE0686

Topic: ggplot

I Diamonds

1) From the diamonds data (included in ggplot2 package), find the relationship between caret, price and color. It is known that the price varies on the carat. But what about color? Generate plots that you feel best shows the relationships and then write in words about what you see.

Minimum Expected graphs:

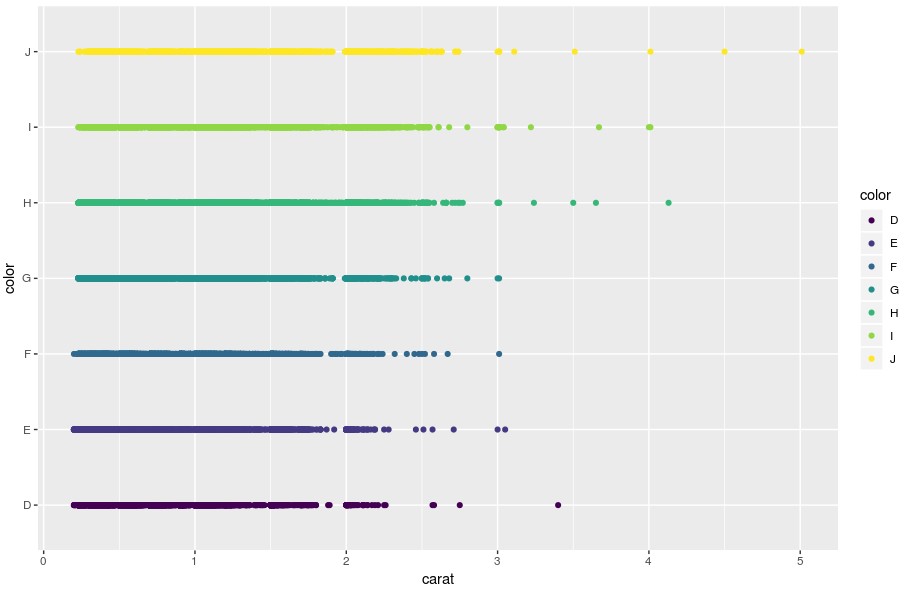
1. Scatterplot (each color in different color)
2. Fit a regression line
3. Box plot
4. Jitter
5. Density

Scatterplot of color vs carat:

> library(ggplot2)

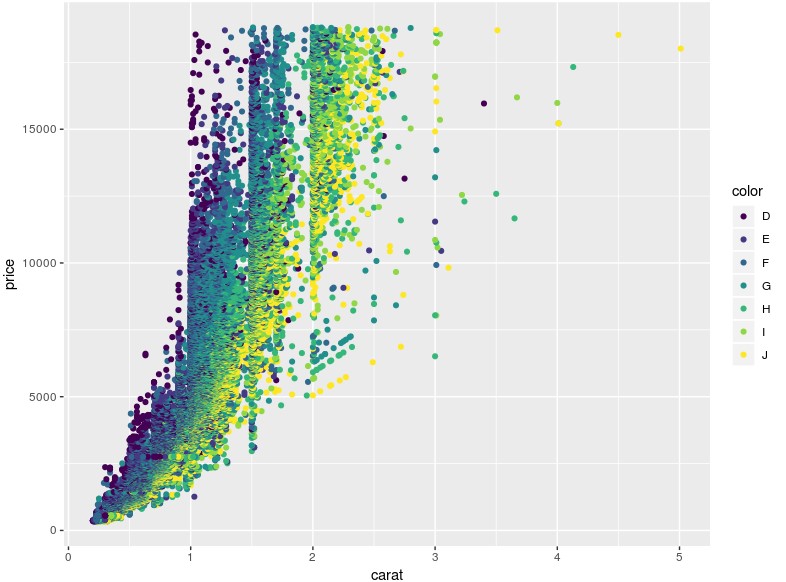
> data1=diamonds

> ggplot(data1,aes(x=carat,y=color,col=color))+geom\_point()



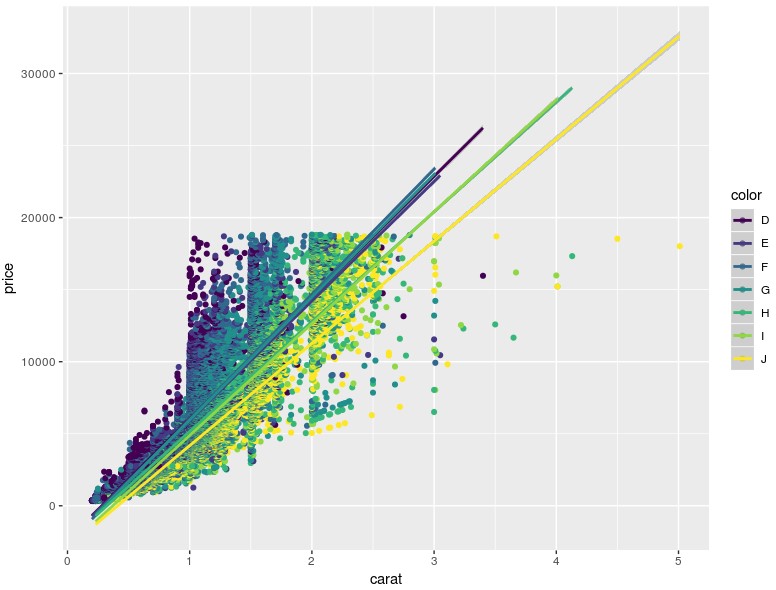
Price vs Carat

> ggplot(data1,aes(x=carat,y=price,col=color))+geom\_point()



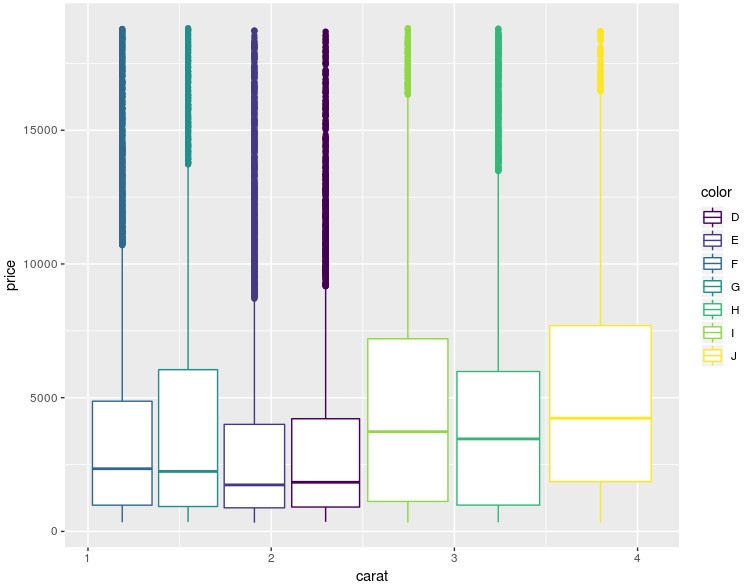
With regression line for each color:

> ggplot(data1,aes(x=carat,y=price,col=color))+geom\_point()+geom\_smooth(method='lm')

