PROJECT WORK
ON
STATISTICAL DATA
ANALYTICS WITH R

Submitted To: -

Mr. Brijesh Dhyani

Submitted By: -

Saurabh Suman

MCA

Semester :- 1

Roll no: - 2001135

Objective Of this project : -

To study the R language and its use in the real world. To learn the use of R language to study a dataset and perform the analysis to predict or keep a record. Determining various statistical information from the dataset which and plotting the various types of graphs.

Description Of the Project: -

For this project I am using the data set of us weather report of the 2014 to 2015. We are using this data set for checking the pattern in the US weather condition.

Code for the Analysis Of the Data Set: -

```r
#setting the directory for a project
setwd("E:/R_project")


#Reading the Csv file and storing in mydata
mydata<-read.csv("UsWeatherReport.csv")


#checking the structure, the of the mydata
str(mydata)


#finding the mean of actual minimum temperature and maximum
temperature
mean_of_min_temp <- mean(mydata$actual_min_temp)
print(mean_of_min_temp)


mean_of_max_temp <- mean(mydata$actual_max_temp)
print(mean_of_max_temp)



#plotting the Bar graph of mean of temperature according to year
barplot(mydata$actual_mean_temp,xlab = "year", ylab =
"temperature" ,main = "year vs mean_temperature")
box()
```

#plotting the histogram of minimum temperature according to year

hist(mydata$actual_min_temp,xlab = "Minimum Temperature" ,main = "Frequency of Minimum Temperature")

box()

#box plot of average maximum temperature, year of minimum temperature and maximum temperature

boxplot(mydata$average_max_temp,mydata$actual_max_temp, mydata$average_min_temp, mydata$actual_min_temp,xlab ="Temperature", main ="comparison of average and real temperature")

box()

plot(mydata$average_min_temp ,type='l', main ="Average Minimum temperature")

box()

plot(mydata$average_min_temp,mydata$record_min_temp_year,xlab = "Average min Temperature" , ylab="Year")

#counting the temperature which is less than 70 for a pie chart

x <- length(which(mydata$actual_max_temp < 70))

#counting the temperature which is greater than 70 for a pie chart

x2 <- length(which(mydata$actual_max_temp >70))

