CARS AND THEIR SPECS DATASET

Group Members :

- Tushar (19/19064)
- Ashish Bachkheti (19/19066)

Description of dataset:

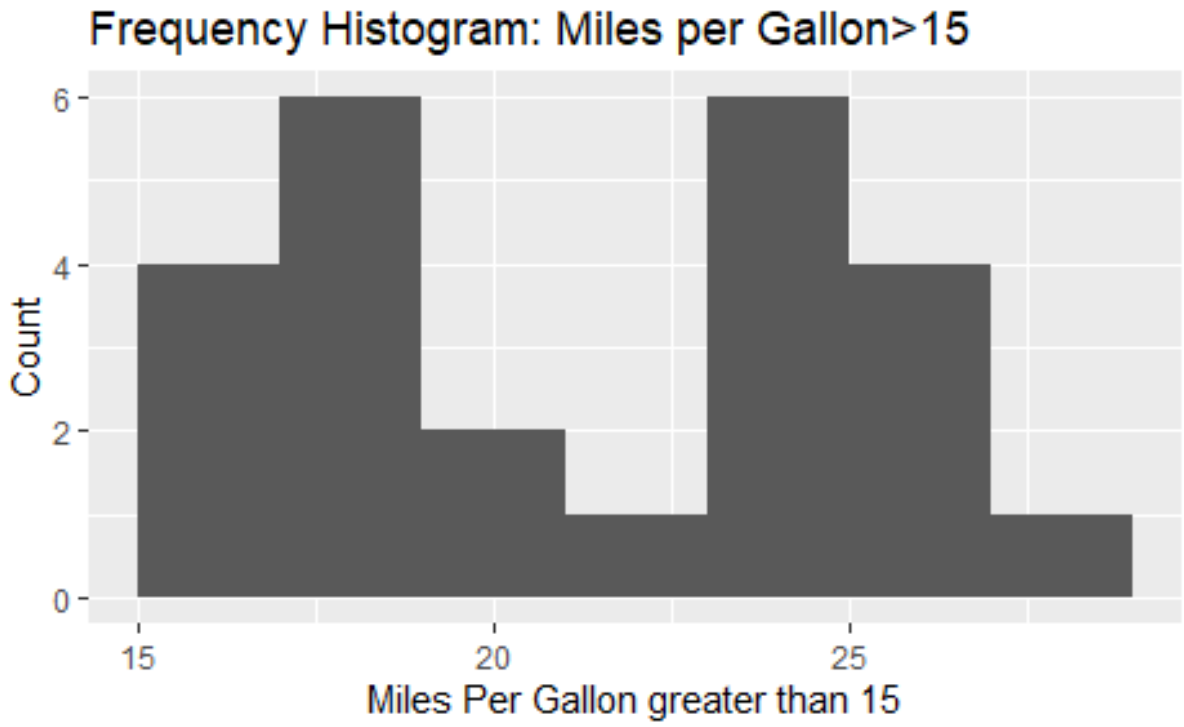
Cars : Various types of cars
MPG : Miles per gallon
Cylinders : Number of Cylinders
Displacement : Measure of the cylinder volume swept by all of the pistons
Horsepower : The power an engine produces
Weight : weight of car in pounds
Acceleration : Time taken from 0mph to 60mph
Model : Model of car
Origin : Origin of car

Information extracted:

- MPG greater than 15
- Percentage of Numbers of cylinders used in cars
- Cars with HP greater than 175 and maximum HP
- Top 5 cars with on basis of HP
- Percentage of cars produced in different origins.
- List of cars produced in different Regions
- Different weights of car
- Weight Comparison by Region of Origin
- Each Region of Origin's Product Mix Over Time

1.Frequency Histogram of Miles per Gallon greater than 15

```
mpg_15=cars$MPG[cars$MPG>15]
print(qplot(mpg_15, xlab = 'Miles Per Gallon greater than 15', ylab = 'Count',
binwidth = 2,
      main='Frequency Histogram: Miles per Gallon>15'))
```