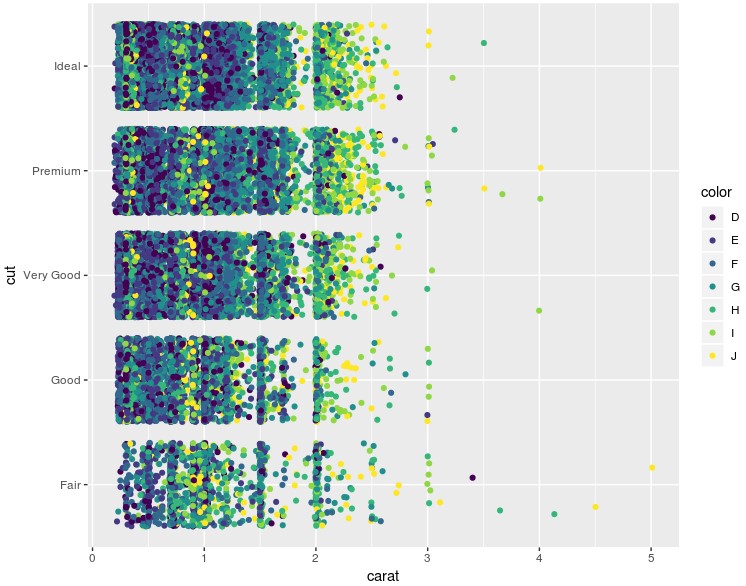


Boxplot price vs carat:

> ggplot(data1,aes(x=carat,y=price,col=color))+geom\_boxplot()



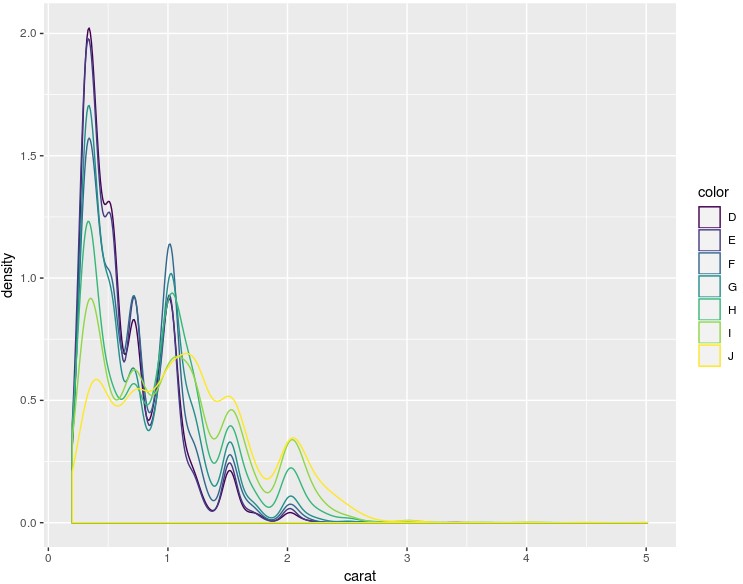
Jitter Plot: Carat vs Cut:



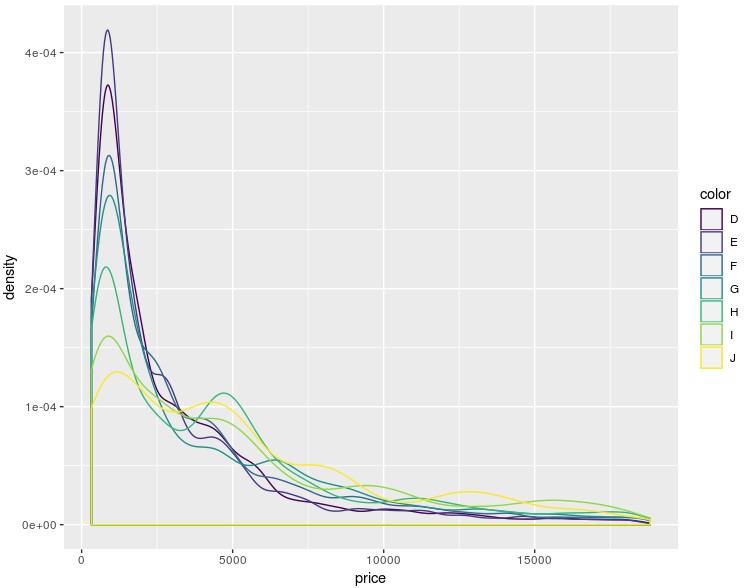
Density Graphs:

1.Carat:

> ggplot(data1,aes(x=carat,col=color))+geom\_density()



2.Price

II Msleep

2) Using msleep

dataset (included in

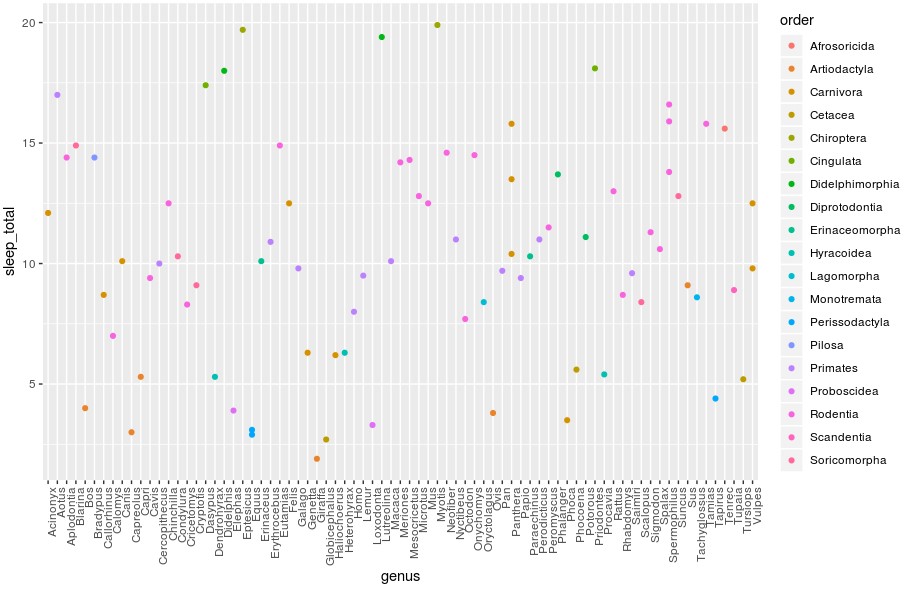
ggplot2) make plots to compare the sleep times of different order and genus. Pick any one order and find if

The sleep time varies with body weight. If so what is the relationship?

