PROGRAMMING FOR DATA ANALYSIS

- **PRESENTED BY: PRESENTED TO:**
- PURUSHOTTAM SAH   MR. R.N. THAKUR
- NP000327   LECTURER
- SECTION B
- IV SEMESTER

ANALYSIS ON WEATHERDATA.CSV FILE OF LAGUARDIA AIRPORT (LGA) AND JOHN F. KENNEDY INTERNATIONAL AIRPORT (JFK)

INTRODUCTION

- R IS DEFINED AS A PROGRAMMING LANGUAGE THAT SET-UPS THE ENVIRONMENT FOR STATISTICAL AND GRAPHICAL DATA COMPUTATION.

- PROPOSED BY JOHN CHAMBERS AND HIS COLLEAGUES IN BELL LABORATORIES.

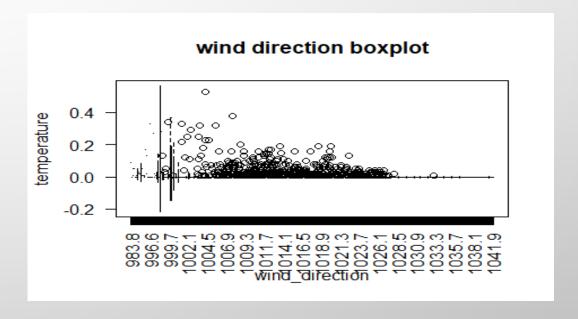- THE R LANGUAGE IS A PROJECT DESIGNED TO CREATE A FREE, OPEN SOURCE LANGUAGE

ASSUMPTIONS

• THE DATA THAT IS PROVIDED IN HOURLY WEATHER DATA THAT CONTAINS HUMIDITY, WIND DIRECTIONS, PRESSURE AS WELL AS VISIBILITY, WIND SPEED AND WIND GUST, TEMPERATURE, DEW POINTS.

• INCREASING OF WIND GUST, INCREASE IN WIND SPEED. AFTER WIND SPEED, THE WIND GUST IS SUDDENLY CHANGED.

• BOTH HUMIDITY AND PRECIPITATION ARE LINKED WITH EACH OTHER. IF THE HUMIDITY IS MORE THEN THE PRECIPITATION IS ALSO AUTOMATICALLY HIGH.
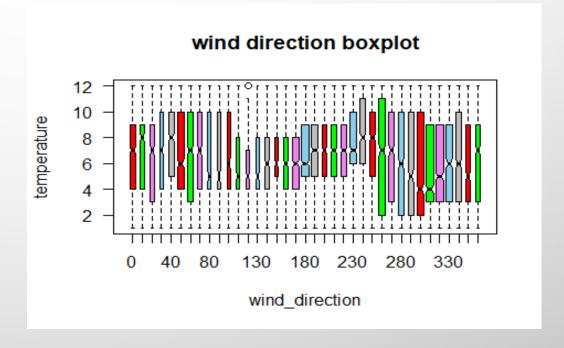
ANALYSIS

ANALYSIS 1: ANALYSIS BETWEEN WIND DIRECTION AND TEMPERATURE
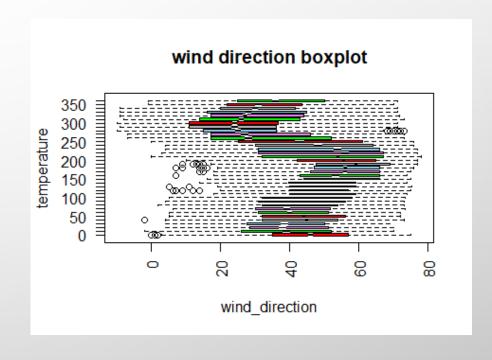
CODE SNIPPET   OUTPUT


```
#TP no: NP000327
val<-read.csv("weatherdata.csv")
boxplot(val$precip~val$pressure,
        main="wind direction boxplot",
        xlab = "wind_direction",
        ylab = "temperature",
        vertical=TRUE,
        las=2,varwidth=TRUE,notch=TRUE,
        col=c("#00FFFF","#5E5A80","#800000",
              "#FFF380","#307D7E", "#2B1B17"))
```

