title: “Analysis Of Car Data” author: “SaM” date: “4/10/2020” output: html_document —.

Intro to Cars Dataset

#This is an intro to `cars` . The data give the speed of cars and the distances taken to stop. Note that the data were recorded in the 1920s. We are giving a name to the dataset. With `dim` we can see the dimension of a matrix/data frame.

```
cars.data <- cars
dim(cars.data)
```

```
## [1] 50  2
```

#We also see that the varibles are `speed` and `distance` . Both of them are numbers.

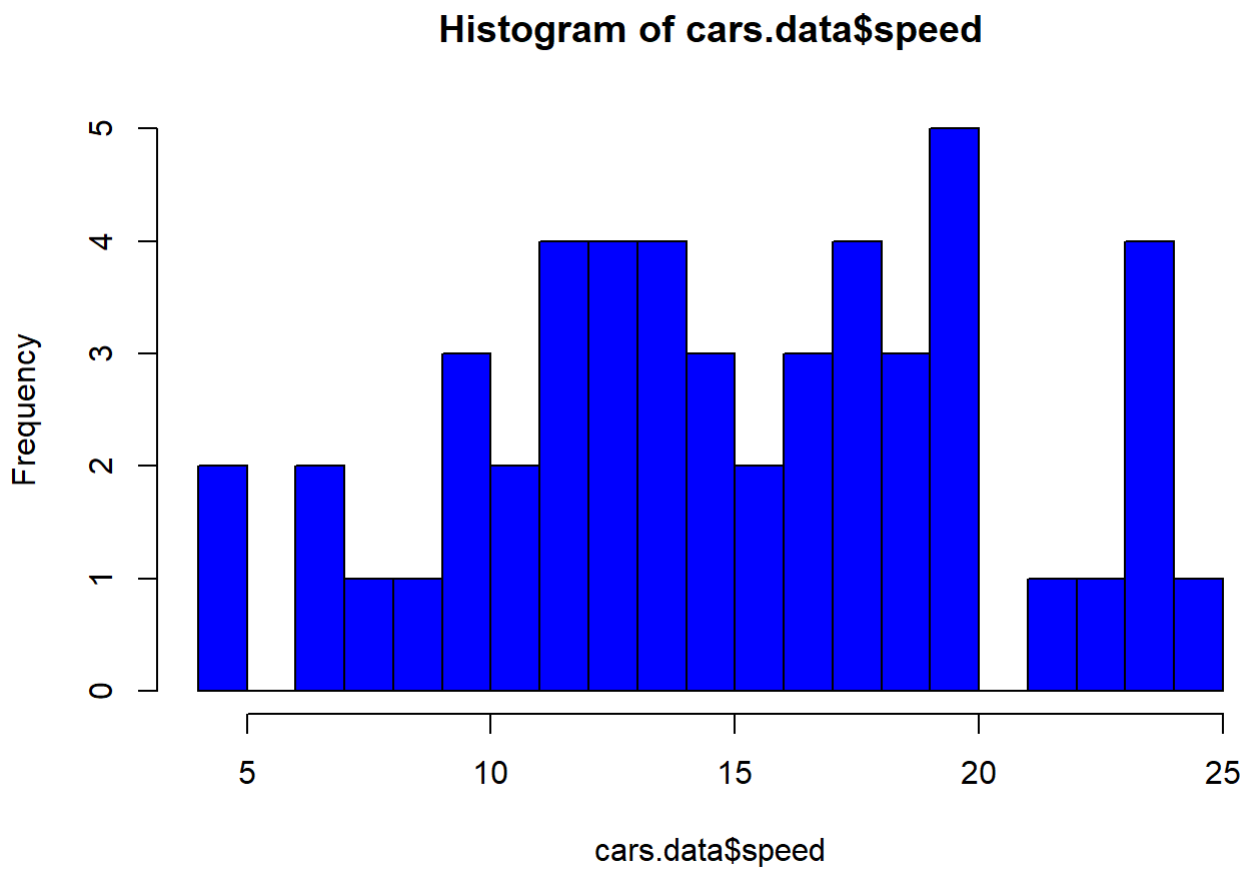
```
str(cars.data)
```

```
## 'data.frame':    50 obs. of  2 variables:
##  $ speed: num  4 4 7 7 8 9 10 10 10 11 ...
##  $ dist : num  2 10 4 22 16 10 18 26 34 17 ...
```