The sleep time varies according to the vore. Yes/No The sleep time varies with brain weight. Prove or disprove

Comparision of sleeping times of different orders and genus:

> ggplot(data2,aes(x=genus,y=sleep\_total,col=order))+geom\_point()+theme(axis.text.x = element\_text(angle = 90,hjust = 1))

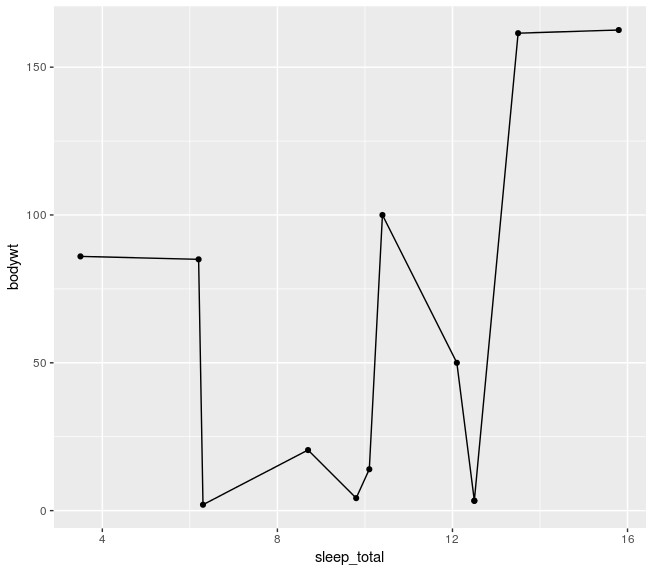


Selected order: Carnivora

> data3=msleep[which(msleep$order == 'Carnivora'),]

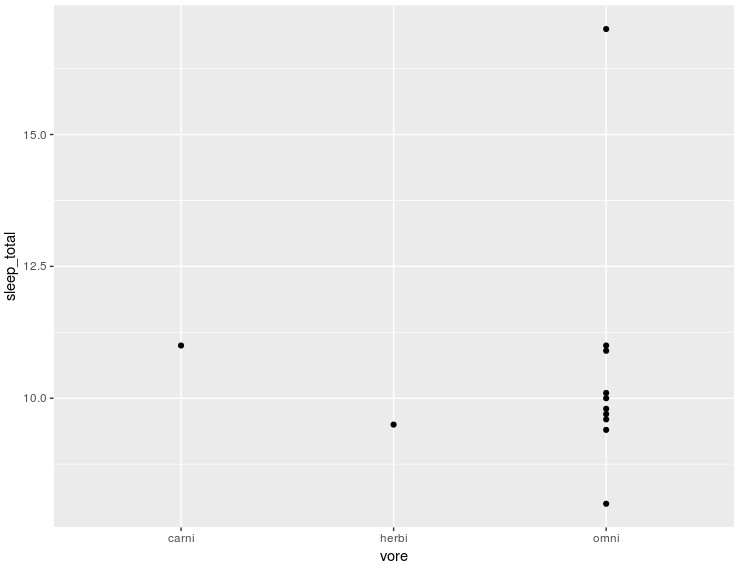
Body Weight vs Sleeping time:

> ggplot(data3,aes(x=sleep\_total,y=bodywt))+geom\_line()+geom\_point()



Sleep time vs vore:(For Primates)

> ggplot(data4,aes(x=vore,y=sleep\_total))+geom\_point()

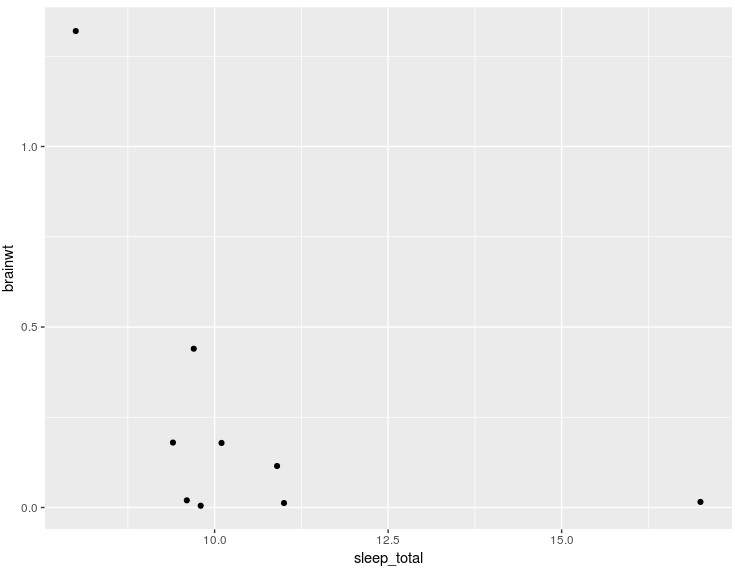


Carnivores tend to sleep longer.

Sleep time vs Brain weight:

> ggplot(data4,aes(x=sleep\_total,y=brainwt))+geom\_point()

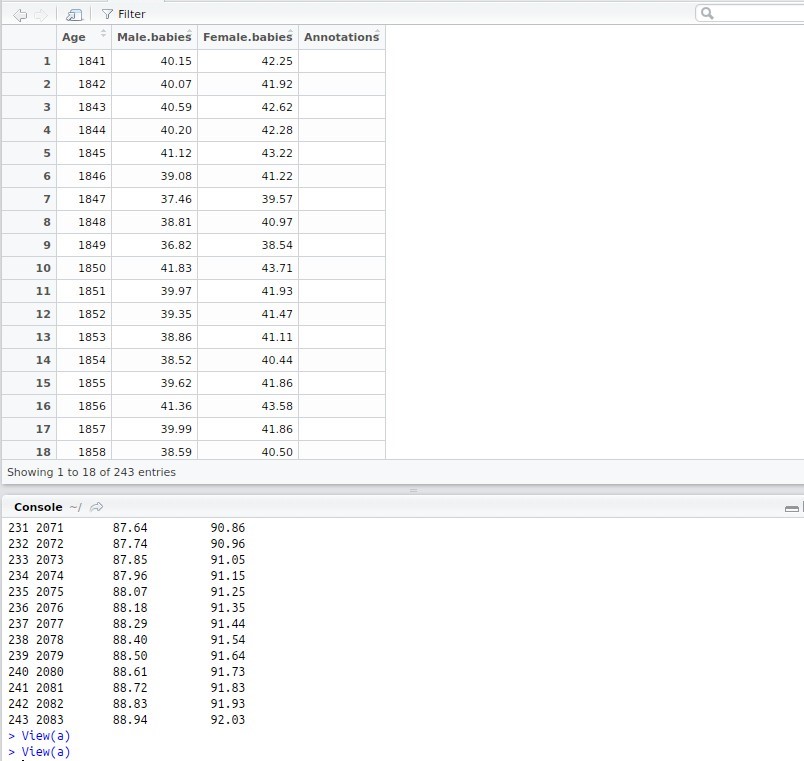
Higher the brain weight lesser sleep.