ANALYSIS 2: ANALYSIS OF WIND DIRECTION IN BOX PLOT

CODE SNIPPET   OUTPUT

```
#Name: Purushottam Sah
#TP no: NP000327
val<-read.csv("weatherdata.csv")
boxplot(val$month~val$wind_dir,
        main= "wind direction boxplot",
        xlab = "wind_direction",
        ylab = "temperature",
        las=1, varwidth=TRUE, notch=TRUE,
        col = c("red","green","violet","skyblue","gray"))
meanval<-by(val$temp,val$wind_dir,mean)
points(meanval,col="white", pch=6, cex=1.5)
```

ANALYSIS 3: HORIZONTAL BOXPLOT OF WIND DIRECTION

CODE SNIPPET   OUTPUT

```
val<-read.csv("weatherdata.csv")
boxplot(val$dewp~val$wind_dir,
        main= "wind direction boxplot",
        xlab = "wind_direction",
        ylab = "temperature",
        horizontal = TRUE,
        las=2, varwidth=TRUE, notch=TRUE,
        col = c("red","green","violet","skyblue","gray"))
```
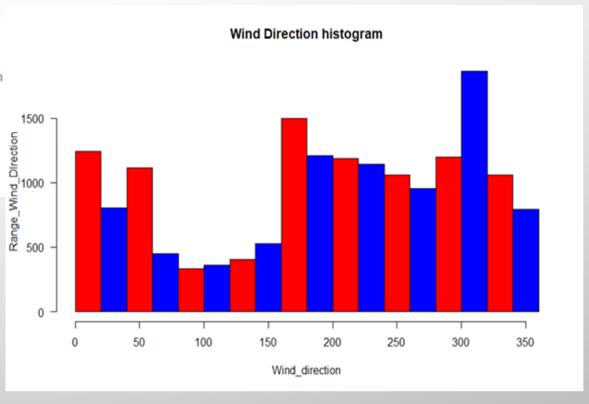
ANALYSIS 4: HISTOGRAM REPRESENTATION OF WIND DIRECTION

CODE SNIPPET

OUTPUT

```
#Name:Ajay Kumar Sah
#TP no:NP000292
#Analysis 4
#An analysis of wind direction data using histogram data visualization
val<-read.csv("weatherdata.csv")
hist(val$wind_dir,
    main = "Wind Direction histogram",
    xlab = "Wind_direction",
    ylab = "Range_Wind_DIrection",
    las=1,
    col = c("red", "blue"))
```


Wind Direction histogram
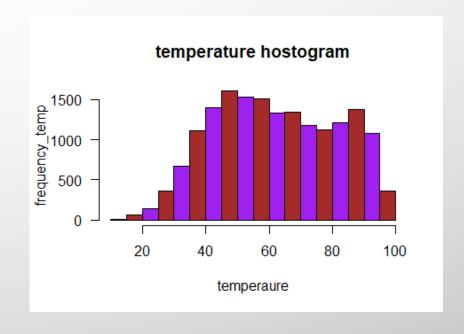
ANALYSIS 5: ANALYSIS OF TEMPERATURE USING HISTOGRAM

CODE SNIPPET   OUTPUT
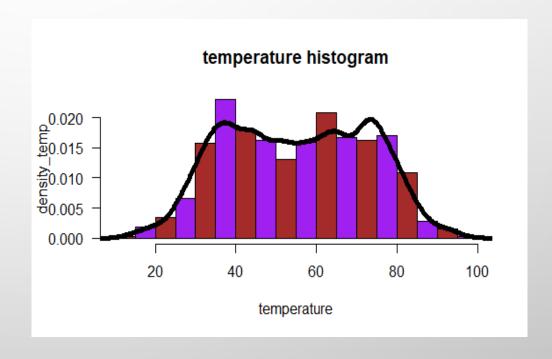
```
val<-read.csv("weatherdata.csv")
hist(val$humid,
    main = "temperature hostogram",
    xlab = "temperaure",
    ylab = "frequency_temp",
    las = 1,
    col = c("purple","brown"))
```