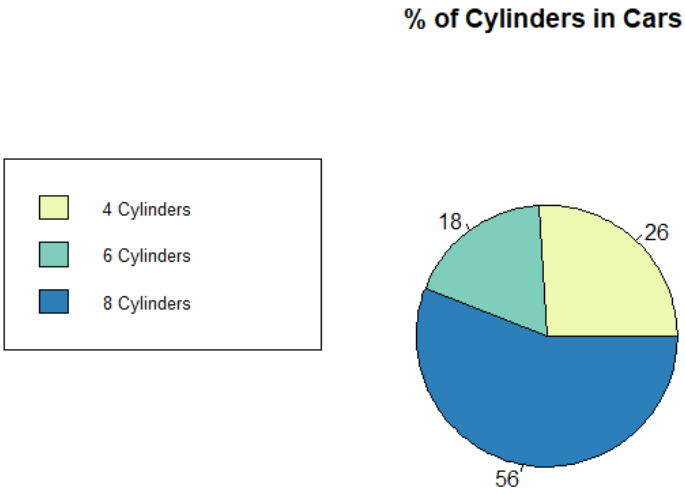
CONCLUSION

This Helps us to maintain the count of cars having MPG greater than 15

2. Percentage of numbers of cylinders used in different Cars

```
a=table(cars$Cylinders)
list1<-unname(a)
labels<-dimnames(a)
labels<-unlist(labels)
```

```
l<-c("4 Cylinders","6 Cylinders","8 Cylinders")
piepercent<-round(100*list1/sum(list1),2)
pie(list1,main = "% of Cylinders in Cars ",labels = piepercent,
    col = brewer.pal(length(list1),"YlGnBu"))
legend("topleft",l,cex = 0.8,fill =  brewer.pal(length(list1),"YlGnBu"))
```



CONCLUSION

This shows that most cars use 8 cylinders followed by 4 and then 6.

3. Cars With HorsePower greater than 175 .

```
a=cars$Car[cars$Horsepower>175]
print("Cars with Horsepower Greater than 175 are :")
print(a)
print(paste("Car with max HorsePower is :",max(cars$Horsepower)))
```

```
> source('C:/Users/Tushar/OneDrive/Desktop/R Project/R project.R')
[1] "Cars with Horsepower Greater than 175 are :"
 [1] "Ford Galaxie 500"         "Chevrolet Impala"
 [3] "Plymouth Fury iii"        "Pontiac Catalina"
 [5] "AMC Ambassador DPL"       "Buick Estate Wagon (sw)"
 [7] "Ford F250"                "Chevy C20"
 [9] "Dodge D200"               "Hi 1200D"
[11] "Dodge Monaco (sw)"
[1] "Car with max HorsePower is : 225"
> 
```