SMALL TEST – DATA VISUALISATION

DHRUV GOEL 16BCE0686

1. How many rows and columns are there?a=read.csv("Downloads//LifeExpectancy.csv") It has 243 rows and 4 columns.

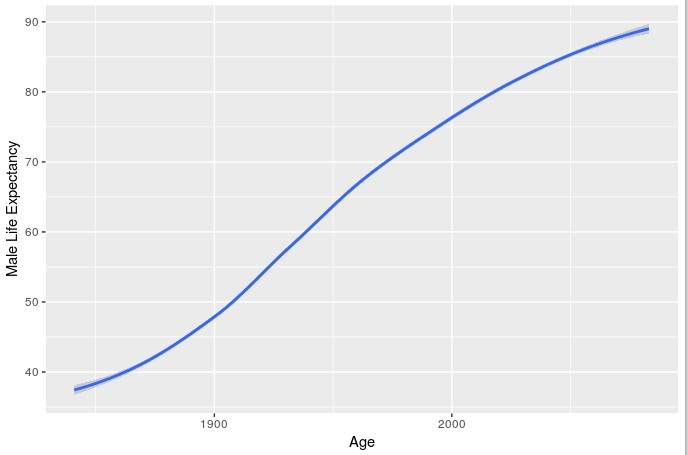


1. Plot how the Male life expectancy rate changes over the years

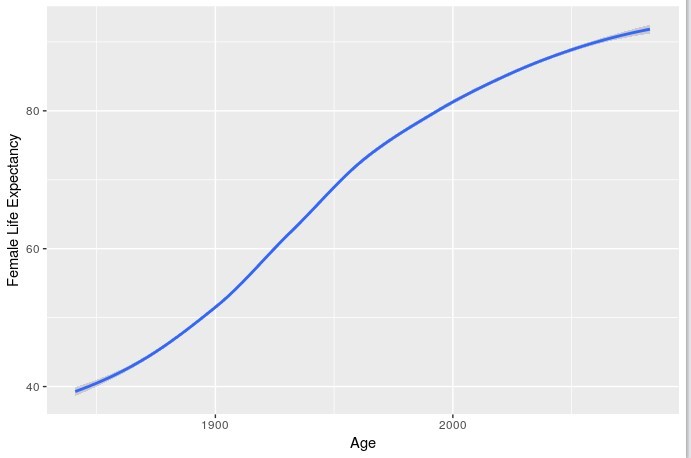
library(ggplot2)

> a=read.csv("Downloads//LifeExpectancy.csv")

> ggplot(a, aes(x=Age, y=Male.babies))+geom\_smooth()



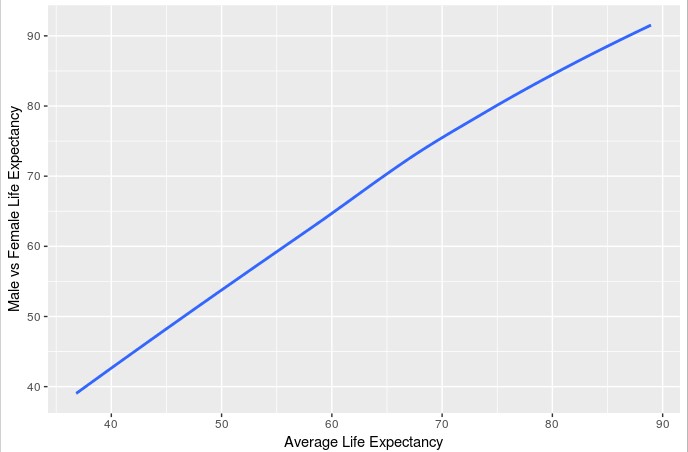
1. Plot how the Female life expectancy rate changes over the yearsggplot(a, aes(x=Age, y=Female.babies))+geom\_smooth()



1. Plot the relationship between Male and Female Life Expectancy

ggplot(a, aes(x=a$Male.babies, y=a$Female.babies))+geom\_smooth()+xlab("Average Life

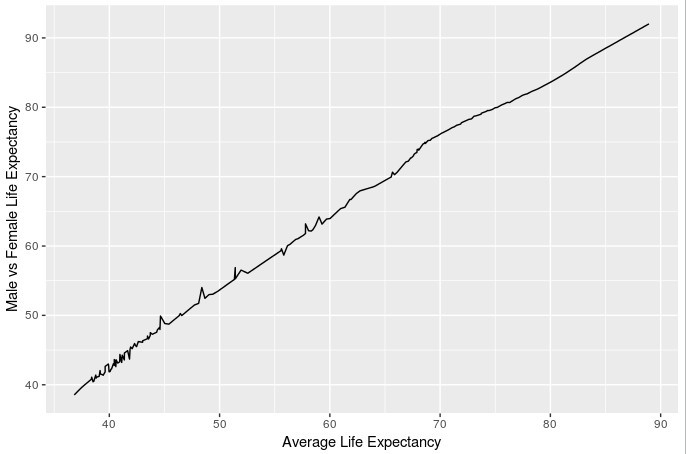
Expectancy")+ylab("Male vs Female Life Expectancy")

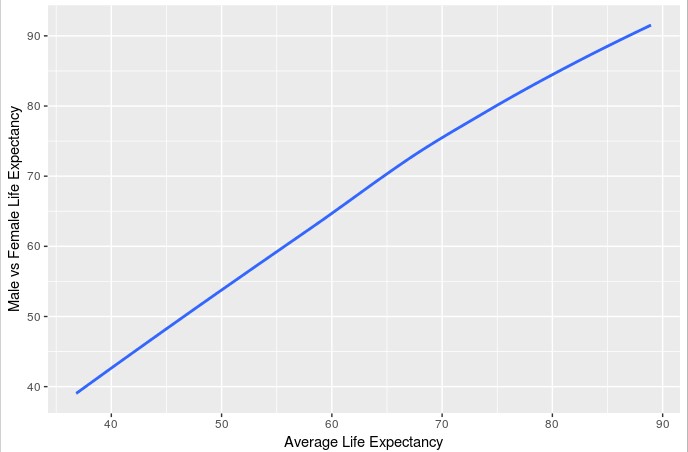


a. What are the different ways to do the above? Show the plots.

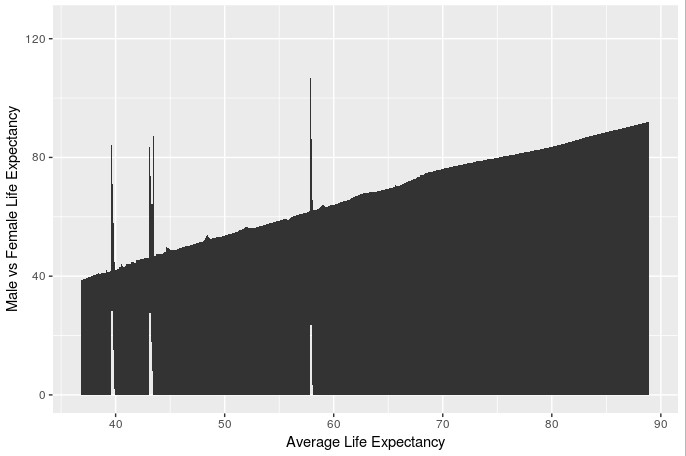
7. Draw any other graphs that you can for the given dataset. (try density, jitter, box , bar(vertical, horizontal, stacked, beside) and other plots)

1. geom\_smooth

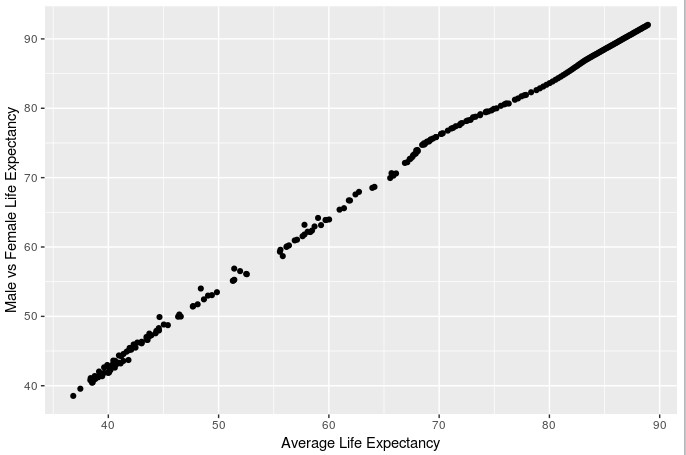
ggplot(a, aes(x=a$Male.babies, y=a$Female.babies))+geom\_smooth()+xlab("Average Life Expectancy")+ylab("Male vs Female Life Expectancy")



