INDIVIDUAL ASSIGNMENT

TECHNOLOGY PARK MALAYSIA

CT127-3-2-PFDA

PROGRAMMING FOR DATA ANALYSIS

UC2F2008IT(ISS)

HAND OUT DATE: 2020

HAND IN DATE: 22 NOVEMBER 2020

WEIGHTAGE: 50%

NAME: TAN YONG QI (TP051880)

LECTURER NAME: MARY TING

Contents

Introduction and Assumption

Exploring and categorise hourly weather data set by various techniques in such a way that it can retrieve the appropriate data that helps to make a decision. The analysis that be done should be deep and detail. In this assignment, I am importing the data from the csv file which named as “Hourly Weather Data” to the RStudio to make an analysis. The dataset is related to two different airports in United States which is LaGuardia Airport (LGA) and John. F Kennedy International Airport (JFK). I had applied the data visualization, exploration, and manipulation techniques in this project. I assume that the data manipulation and data visualisation that I did for this assignment is understandable and the graph I created is easy to understand and the information generated in the graph is completed. The graph that I created in this assignment is histogram, scatter plot and boxplot. The explanation that I did for each analysis is assumed to be easy to understand by everyone. Lastly, if the data being well analysed, it can benefit a lot of people and help them to make a right decision.

2.0 Import Data

```
setwd("/Users/User/Desktop/PFDA/PFDA Individual Assignment/PFDA Individual Assignment")
data = read.csv("4. Hourly weather data.csv")
library(ggplot2)
library(dplyr)
summary(data)
```

```
> summary(data)
    origin               year          month             day              hour            temp            dewp           humid       
 Length:17412       Min.   :2013   Min.   : 1.000   Min.   : 1.00   Min.   : 0.00   Min.   :12.02   Min.   :-9.94   Min.   : 12.74  
 Class :character   1st Qu.:2013   1st Qu.: 4.000   1st Qu.: 8.00   1st Qu.: 6.00   1st Qu.:39.92   1st Qu.:26.06   1st Qu.: 46.85  
 Mode  :character   Median :2013   Median : 7.000   Median :16.00   Median :11.00   Median :55.04   Median :42.08   Median : 61.15  
                    Mean   :2013   Mean   : 6.504   Mean   :15.68   Mean   :11.49   Mean   :55.12   Mean   :41.23   Mean   : 62.26  
                    3rd Qu.:2013   3rd Qu.: 9.000   3rd Qu.:23.00   3rd Qu.:17.00   3rd Qu.:69.98   3rd Qu.:57.02   3rd Qu.: 78.66  
                    Max.   :2013   Max.   :12.000   Max.   :31.00   Max.   :23.00   Max.   :98.96   Max.   :78.08   Max.   :100.00  

    wind_dir       wind_speed       wind_gust         precip            pressure         visib         time_hour     
 Min.   :  0.0   Min.   : 0.000   Min.   :16.11   Min.   :0.000000   Min.   : 983.8   Min.   : 0.000   Length:17412  
 1st Qu.:120.0   1st Qu.: 6.905   1st Qu.:21.86   1st Qu.:0.000000   1st Qu.:1012.9   1st Qu.:10.000   Class :character  
 Median :220.0   Median :10.357   Median :25.32   Median :0.000000   Median :1017.7   Median :10.000   Mode  :character  
 Mean   :201.9   Mean   :11.046   Mean   :26.18   Mean   :0.004183   Mean   :1017.9   Mean   : 9.245                     
 3rd Qu.:300.0   3rd Qu.:14.960   3rd Qu.:29.92   3rd Qu.:0.000000   3rd Qu.:1023.1   3rd Qu.:10.000                     
 Max.   :360.0   Max.   :42.579   Max.   :66.75   Max.   :0.820000   Max.   :1042.1   Max.   :10.000                     
 NA's   :204     NA's   :3        NA's   :13877                      NA's   :1794                                        
```

Figure 1: Import Data

Before starting the analysis, we need to import the data to the RStudio from our desktop. Then, we need to install and load the “ggplot2” and “dplyr” function using library. After importing the data and load the package, we can summary the data to check whether our data is successfully import or not.