ANALYSIS 6: LINE GRAPH REPRESENTATION OF TEMPERATURE DATA

CODE SNIPPET

OUTPUT

```r
#Name: Purushottam Sah
#TP no: NP000327
val<-read.csv("weatherdata.csv")
hist(val$temp,
     freq = FALSE,
     main = "temperature histogram",
     xlab = "temperature",
     ylab = "density_temp",
     las = 1,
     col = c("Brown","Purple"))
lines(density(val$temp), lwd = 5, col = "black")
```
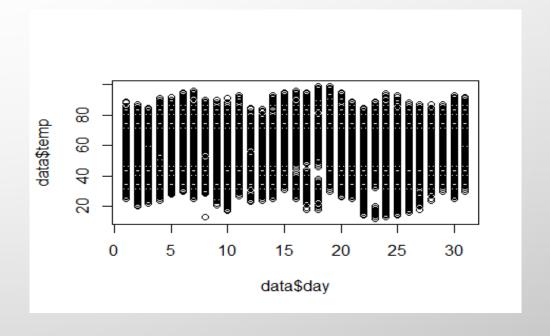
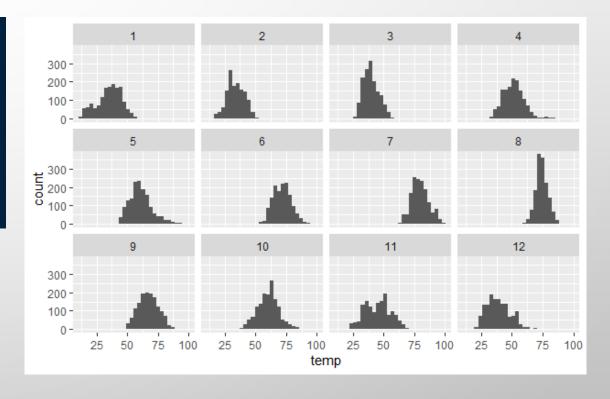# ANALYSIS 7: ANALYSIS BETWEEN DAY AND TEMPERATURE

CODE SNIPPET

OUTPUT

```
data<-read.csv("weatherdata.csv")
plot(x=data$day, y=data$temp)
```

ANALYSIS 8: ANALYSIS BETWEEN MONTH AND TEMPERATURE

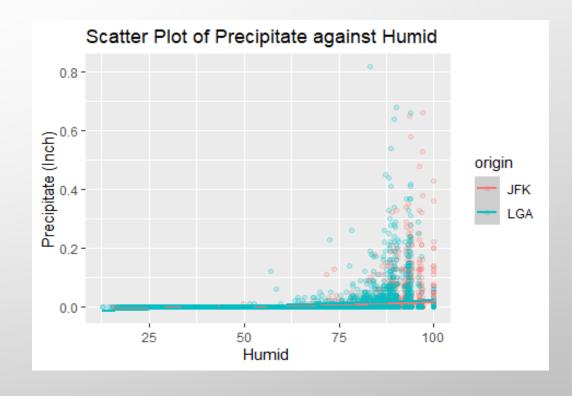CODE SNIPPET   OUTPUT

```
data<-read.csv("weatherdata.csv")
print(data)
ggplot(data = data, mapping = aes(x = temp)) +
  geom_histogram() +
  labs(title = 'Histogram of Temperature',x = 'Temperature (F)') +
facet_wrap(~month)
```