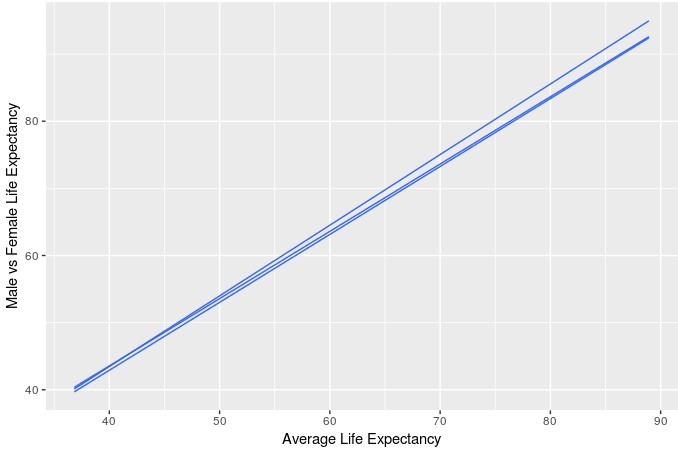
1. geom\_line
2. geom\_area



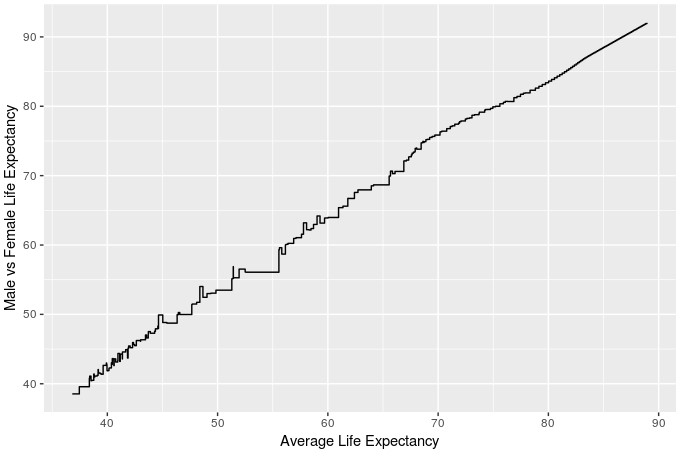
1. geom\_jitter



1. geom\_quantile



1. geom\_step

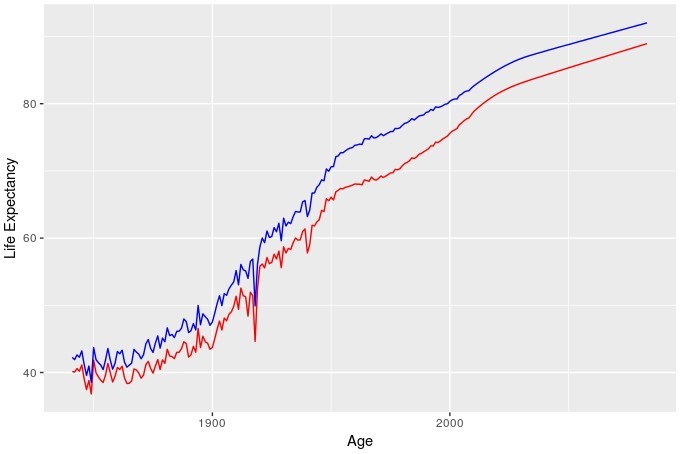


1. What was the effect of world war on life expectancy of both male and female?
2. Plot female and male life expectancy during WWII in the same graph

The WW1 took place from 1914 to 1918 and dring that time there was a sharp decline in the life expectancy for both male and female. Again, the second world war (WW2) took place from 1939 to 1945, and there was a decline in life expectancy.

ggplot(a)+geom\_line(color='red', aes(x=Age, y=Male.babies))+geom\_line(color='blue',

aes(x=Age, y=Female.babies))+ylab("Life Expectancy")



Name: Priyal Agrawal

Reg:16BCE0776

LAB 1

CGPA: histo

> cgpa<-c(-5.0,-3.0,2.0,2.5,3.0,3.2,6.1,6.2,7.6,7.8,9,9.1,9.2,9.8,20,20.1,20.2)

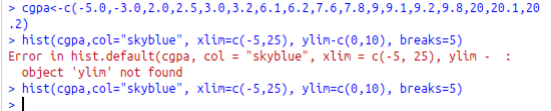
> hist(cgpa,col="skyblue", xlim=c(-5,25), ylim-c(0,10), breaks=5)

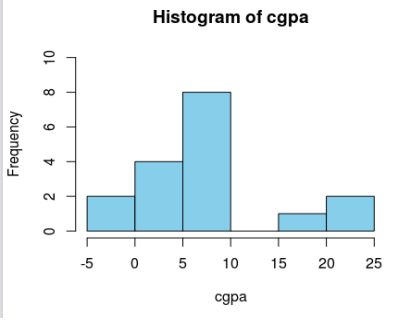
Error in hist.default(cgpa, col = "skyblue", xlim = c(-5, 25), ylim - :

object 'ylim' not found

> hist(cgpa,col="skyblue", xlim=c(-5,25), ylim=c(0,10), breaks=5)

>





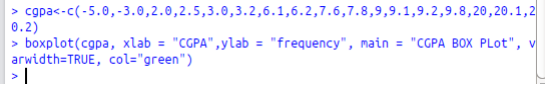
CGPA: box

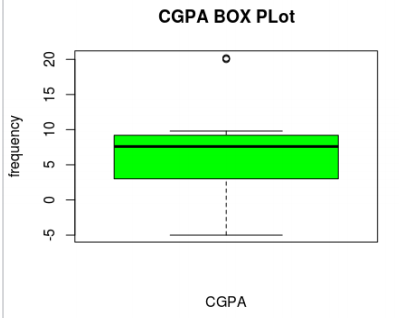
> cgpa<-c(-5.0,-3.0,2.0,2.5,3.0,3.2,6.1,6.2,7.6,7.8,9,9.1,9.2,9.8,20,20.1,20.2)

> boxplot(cgpa, xlab = "CGPA",ylab = "frequency", main = "CGPA BOX PLot", varwidth=TRUE,

col="green")

>



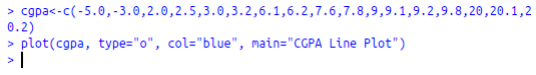


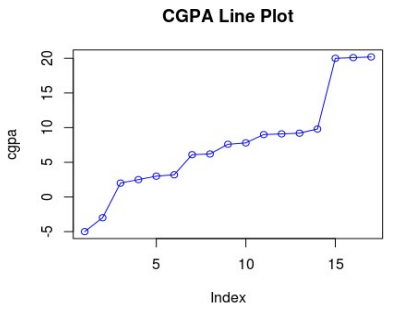
CGPA: line

> cgpa<-c(-5.0,-3.0,2.0,2.5,3.0,3.2,6.1,6.2,7.6,7.8,9,9.1,9.2,9.8,20,20.1,20.2)