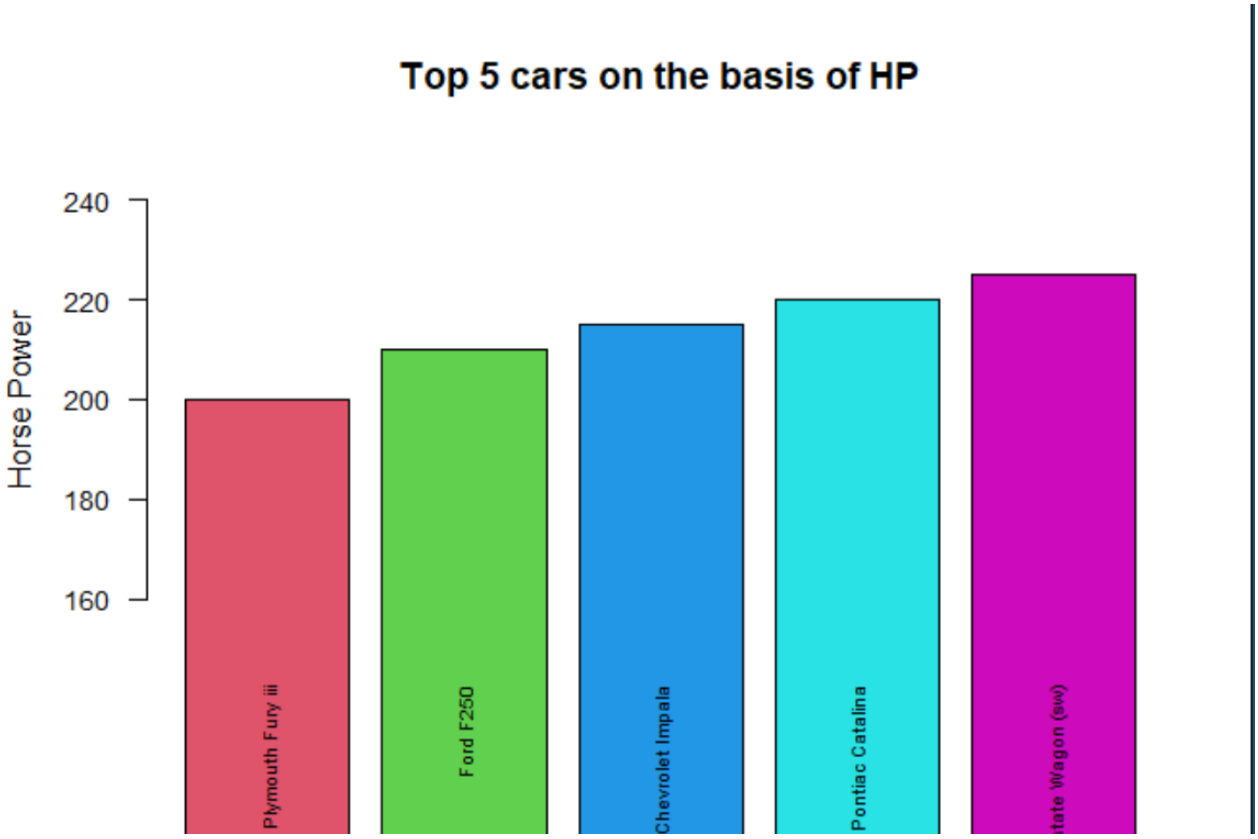
CONCLUSION

We see that there are 11 cars having HP greater than 175 and max HP is 225 in Dodge Monaco.

4.Selecting Top 5 Cars on the basis of HorsePower.

```
cars<-read.csv('cars.csv',header=TRUE,sep=';')
asc=cars$Car[order(cars$Horsepower)]
asc
x <- table(cars$Horsepower)
top5 <- tail(x, 5)
top5 <- as.integer(unlist(dimnames(top5)))
top5
labels<-c("Plymouth Fury iii",
          "Ford F250",
          "Chevrolet Impala",
          "Pontiac Catalina",
          "Buick Estate Wagon (sw)" )

barplot(top5,names.arg=labels,las = 2,ylim=c(150,250),
        col = c(2:6),
        ylab = "Horse Power",
        main = "Top 5 cars on the basis of HP",
        cex.axis = 0.90,cex.names = 0.65
)
```

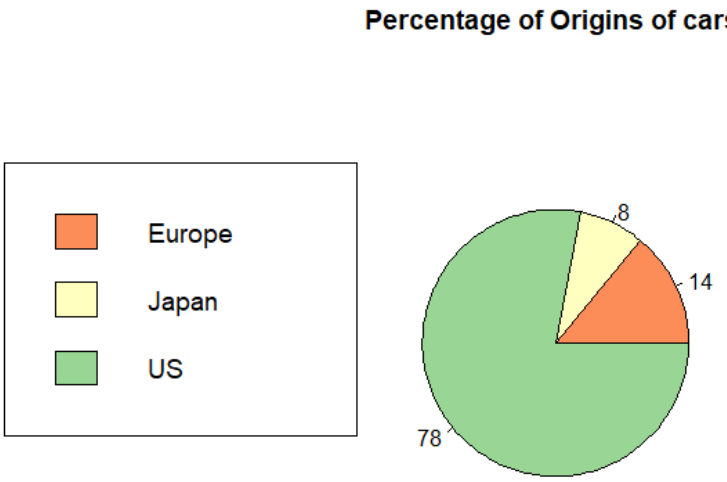


CONCLUSION

This shows us the top 5 cars with highest horsepower among them.