Plotting

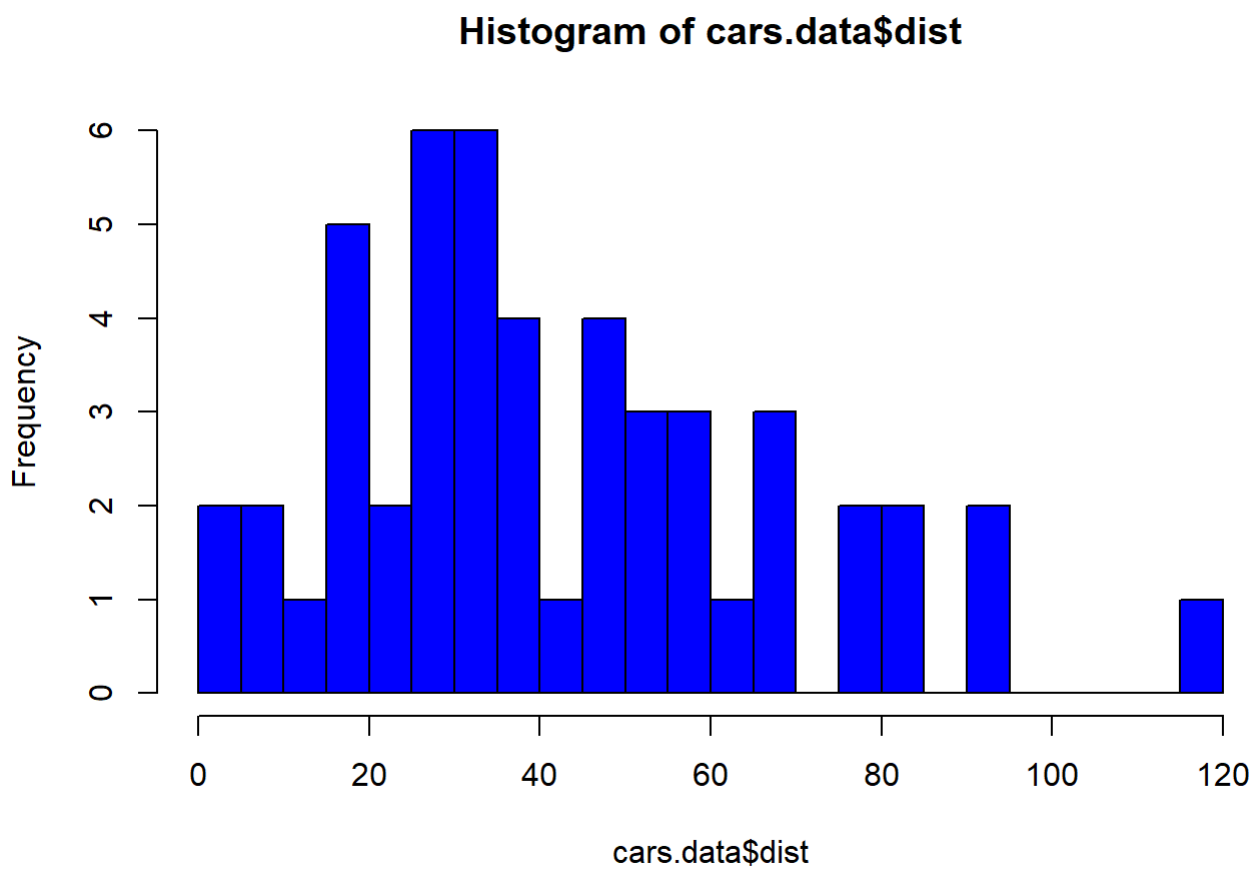
#We will begin by making a `histogram` .We can see that most of the cars have a top speed of `18-20` km.

```
hist(cars.data$speed, 20, col="blue")
```