2.1 Analysis 1 Temperature

```
#Analysis 1 Temperature
ggplot(data,aes(x=temp)) +
  geom_histogram()+
  labs(title = 'Histogram of Temperature', x = 'Temperature (F)') +
  facet_wrap(~month)+
  theme_bw() + theme(panel.background = element_rect(fill = "darkseagreen1"), plot.background = element_rect(fill = "darkseagreen"))
```

Figure 2: Temperature (Code)

For the analysis one, I am using ggplot function to plot the histogram. The first line of the code is means that the graph is form with the temperature from the dataset. The second line is the type of graph that I choose. The third line is I am naming the histogram with 'History of Temperature' and x-axis represent 'Temperature'. The fourth line is I am arranging and showing the histogram in 12 months and the code for last line is the design of the graph's background.

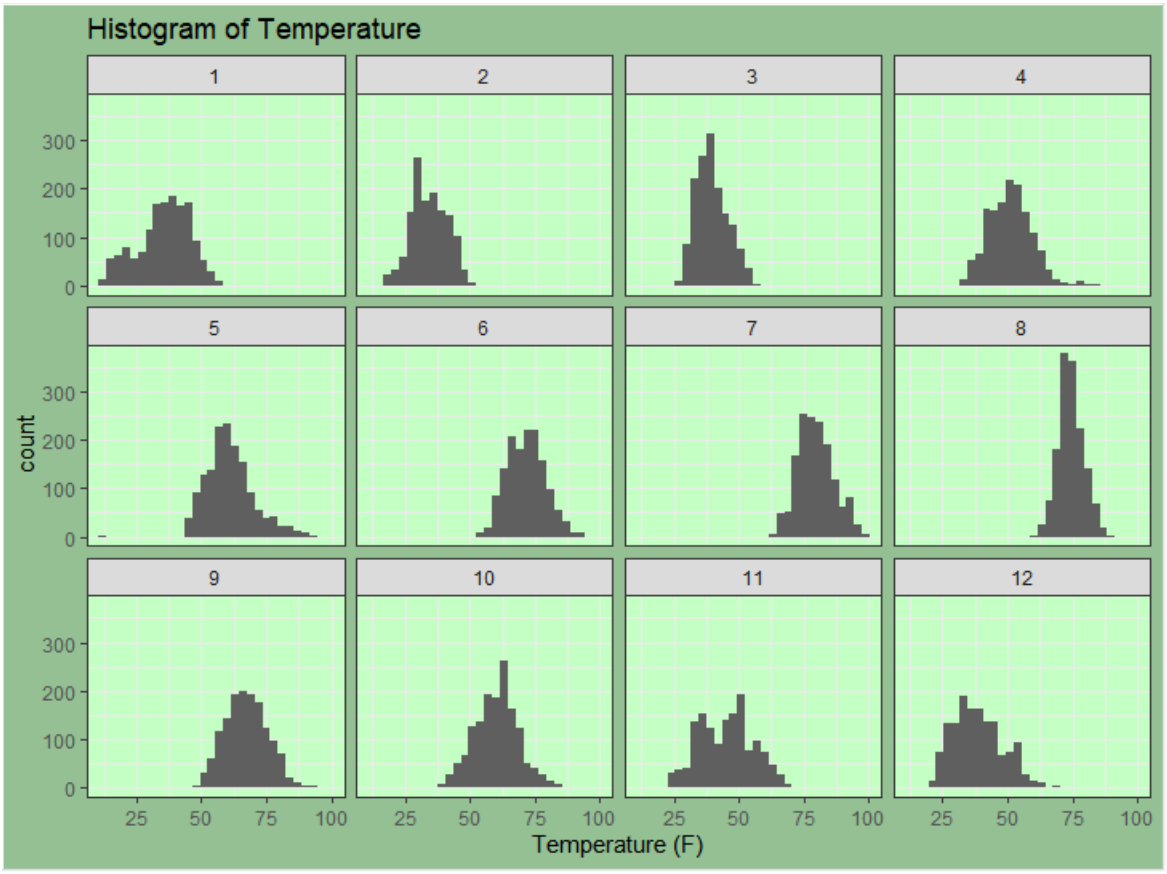

Figure 3: Temperature

The histogram above is showing the temperature of JFK airport and LGA airport along 12 months of year 2013. From the histogram we can know that the highest temperature of two airport is July and the lowest temperature is January.