5.Percentage of cars produced in different origins.

```
a=table(cars$Origin)
list1<-unname(a)
labels<-dimnames(a)
labels<-unlist(labels)
piepercent<-round(100*list1/sum(list1),2)
pie(list1,main = "Percentage of Origins of cars ",labels = piepercent,
    col = brewer.pal(length(list1),"Spectral"))
legend("topleft",labels,cex = 1.2,fill =  brewer.pal(length(list1),"Spectral"))
```



CONCLUSION

This data shows that most of the cars are made in the USA and least in Japan.

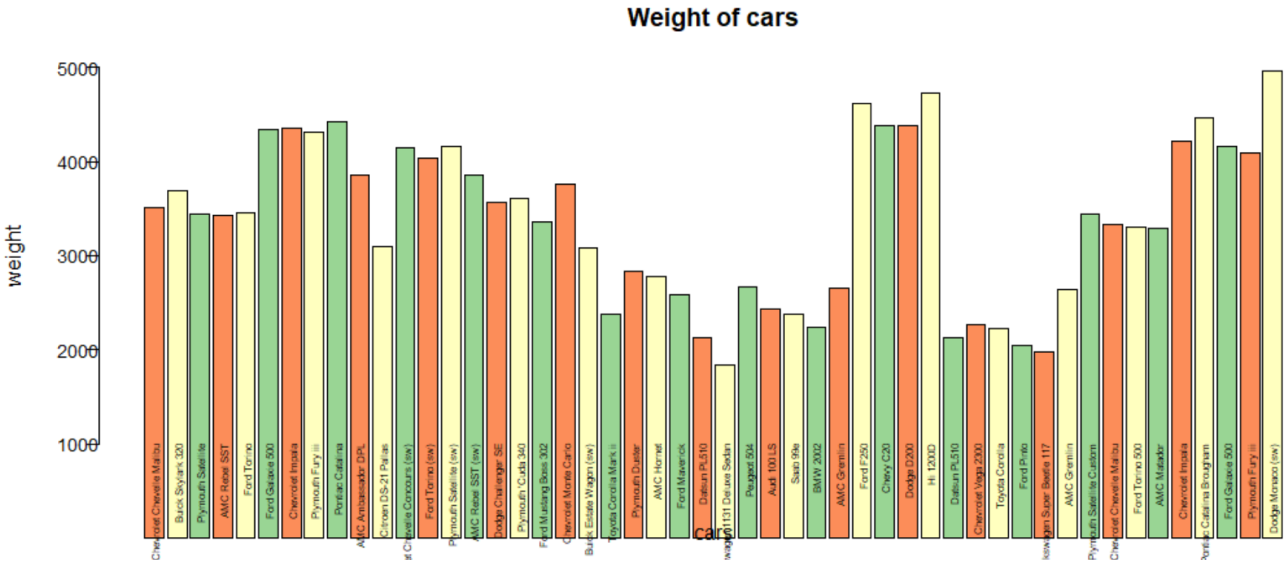
6.List of all Cars Made in Europe and Japan.

```
cars<-read.csv('cars.csv',header=TRUE,sep=';')
y=cars$Car[cars$Origin=="Japan"]
print(paste("The cars with origin Japan ",y))
x=cars$Car[cars$Origin=="Europe"]
print(paste("The cars with origin Europe ",x))
```

```
> source('C:/Users/Tushar/OneDrive/Desktop/R Project/R project.R')
[1] "The cars with origin Japan  Toyota Corolla Mark ii"
[2] "The cars with origin Japan  Datsun PL510"
[3] "The cars with origin Japan  Datsun PL510"
[4] "The cars with origin Japan  Toyota Corolla"
[1] "The cars with origin Europe  Citroen DS-21 Pallas"
[2] "The cars with origin Europe  Volkswagen 1131 Deluxe Sedan"
[3] "The cars with origin Europe  Peugeot 504"
[4] "The cars with origin Europe  Audi 100 LS"
[5] "The cars with origin Europe  Saab 99e"
[6] "The cars with origin Europe  BMW 2002"
[7] "The cars with origin Europe  Volkswagen Super Beetle 117"
>
```