> plot(cgpa, type="o", col="blue", main="CGPA Line Plot")

>



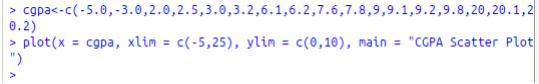


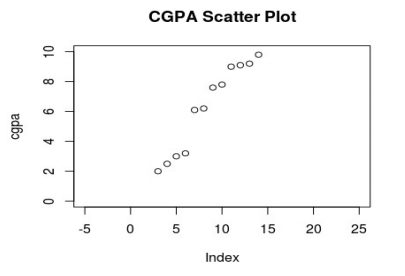
CGPA: Scatter

> cgpa<-c(-5.0,-3.0,2.0,2.5,3.0,3.2,6.1,6.2,7.6,7.8,9,9.1,9.2,9.8,20,20.1,20.2)

> plot(x = cgpa, xlim = c(-5,25), ylim = c(0,10), main = "CGPA Scatter Plot")

>





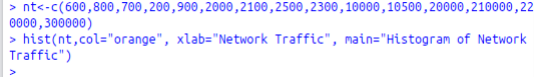
Network Traffic: histo

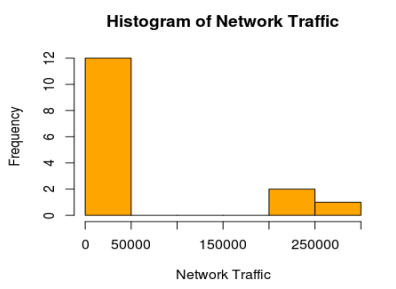
> nt<-

c(600,800,700,200,900,2000,2100,2500,2300,10000,10500,20000,210000,22000

0,300000)

> hist(nt,col="orange", xlab="Network Traffic", main="Histogram of Network Traffic")





Network Traffic: box

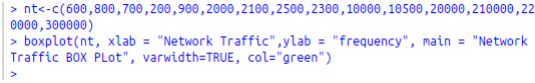
> nt<-

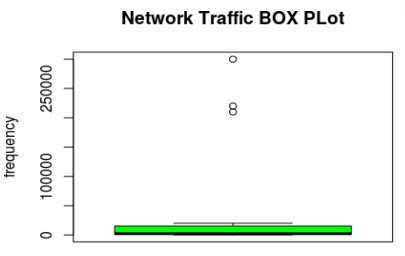
c(600,800,700,200,900,2000,2100,2500,2300,10000,10500,20000,210000,22000

0,300000)

> boxplot(nt, xlab = "Network Traffic",ylab = "frequency", main = "Network Traffic BOX

PLot", varwidth=TRUE, col="green")





Network Traffic: line

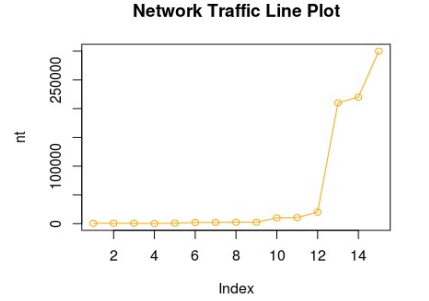
> nt<-

c(600,800,700,200,900,2000,2100,2500,2300,10000,10500,20000,210000,22000

0,300000)

> plot(nt, type="o", col="orange", main="Network Traffic Line Plot")





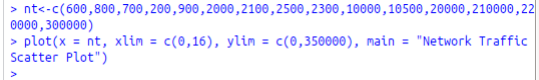
Network Traffic: Scatter

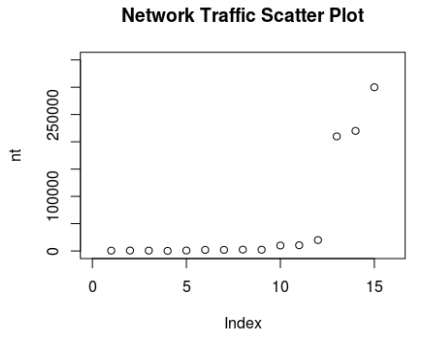
> nt<-

c(600,800,700,200,900,2000,2100,2500,2300,10000,10500,20000,210000,22000

0,300000)

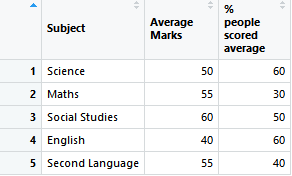
> plot(x = nt, xlim = c(0,16), ylim = c(0,350000), main = "Network Traffic Scatter Plot")





Movie Rating: Histo

Use appropriate method to show the data visually



>library(readxl)

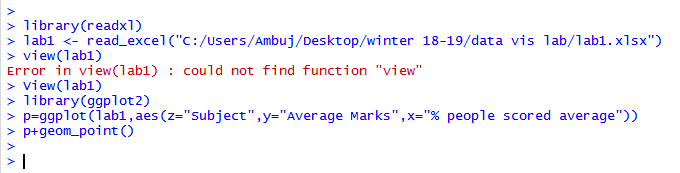
> lab1 <- read\_excel("C:/Users/Ambuj/Desktop/winter 18-19/data vis lab/lab1.xlsx")

> view(lab1)

> library(ggplot2)