ANALYSIS 9: ANALYSIS OF WIND SPEED

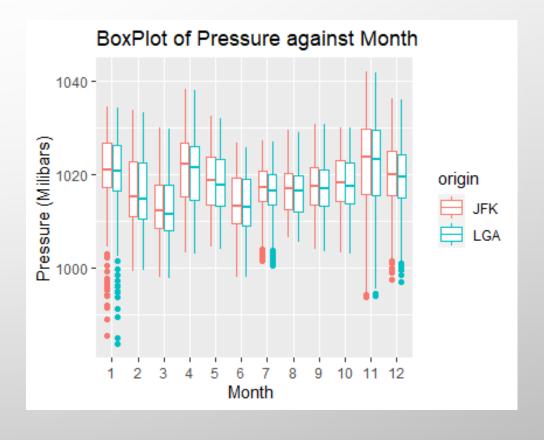CODE SNIPPET   OUTPUT

```
ggplot(data = data, mapping = aes(x = humid, y = precip, color = origin)) +
  geom_point(alpha = 0.2) + stat_smooth(method = "lm") +
  labs(title = 'Scatter Plot of Precipitate against Humid',x = 'Humid ', y = 'Precipitate (Inch)')
```


Scatter Plot of Precipitate against Humid

ANALYSIS 10: ANALYSIS BETWEEN PRESSURE AND MONTH
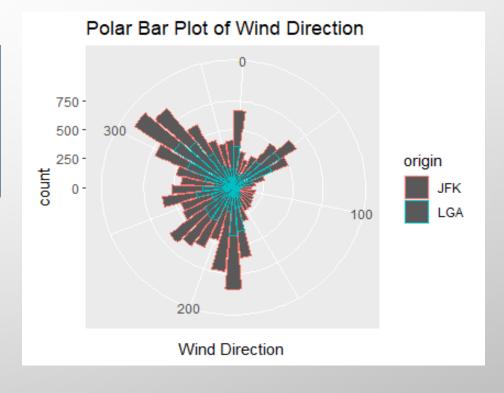
CODE SNIPPET

OUTPUT

```
ggplot(data = data, mapping = aes(x = factor(month), y = pressure, na.rm = TRUE, color=origin)) +
  geom_boxplot() +
  labs(title = 'BoxPlot of Pressure against Month',x = 'Month ', y = 'Pressure (Milibars)')
```


BoxPlot of Pressure against Month

ANALYSIS 11: ANALYSIS OF POLAR BAR POINT OF WIND DIRECTION

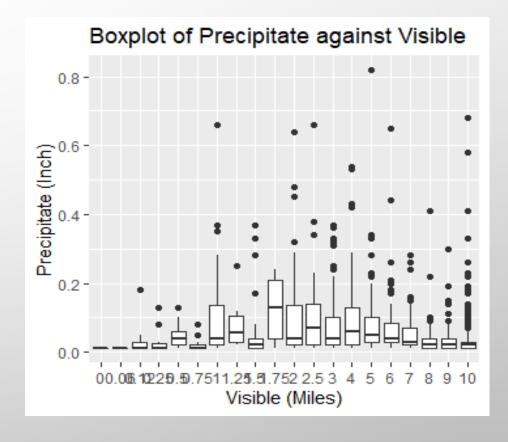CODE SNIPPET   OUTPUT

```
#NP000327
ggplot(data = data) +
  geom_bar(mapping = aes(x = wind_dir, na.rm = TRUE, color=origin)) +
  coord_polar() +
  labs(title = 'Polar Bar Plot of Wind Direction',x = 'Wind Direction ')
```