7.Different weights of car.

```
cars<-read.csv('cars.csv',header=TRUE,sep=';')
barplot(names.arg=cars$Car,mgp=c(3,0,0),
        main="Weight of cars",
        xlab="cars",
        ylab="weight",
        ylim=c(1000,5000),
        cars$Weight,
        col=brewer.pal(length(list1),"Spectral"),
        las=2,cex.axis = 0.90,cex.names = 0.50)
```

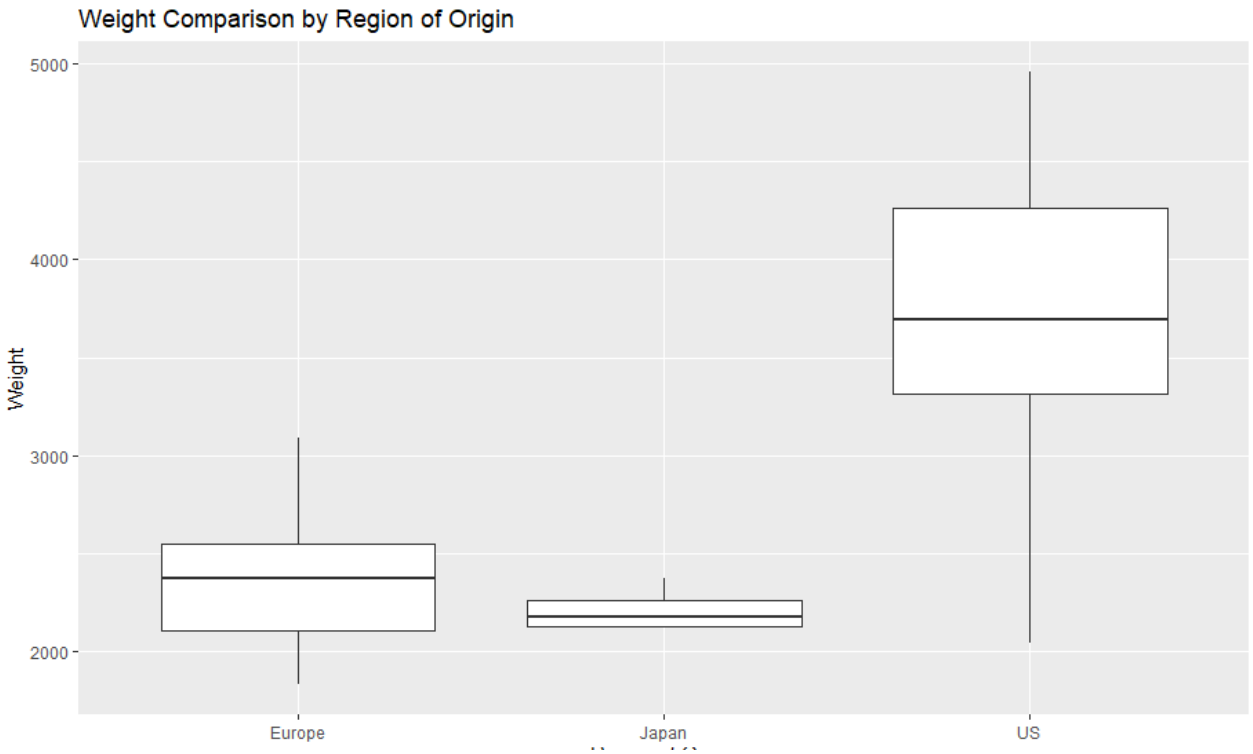


CONCLUSION

This graph shows weights of different cars which might help us to know which car we should buy.