#printing both the counter a and x2

```
print(x)
print(x2)


vec <- c(x,x2)
print(vec)


leb <- c("Temperature is less than 70", "Temperature is greater than
70")


#printing the pie chart of the maximum temperature in comparison of
less than 70 and more than 70

pie (vec,leb,main ="Maximum Temperature between 2014-15 in
USA")

box()
```
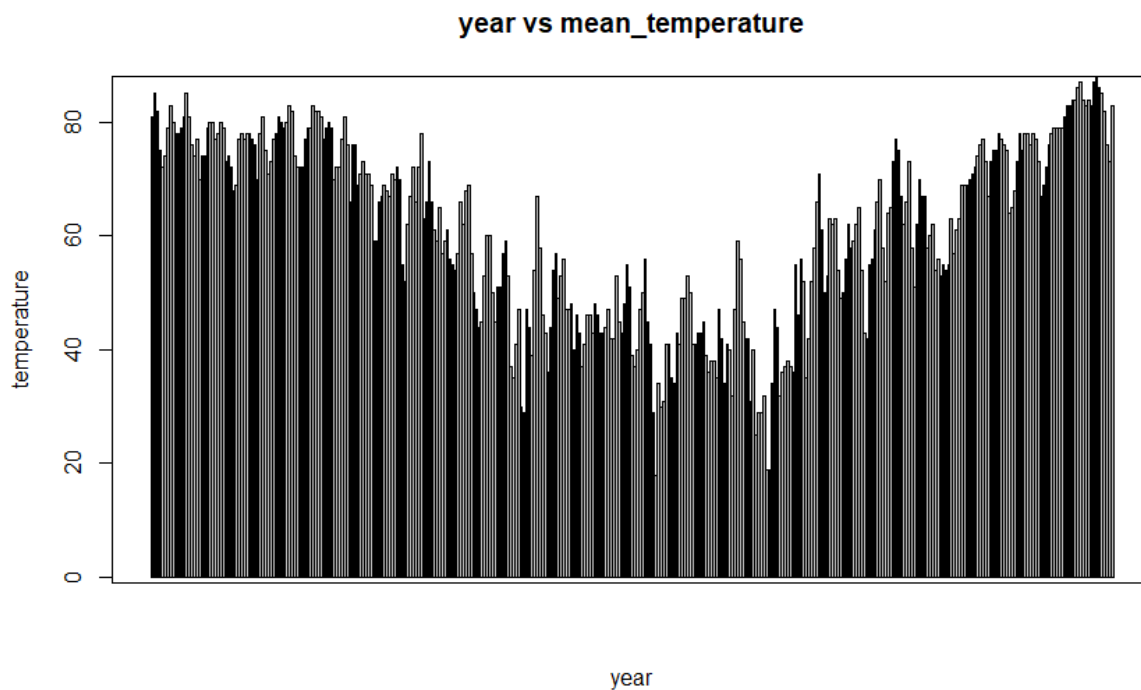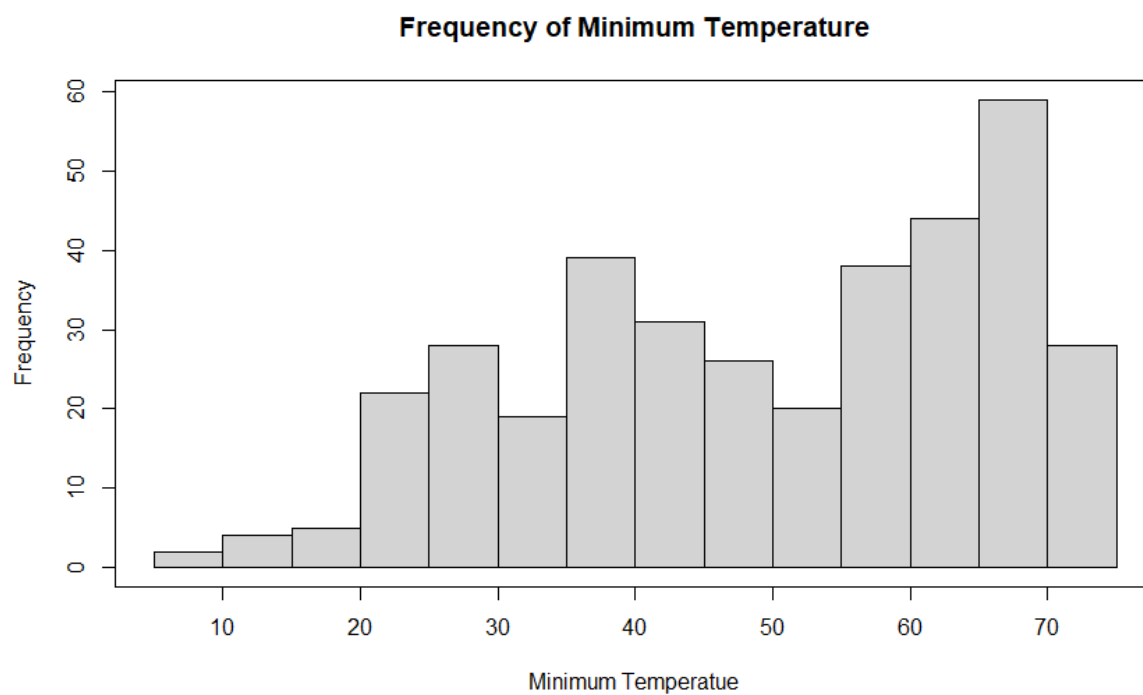
1. Bar Graph: -
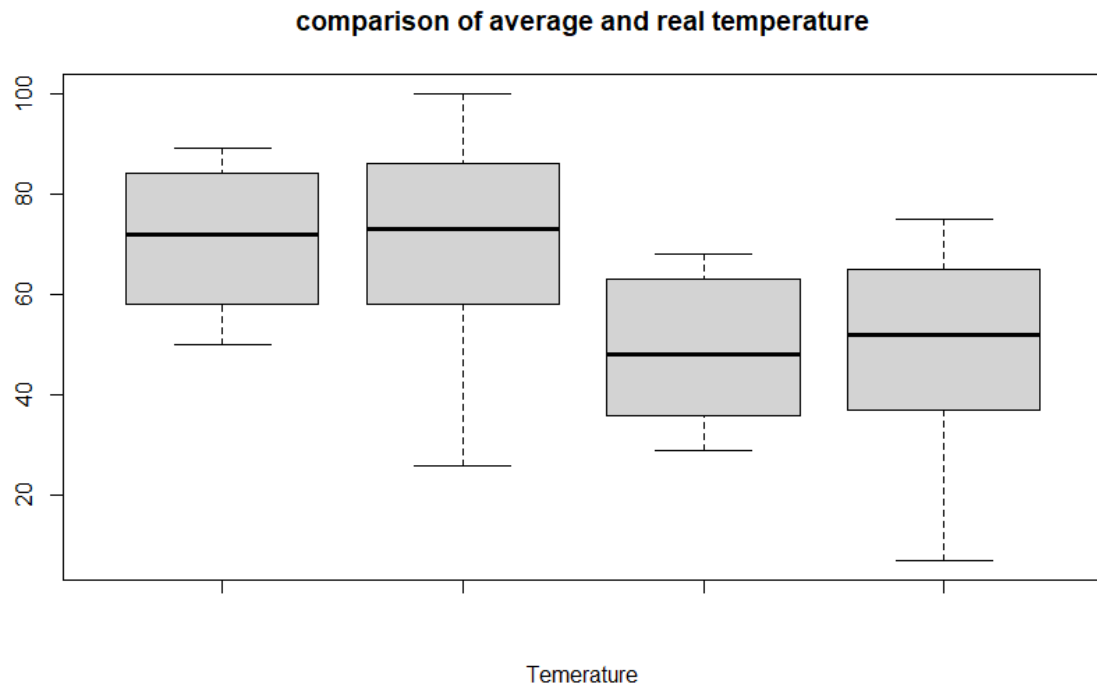
year vs mean_temperature