2.2 Analysis 2 Wind Speed

```
#Analysis 2 Wind Speed
ggplot(data,aes(x = wind_speed, na.rm = TRUE)) +
  geom_histogram() +
  labs(title = 'Histogram of Wind speed',x = 'Wind speed (MPH)') +
  facet_wrap(~origin)+
  theme_bw() + theme(panel.background = element_rect(fill = "darkseagreen1"), plot.background = element_rect(fill = "darkseagreen"))
```

Figure 4: Wind Speed (Code)

For the analysis 2, I am using ggplot function to plot the histogram. The first line of the code is means that the graph is form with the wind speed from the dataset. Remove not available function also including after fetch the data from dataset. The second line is the type of graph that I choose. The third line is I am naming the histogram with 'History of Wind Speed' and x-axis represent 'Wind Speed (MPH)'. The fourth line is I am arranging and showing the histogram with two different airports and the code for last line is the design of the graph's background.

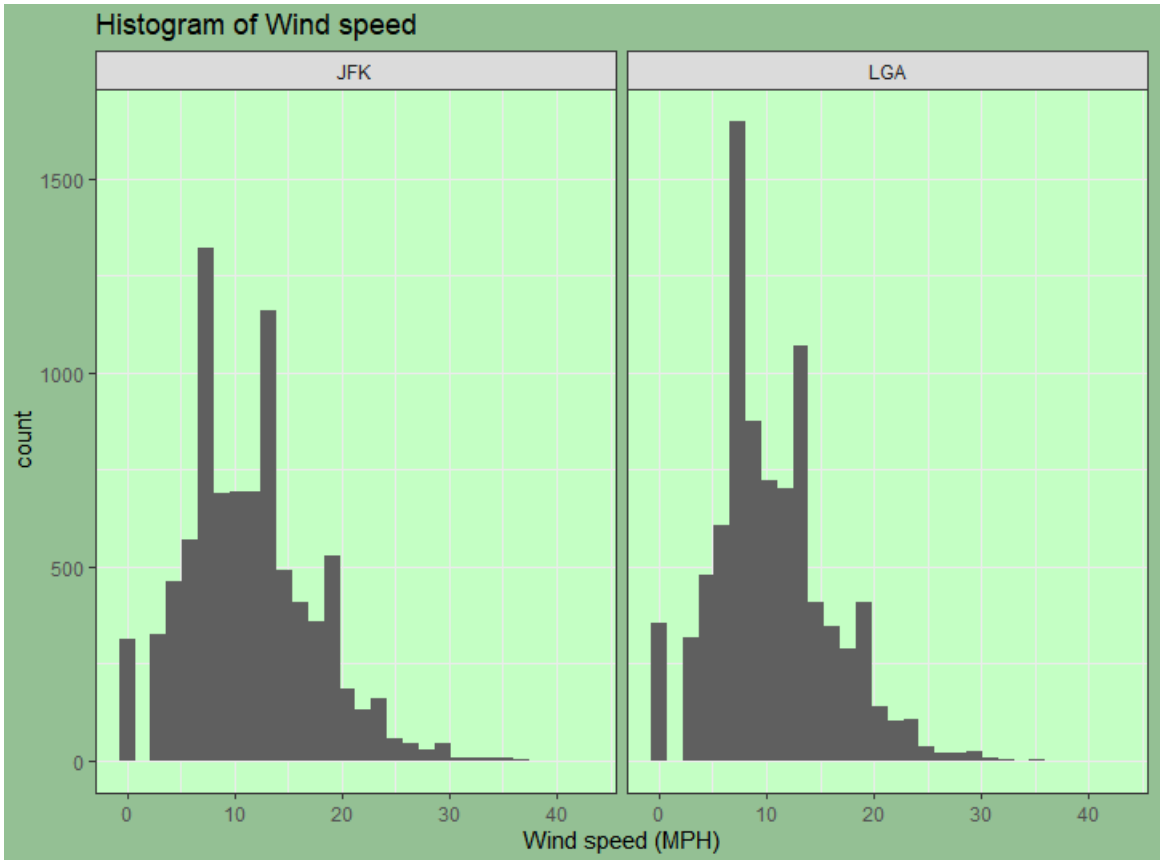

Figure 5: Wind Speed

The histogram above is showing wind speed of two airport. Most of the wind speed is in the range of 0-30 mph. According to a researcher [CITATION RobertSchrader2019 \l 17417], the high-speed winds will lead to a flight delay even take-off or landing. The reason is every

flight will deal with high speed winds at some point during its climb or descent The take-off and landing of horizontal winds (also known as crosswinds) in excess of 30-35 kts (about 34-40 mph) is typically prohibitive Owing to extremely high wind speeds, an aeroplane can have to abandon its landing As to how this works, it depends on where you are on the aircraft If crosswinds are high while the plane is at the gate, air traffic controllers will simply delay departure, as they would during heavy snow. As the result, the wind speed may affect to a flight schedule.

2.3 Analysis 3 Visibility

```
#Analysis 3 Visible
ggplot(data,aes(x =visib)) +
  geom_histogram() +
  labs(title = 'Histogram of Visible',x = 'Visible (Miles)')+
  facet_wrap(~origin)+
  theme_bw() + theme(panel.background = element_rect(fill = "darkseagreen1"), plot.background = element_rect(fill = "darkseagreen"))
```

Figure 6: Visibility (Code)

For the analysis 3, I am using ggplot function to plot the histogram. The first line of the code is means that the graph is form with the visibility from the dataset. The second line is the type of graph that I choose. The third line is I am naming the histogram with 'Histogram of Visible' and x-axis represent 'Visible (Miles)'. The fourth line is I am arranging and showing the histogram with two different airports and the code for last line is the design of the graph's background.