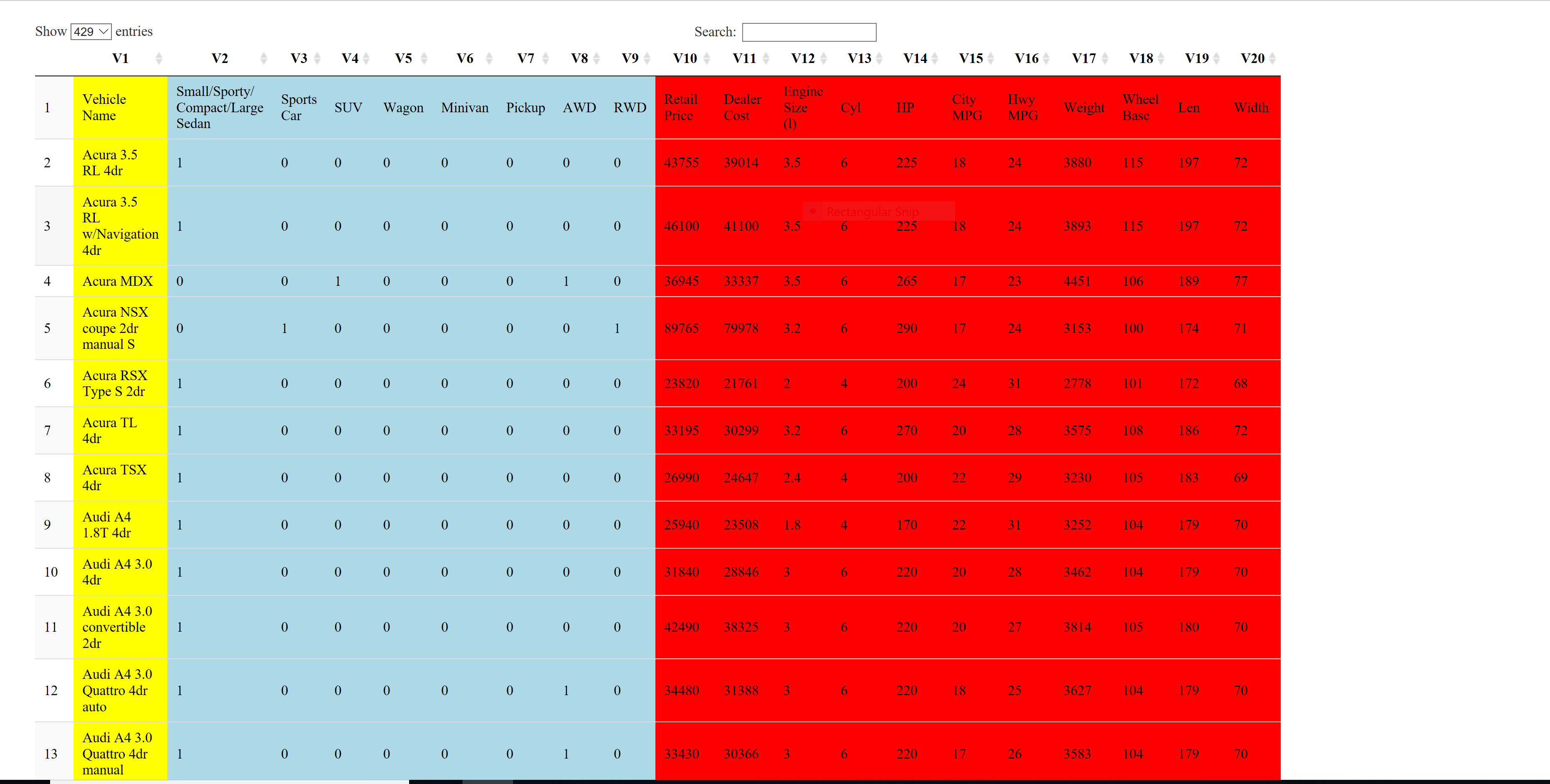
> p=ggplot(lab1,aes(z="Subject",y="Average Marks",x="% people scored average"))

> p+geom\_point()





Q3 a)



|  |  |
| --- | --- |
|  | Quantitative |
|  | Categorical |
|  | Ordinal |

Commands Used

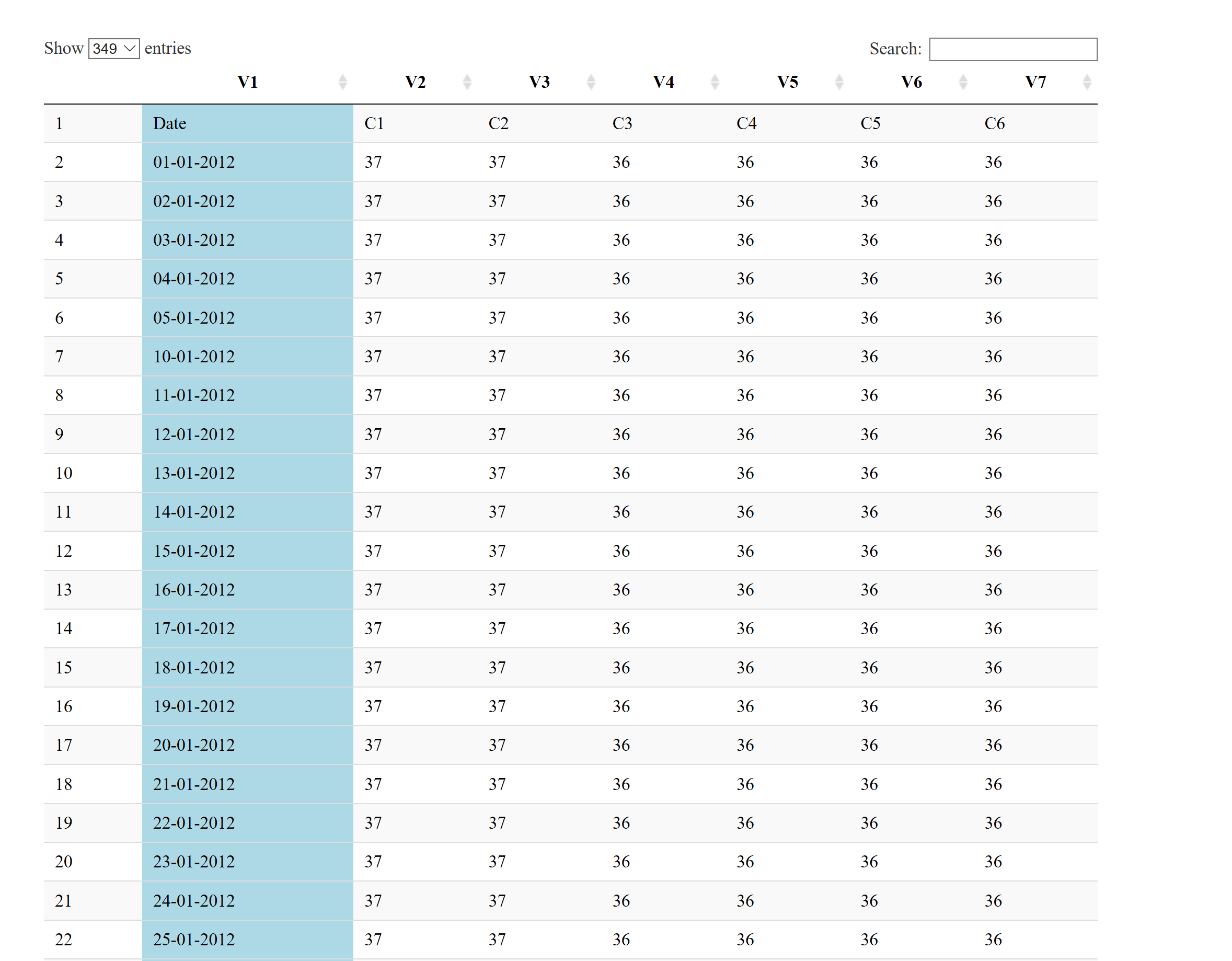
> library(DT)

> data = read.table("C:/Users/Aditya Dhall/Desktop/Cars.csv", header = FALSE, sep=',')

> options(DT.options = list(pageLength = 429))

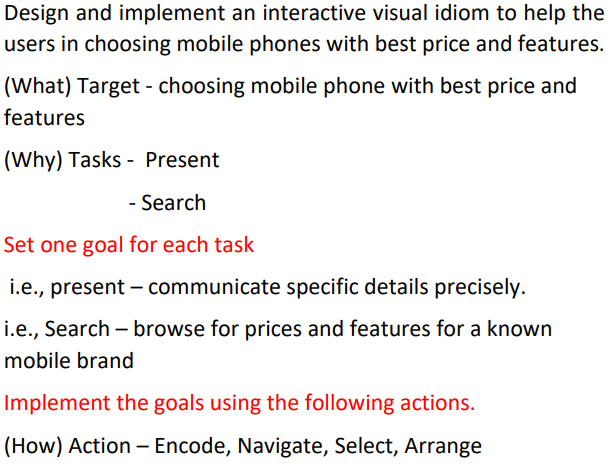
|  |
| --- |
| > catagorical = c("V2","V3","V4","V5","V6","V7","V8","V9")  > quantitative = c("V10","V11","V12","V13","V14","V15","V16","V17","V18", "V19", "V20")  > ordinal = "V1"  > datatable(data)  %>% formatStyle(catagorical,backgroundColor = "lightblue")  %>% formatStyle(ordinal,backgroundColor = "yellow")  %>% formatStyle(quantitative,backgroundColor = "red") |