2. histogram: -

Frequency of Minimum Temperature

3. Box Plot: -

comparison of average and real temperature

Temerature

4. Line Graph: -

Average Minimum temperature

Index
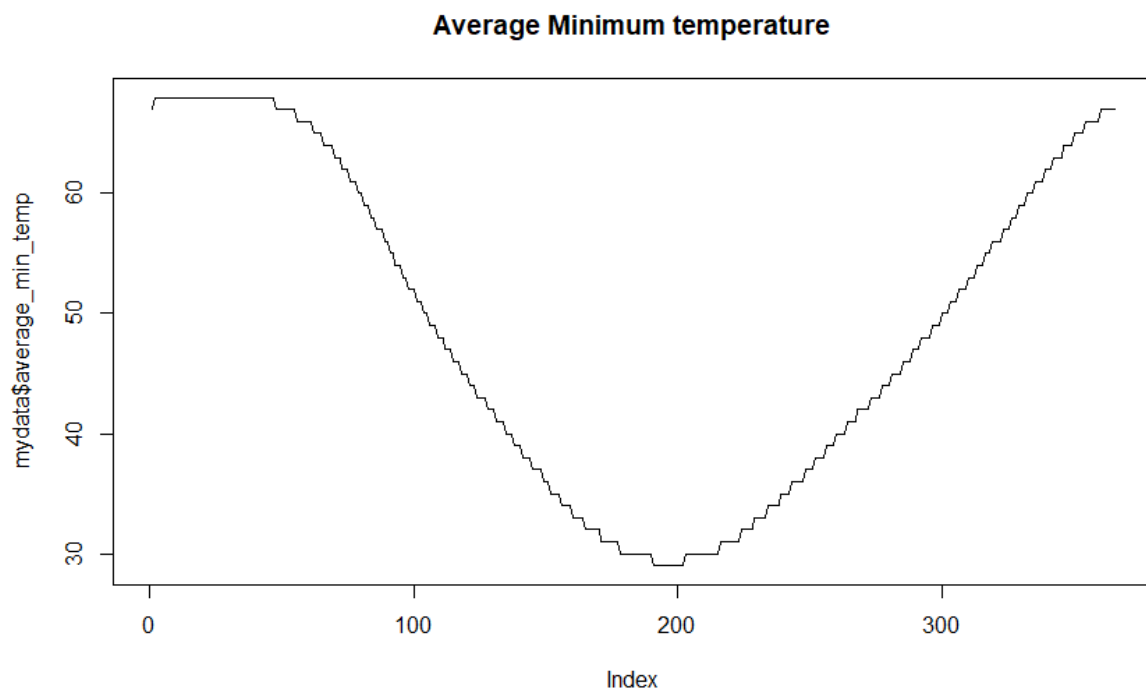
5. <u>Scatter Plot: -</u>
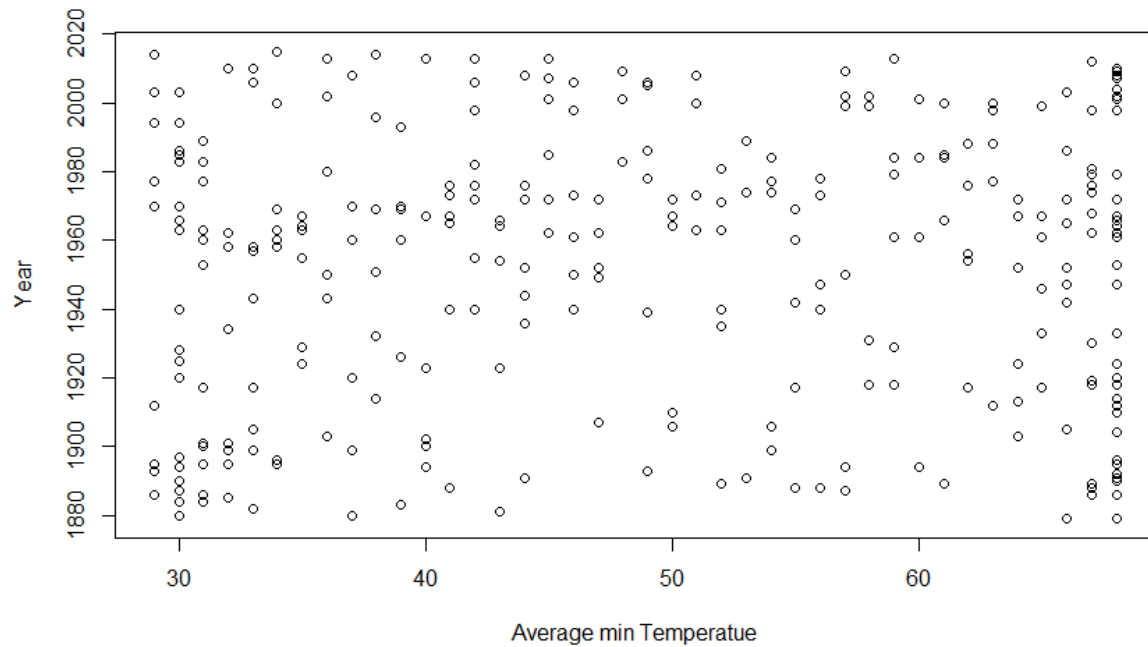
6. <u>Pie Chart: -</u>