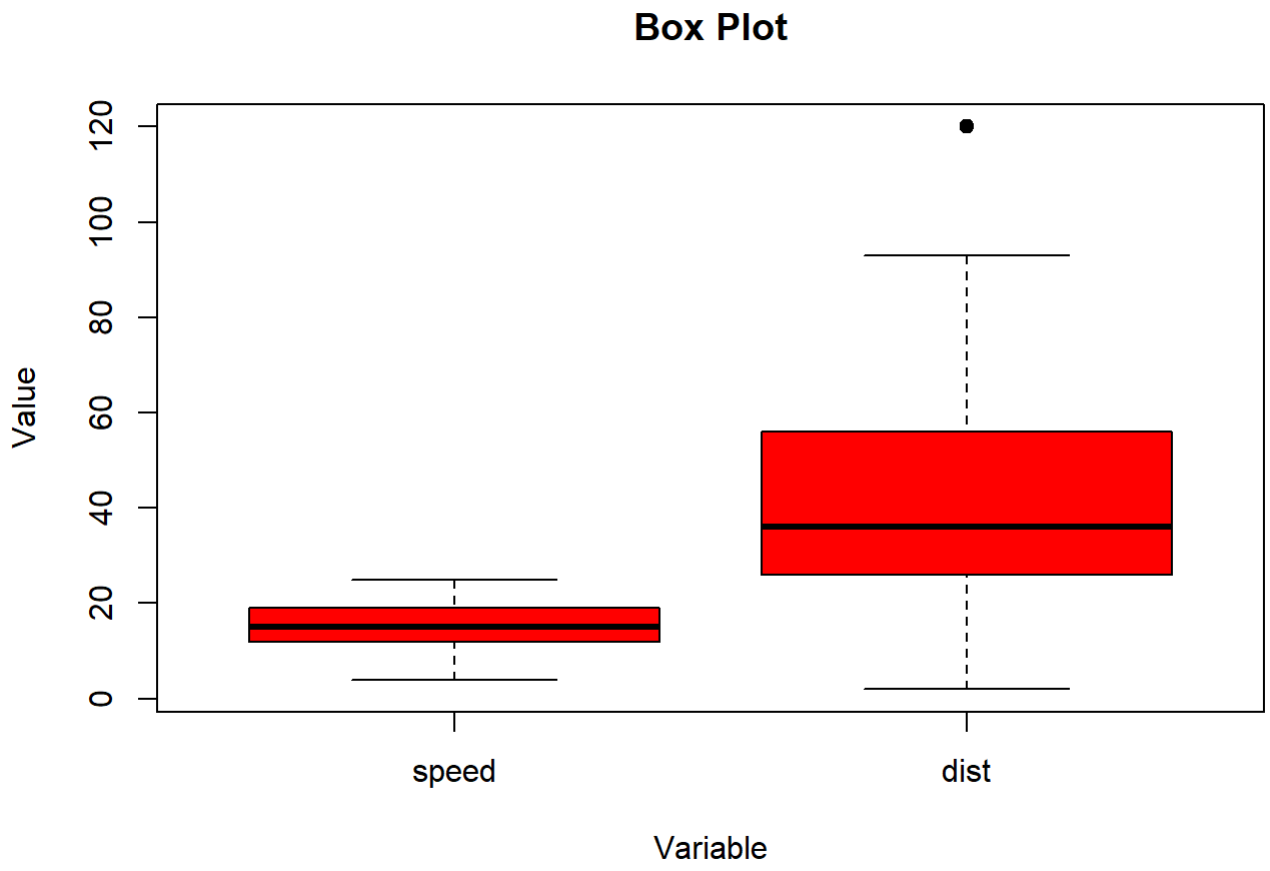
#We can also see that the distance travelled is around `20km`

```
hist(cars.data$dist, 20, col="blue")
```