8.Weight Comparison by region of Origin.

```
cars<-read.csv('cars.csv',header=TRUE,sep=';')
print(ggplot(data = cars, aes(x = Origin, y = Weight)) +
      geom_boxplot() +
          xlab('Region of Origin') +
      ylab('Weight') +
      ggtitle('Weight Comparison by Region of Origin') )
```

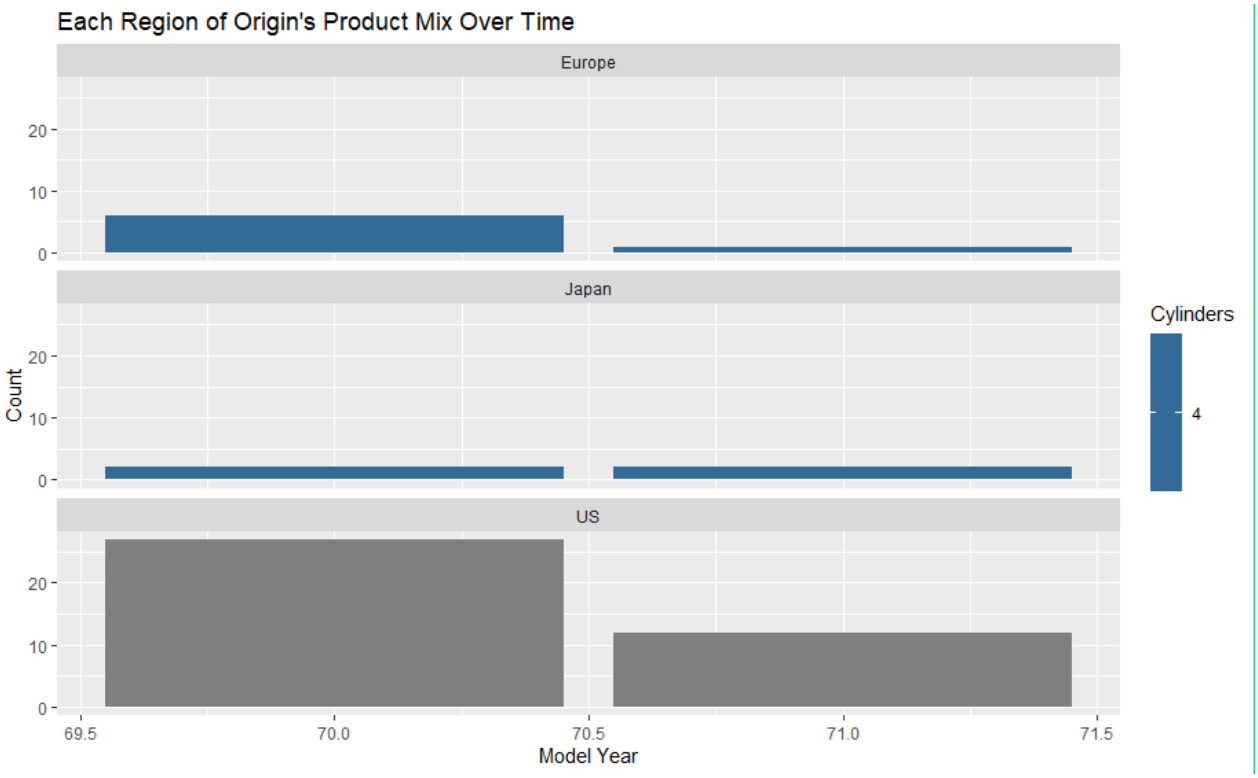


CONCLUSION

It shows that the US produces the most number of cars and also heavy cars.

9.Each Region of Origin's Product Mix Over Time.

```
cars<-read.csv('cars.csv',header=TRUE,sep=';')
print(ggplot(data = cars, aes(x = Model, fill = Cylinders)) +
  geom_bar() +
  facet_wrap(~ Origin, ncol = 1) +
  xlab('Model Year') +
  ylab('Count') +
  ggtitle('Each Region of Origin\'s Product Mix Over Time'))
```



CONCLUSION

This data shows us numbers of cars produced according to their model year in different origins.