Maximum Temperature

Temperature is less than 70

Temperature is greater than 70
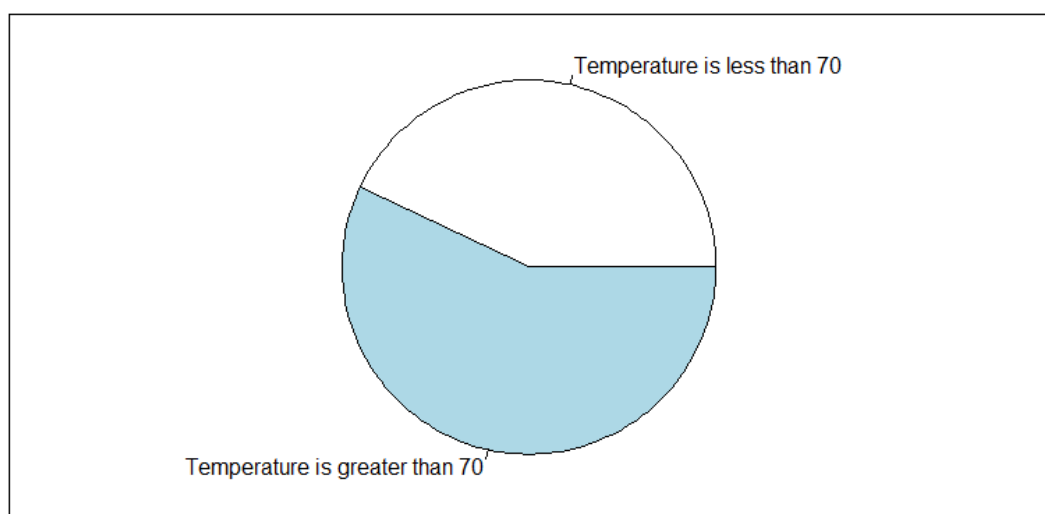
Console Output : -

```
Console   Terminal ×   Jobs ×                                                     ─ ⊡
E:/R_project/

> #setting the directory for a project
> setwd("E:/R_project")
> #Reading the Csv file and storing in mydata
> mydata<-read.csv("UsWeatherReport.csv")
> #checking the structure the of the mydata
> str(mydata)
'data.frame':   365 obs. of 13 variables:
 $ date                : chr  "2014-7-1" "2014-7-2" "2014-7-3" "2014-7-4" ...
 $ actual_mean_temp    : int  81 85 82 75 72 74 79 83 80 78 ...
 $ actual_min_temp     : int  70 74 71 64 60 61 67 72 71 71 ...
 $ actual_max_temp     : int  91 95 93 86 84 87 91 94 89 85 ...
 $ average_min_temp    : int  67 68 68 68 68 68 68 68 68 68 ...
 $ average_max_temp    : int  89 89 89 89 89 89 89 89 89 89 ...
 $ record_min_temp     : int  56 56 56 55 57 57 55 58 57 53 ...
 $ record_max_temp     : int  104 101 99 99 100 99 100 101 101 101 ...
 $ record_min_temp_year : int  1919 2008 2010 1933 1967 1964 1972 1892 1891 1961 ...
 $ record_max_temp_year : int  2012 1931 1931 1955 1954 1948 2010 1986 1926 ...
 $ actual_precipitation : num  0 0 0.14 0 0 0 0 0 0.15 0 ...
 $ average_precipitation: num  0.1 0.1 0.11 0.1 0.1 0.1 0.11 0.11 0.12 0.11 ...
 $ record_precipitation : num  5.91 1.53 2.5 2.63 1.65 1.95 2.37 1.87 3.71 2.45 ...
> #finding the mean of actual minimum temperature and maximum temperature
> mean_of_min_temp <- mean(mydata$actual_min_temp)
> print(mean_of_min_temp)
[1] 49.9589
> mean_of_max_temp <- mean(mydata$actual_max_temp)
> print(mean_of_max_temp)
[1] 71.63014
> barplot(mydata$actual_mean_temp,xlab = "year",ylab = "temperature" ,main = "year vs mean_temperature")
> box()
> hist(mydata$actual_min_temp)
> box()
> hist(mydata$actual_min_temp,xlab = "Minimum Temperatue")
> box()
> hist(mydata$actual_min_temp,xlab = "Minimum Temperatue" ,main = "Frequency of Minimum Temperature")
> box()
> boxplot(mydata$average_max_temp, mydata$average_min_temp, mydata$actual_max_temp, mydata$actual_min_tem
p)
> box()
> boxplot(mydata$average_max_temp,mydata$actual_max_temp, mydata$average_min_temp, mydata$actual_min_temp)
> box()
> #box plot of average maximum temperature, year of minimum temperature and maximum temperature
> boxplot(mydata$average_max_temp,mydata$actual_max_temp, mydata$average_min_temp, mydata$actual_min_temp,
 main ="comparison of average and real temperature")
> #box plot of average maximum temperature, year of minimum temperature and maximum temperature
> boxplot(mydata$average_max_temp,mydata$actual_max_temp, mydata$average_min_temp, mydata$actual_min_temp,
xlab ="Temerature", main ="comparison of average and real temperature")
> plot(mydata$average_min_temp ,type='l')
> box()
> plot(mydata$average_min_temp ,type='l',main ="Average Minimum temperature")
> box()
> plot(mydata$average_min_temp,mydata$record_min_temp,xlab = "Average min Temperatue", ylab = "Year")
> #counting the temperature which is less than 70 for a pie chart
> x <- length(which(mydata$actual_max_temp < 70))
> #counting the temperature which is greater than 70 for a pie chart
> x2 <- length(which(mydata$actual_max_temp >70))