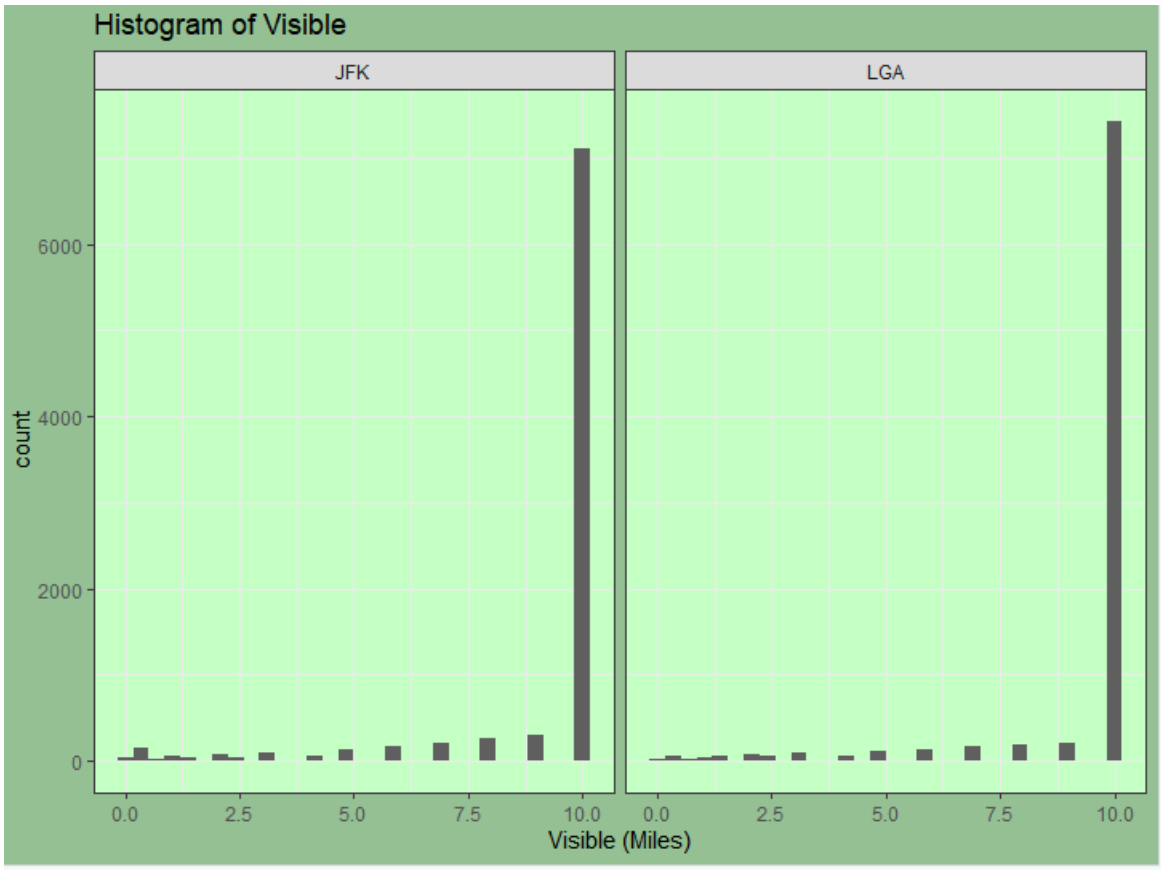

Figure 7: Visibility

The histogram above is showing the visibility of the two different airports in year 2013. Most of the time, these two airports are maintaining at 10 miles. 10-mile visibility implies that in the daytime "a prominent dark object against the sky at the horizon" and at night "a recognised, ideally unfocused, moderately intense light source" that is 10 miles away should

be able to see and recognise. Based on the result, most of the flight can be depart and landing as usual.

2.4 Analysis 4 Precipitate

```
#Analysis 4 Precipitate
ggplot(data, aes(x = precip)) +
  geom_histogram() +
  labs(title = 'Histogram of Precipitate',x = 'Precipitate (Inch)')+
  theme_bw() + theme(panel.background = element_rect(fill = "darkseagreen1"), plot.background = element_rect(fill = "darkseagreen"))
```

Figure 8: Precipitate (Code)

For the analysis 4, I am using ggplot function to plot the histogram. The first line of the code is means that the graph is form with the precipitation from the dataset. The second line is the type of graph that I choose. The third line is I am naming the histogram with 'Histogram of Precipitate' and x-axis represent 'Precipitate (Inch)'. The code for last line is the design of the graph's background.