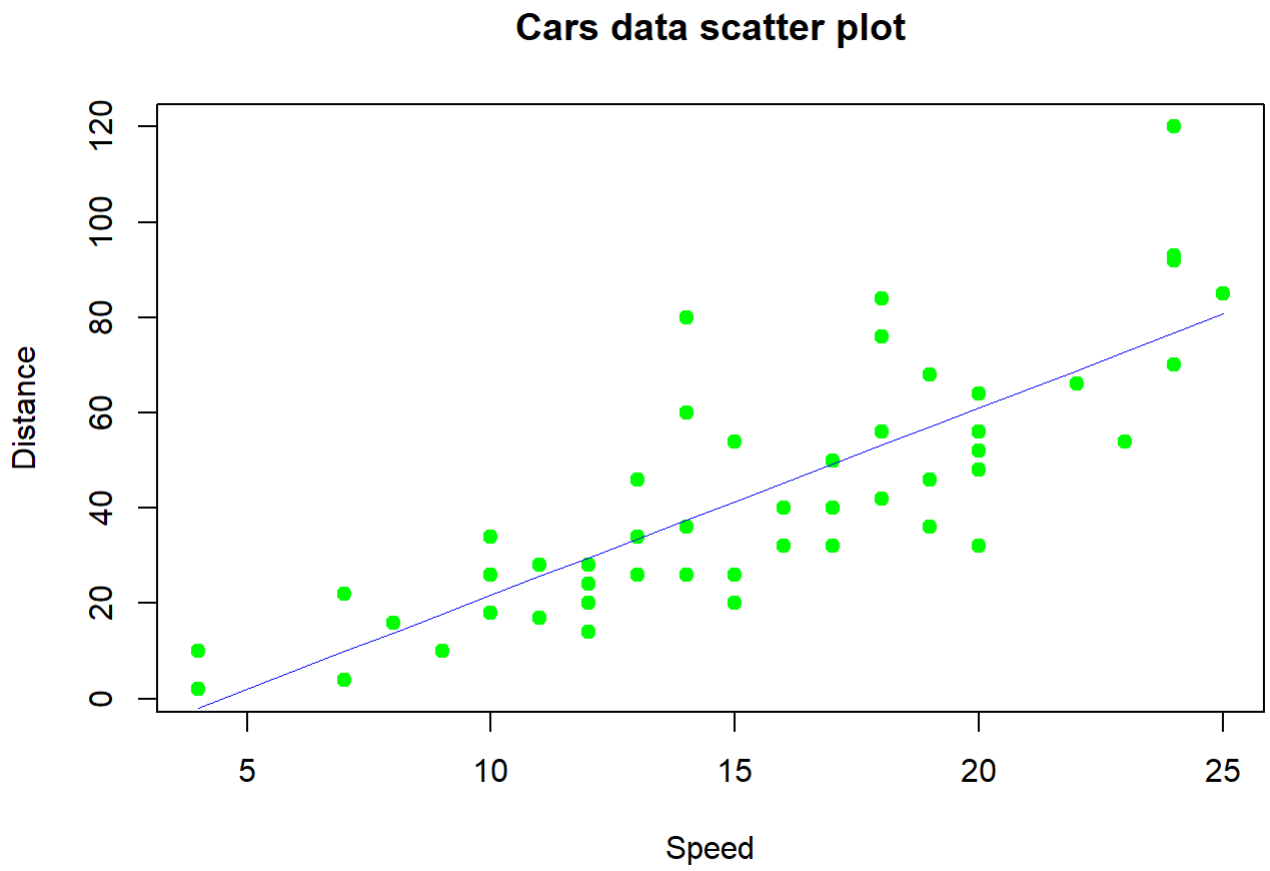
#Now we will make a box plot. We can see that `speed` has less variation compared to `distance` .

```
boxplot(cars.data, main="Box Plot",
 xlab="Variable", ylab="Value", col = "red", pch=19)
```



#Scatter plot shows us that the data is positively correlated. Speed depends on Distance or vice verca.

```
plot(x=cars.data$speed,y=cars.data$dist, main="Cars data scatter plot", xlab = "Speed",
 ylab = "Distance", pch=19, col="green")
lin.mod <- lm(cars.data$dist~cars.data$speed)
pr.lm <- predict(lin.mod)
lines(pr.lm~cars.data$speed, col="blue", lwd=0.7)
```


Cars data scatter plot
Distance
0
20
40
60
80
100
120
5
10
15
20
25
Speed