> #printing both the counter a and x2
> print(x)
```

```
> box()
> boxplot(mydata$average_max_temp,mydata$actual_max_temp, mydata$average_min_temp, mydata$actual_min_temp)
> box()
> #box plot of average maximum temperature, year of minimum temperature and maximum temperature
> boxplot(mydata$average_max_temp,mydata$actual_max_temp, mydata$average_min_temp, mydata$actual_min_temp,
 main ="comparison of average and real temperature")
> #box plot of average maximum temperature, year of minimum temperature and maximum temperature
> boxplot(mydata$average_max_temp,mydata$actual_max_temp, mydata$average_min_temp, mydata$actual_min_temp,
xlab ="Temerature", main ="comparison of average and real temperature")
> plot(mydata$average_min_temp ,type='l')
> box()
> plot(mydata$average_min_temp ,type='l',main ="Average Minimum temperature")
> box()
> plot(mydata$average_min_temp,mydata$record_min_temp,xlab = "Average min Temperatue", ylab = "Year")
> #counting the temperature which is less than 70 for a pie chart
> x <- length(which(mydata$actual_max_temp < 70))
> #counting the temperature which is greater than 70 for a pie chart
> x2 <- length(which(mydata$actual_max_temp >70))
> #printing both the counter a and x2
> print(x)
[1] 152
> print(x2)
[1] 202
> vec <- c(x,x2)
> print(vec)
[1] 152 202
> #printing the pie chart of the maximum temeperature in comparison of less than 70 and more than 70
> pie(vec)
> #printing the pie chart of the maximum temperature in comparison of less than 70 and more than 70
> pie(vec,lab)
Error in pie(vec, lab) : object 'lab' not found
> #printing the pie chart of the maximum temperature in comparison of less than 70 and more than 70
> pie(vec,leb)
Error in pie(vec, leb) : object 'leb' not found
> leb <- labels("less than 70", "greater than 70")
> #printing the pie chart of the maximum temperature in comparison of less than 70 and more than 70
> pie(vec,leb)
> leb <- c("less than 70", "greater than 70")
> #printing the pie chart of the maximum temperature in comparison of less than 70 and more than 70
> pie(vec,leb)
> pie(vec,leb,main ="comparing temperature")
> box()
> leb <- c("Temperature is less than 70", "Temperature is greater than 70")
> pie(vec,leb,main ="comparing temperature")
> box()
> pie(vec,leb,main ="Maximum Temperature")
> box()
> pie(vec,leb,main ="Maximum Temperature between 2014-15")
> box()
> pie(vec,leb,main ="Maximum Temperature between 2014-15 in USA")
> box()
> plot(mydata$average_min_temp,mydata$record_min_temp,xlab = "Average min Temperatue")
> #counting the temperature which is less than 70 for a pie chart
> x <- length(which(mydata$actual_max_temp < 70))
> plot(mydata$average_min_temp,mydata$record_min_temp,xlab = "Average min Temperatue" ,ylab="")
> plot(mydata$average_min_temp,mydata$record_min_temp_year,xlab = "Average min Temperatue" ,ylab="Year")
>
```

Inference from the statistical data from the above data set: -

1. Mean of the minimum temperature   49.9589
2. Mean of the maximum temperature   71.63014
3. From the above data analysis, we can say that temperature is gradually increasing.
4. Minimum temperature is recorded in 1991 which is 56.
5. Maximum temperature is recorded in 2012 which is 104.