Q3b)



|  |  |
| --- | --- |
|  | Key Attribute |

|  |
| --- |
|  |
| |  | | --- | |  | |



> data<-read.csv('/home/anish/Desktop/SEM6/data visualization/Lab/data.csv')

> head(data)

Mobile cost touch back.camera front.camera NFC Memory

1 Iphone X 94000 Yes 18 5 Yes 32

2 Mi A1 15000 Yes 15 12 No 64

3 Samsung S9 70000 Yes 18 18 Yes 64

4 Iphone 10 75000 Yes 18 18 Yes 32

5 Redmi Y2 9000 Yes 16 12 No 16

6 Poco F1 20000 Yes 20 20 No 64

> is.factor(data$NFC)

[1] TRUE

> data$NFC<-c("1","0","1","1","0","0")

> is.factor(data$touch)

[1] TRUE

> levels(data$touch)<-c("0","1")

> head(data)

Mobile cost touch back.camera front.camera NFC Memory

1 Iphone X 94000 0 18 5 1 32

2 Mi A1 15000 0 15 12 0 64

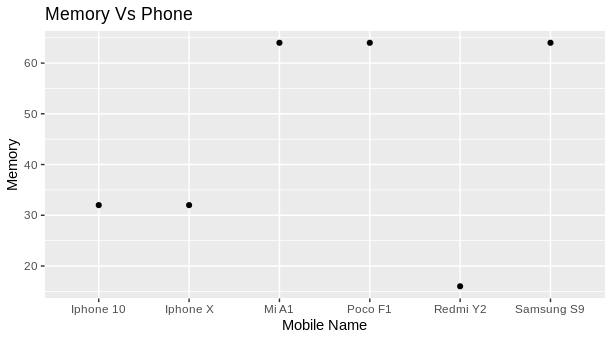
3 Samsung S9 70000 0 18 18 1 64

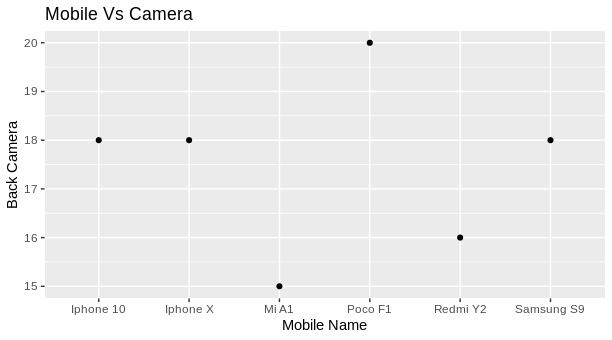
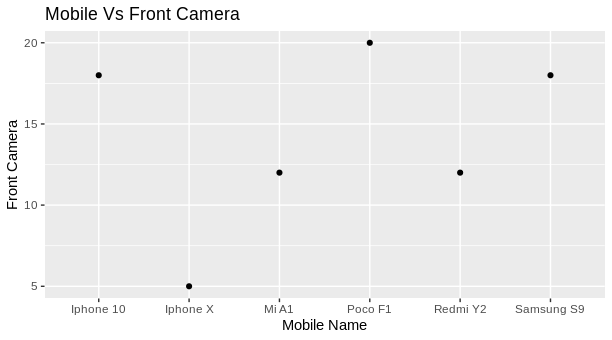
4 Iphone 10 75000 0 18 18 1 32

5 Redmi Y2 9000 0 16 12 0 16

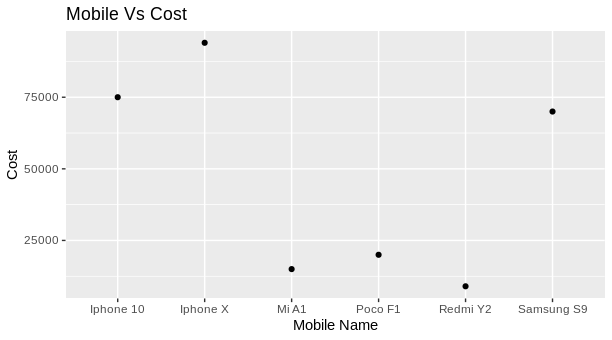
6 Poco F1 20000 0 20 20 0 64

> qplot(data$Mobile,data$Memory,xlab ="Mobile Name",ylab ="Memory",main ="Memory Vs Phone")

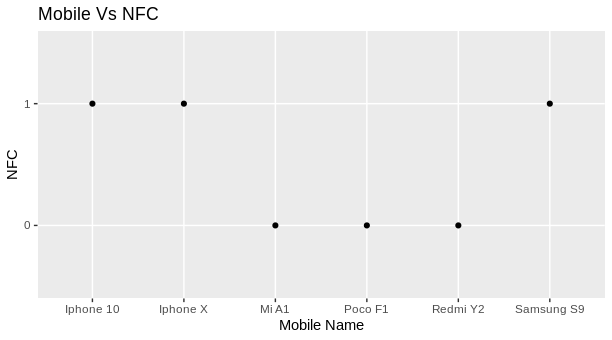


> qplot(data$Mobile,data$back.camera,xlab = "Mobile Name",ylab = "Back Camera",main = "Mobile Vs Camera")> qplot(data$Mobile,data$front.camera,xlab = "Mobile Name",ylab = "Front Camera",main = "Mobile Vs Front Camera")

> qplot(data$Mobile,data$cost,xlab = "Mobile Name",ylab = "Cost",main = "Mobile Vs Cost")



> qplot(data$Mobile,data$NFC,xlab = "Mobile Name",ylab = "NFC",main = "Mobile Vs NFC")



**Questtion 2:- Generate a Node Edge Diagram (directed) for the data given in the table below. The**

**attributes : Relationship, Age and Lives in should be displayed when there is a mouse click on each node. Use different shapes to differentiate gender. Also use similar node colors for the same places in which persons are living.**

data = read.csv("data.csv",header = TRUE)

Adata = data.frame(Node = 'A', Age = 100, Lives.in = 'Coimbatore')

vertices = data.frame(Node = data$Destination, Age = data$Age, Lives.in = data$Lives.in)

vertices = rbind(Adata,vertices)

library(visNetwork)

nodes = data.frame(id = vertices$Node, label= vertices$Node)

nodes$title = vertices$Age

nodes$color.background = c("slategrey", "tomato", "gold","orange","red","green")[as.factor(vertices$Lives.in)]

nodes$title = paste0("<p>Age:", vertices$Age,"<br>Lives in :", vertices$Lives.in ,"</p>")

links = data.frame(from = data$Source, to = data$Destination, label = data$Relationship)

visNetwork(nodes, links) %>% visEdges(arrows = 'to')