Polar Bar Plot of Wind Direction

ANALYSIS 12: ANALYSIS BETWEEN WIND SPEED AND WIND GUST
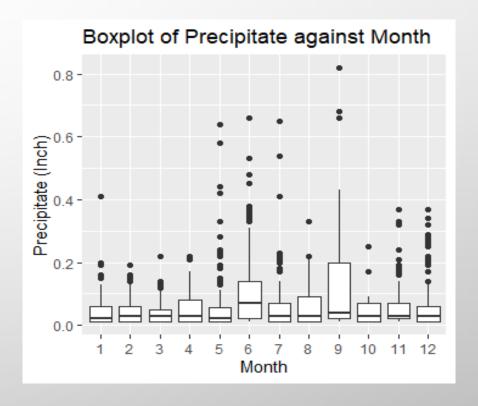
CODE SNIPPET

OUTPUT

```
data %>%
  filter(data$precip>0) %>%
  ggplot(mapping = aes(x = factor(visib), y = precip)) +
  geom_boxplot() +
  labs(title = 'Boxplot of Precipitate against Visible',x = 'Visible (Miles)', y = 'Precipitate (Inch)')
```


Boxplot of Precipitate against Visible

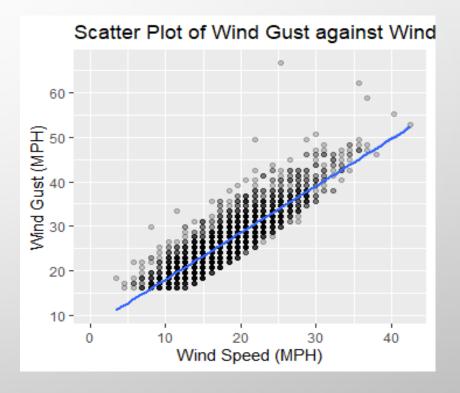# ANALYSIS 13: ANALYSIS PRECIPITATE AND MONTH

CODE SNIPPET

OUTPUT

```
data %>%
  filter(data$precip>0) %>%
  ggplot(mapping = aes(x = factor(month), y = precip)) +
  geom_boxplot() +
  labs(title = 'Boxplot of Precipitate against Month',x = 'Month', y = 'Precipitate (Inch)')
```


Boxplot of Precipitate against Month

ANALYSIS 14: ANALYSIS OF SCATTER PLOT BETWEEN WIND GUST AND WIND SPEED

CODE SNIPPET   OUTPUT

```
ggplot(data = data, mapping = aes(x = wind_speed, y = wind_gust, na.rm = TRUE)) +
  geom_point(alpha = 0.2)+ stat_smooth(method = "lm") +
  labs(title = 'Scatter Plot of Wind Gust against Wind Speed',x = 'Wind Speed (MPH)', y = 'Wind Gust (MPH)')
```

CONCLUSION

THE ASSIGNMENT ATTACHED WITH THIS MODULE WAS VERY HELPFUL IN UNDERSTANDING THE CONCEPT OF PACKAGES AND LIBRARIES IN R PROGRAMMING LANGUAGE. THIS MODULE HELPED A LOT IN EXPANDING MY KNOWLEDGE IN THE FIELD OF DATA ANALYSIS. THIS MODULE TAUGHT ME THAT HOW A REAL DATA IS PRESENTED IN REAL WORLD THE GIVEN HOUR WEATHER DATA IS PLOTTED THAT HELPS TO SHOW THE RELATIONSHIP BETWEEN DIFFERENT ELEMENT.

REFERENCES

R-PROJECT.ORG. 2021. R: WHAT IS R?. [ONLINE]

AVAILABLE AT: <HTTPS://WWW.R-PROJECT.ORG/ABOUT.HTML>

[ACCESSED 10 JANUARY 2021].

DATAMENTOR. 2021. R TUTORIAL - LEARN R PROGRAMMING. [ONLINE]

AVAILABLE AT: <HTTPS://WWW.DATAMENTOR.IO/R-PROGRAMMING/>

[ACCESSED 10 JANUARY 2021].

HOLTZ, Y., 2021. THE R GRAPH GALLERY – HELP AND INSPIRATION FOR R CHARTS. [ONLINE] THE R GRAPH GALLERY. AVAILABLE AT: <HTTPS://WWW.R-GRAPH-GALLERY.COM/> [ACCESSED 20 JANUARY 2021].

THANK YOU