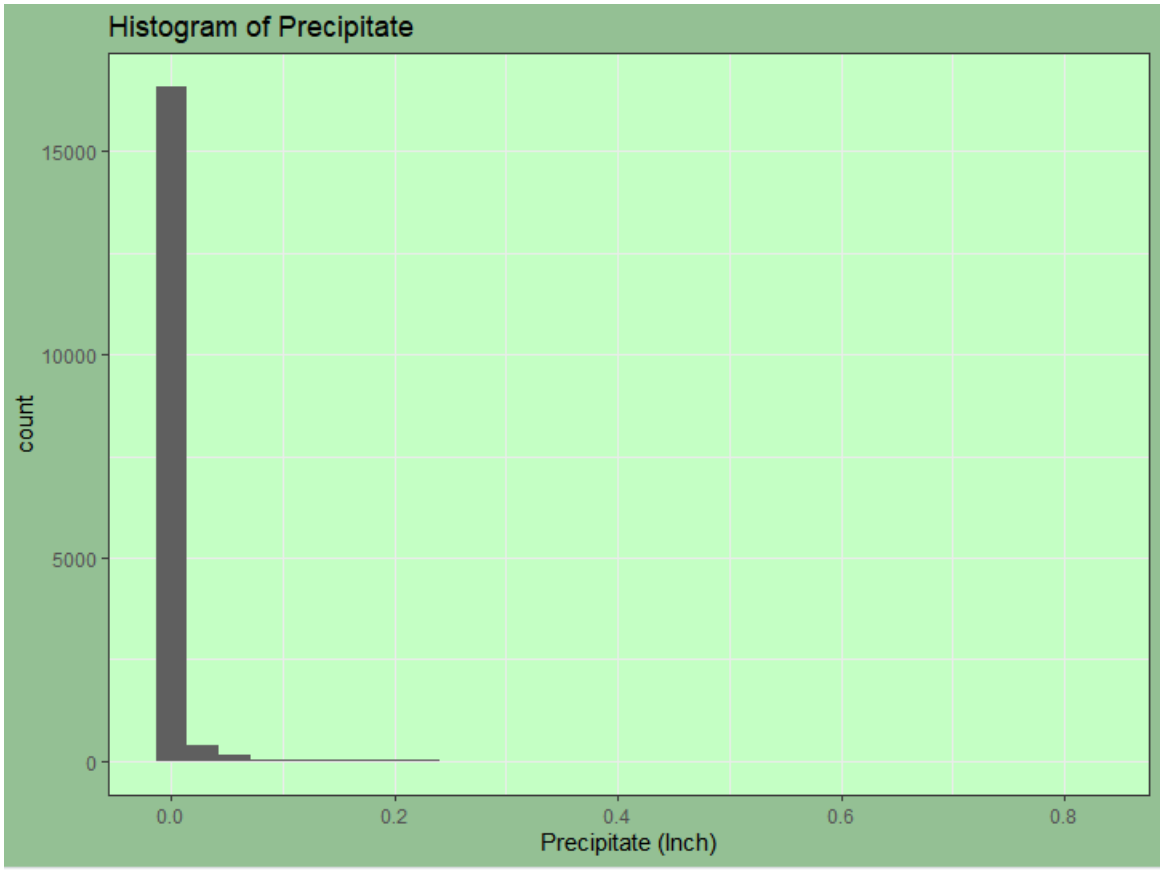

Figure 9: Precipitate

The histogram above is showing the precipitation of two airport in year 2013. The result is showing that the precipitate is in a range of 0.0 to 0.2 inch but mostly at 0.0 inch in year 2013. Water that release in the form of rain, snow and hail from clouds is call as precipitation. The rate of rainfall is typically characterised as light, moderate or heavy. According to the research website [CITATION WeatherShack \l 17417], it is considered as light rainfall is less than 0.10 inches of rain per hour. 0.1 to 0.3 inches of rain per hour considered as

moderate and heavy rain is more than 0.30 per hour. Based on the result, most of the weather of the two airports during year 2013 is sunny and it can attract the tourist who visit to New York.

2.5 Analysis 5 Relationship between Dew Point and Humid

```
#Analysis 5 Dew point against Humid
ggplot(data, aes(x = humid, y = dewp, color = origin)) +
  geom_point(alpha=0.4)+ geom_smooth(method = "lm") +
  labs(title = 'Relationship between Dew Point and Humid',x = 'Humid ', y = 'Dew Point (F)')+
  facet_wrap(~origin)+
  theme_bw() + theme(panel.background = element_rect(fill = "darkseagreen1"), plot.background = element_rect(fill = "darkseagreen"))
```

Figure 10: Dew Point and Humid (Code)

For the analysis 5, I am using ggplot function to plot the scatterplot diagram. The first line of the code is means that the graph is form with the humidity and dew point from the dataset. The X-axis is the humidity and Y-axis is the dew point. The colour is also using to show the two different airports. The second line is the type of graph that I choose. I am using 'alpha=0.4' to make the point in the graph more lightly. I am also using geom_smooth method to make the graph more visualize. The third line is I am naming the scatterplot diagram with 'Relationship between dew point and humidity', x-axis represent 'Humid' and y-axis represent 'Dew Point(F)'. The fourth line is I am arranging and showing the histogram with two different airports and the code for last line is the design of the graph's background.

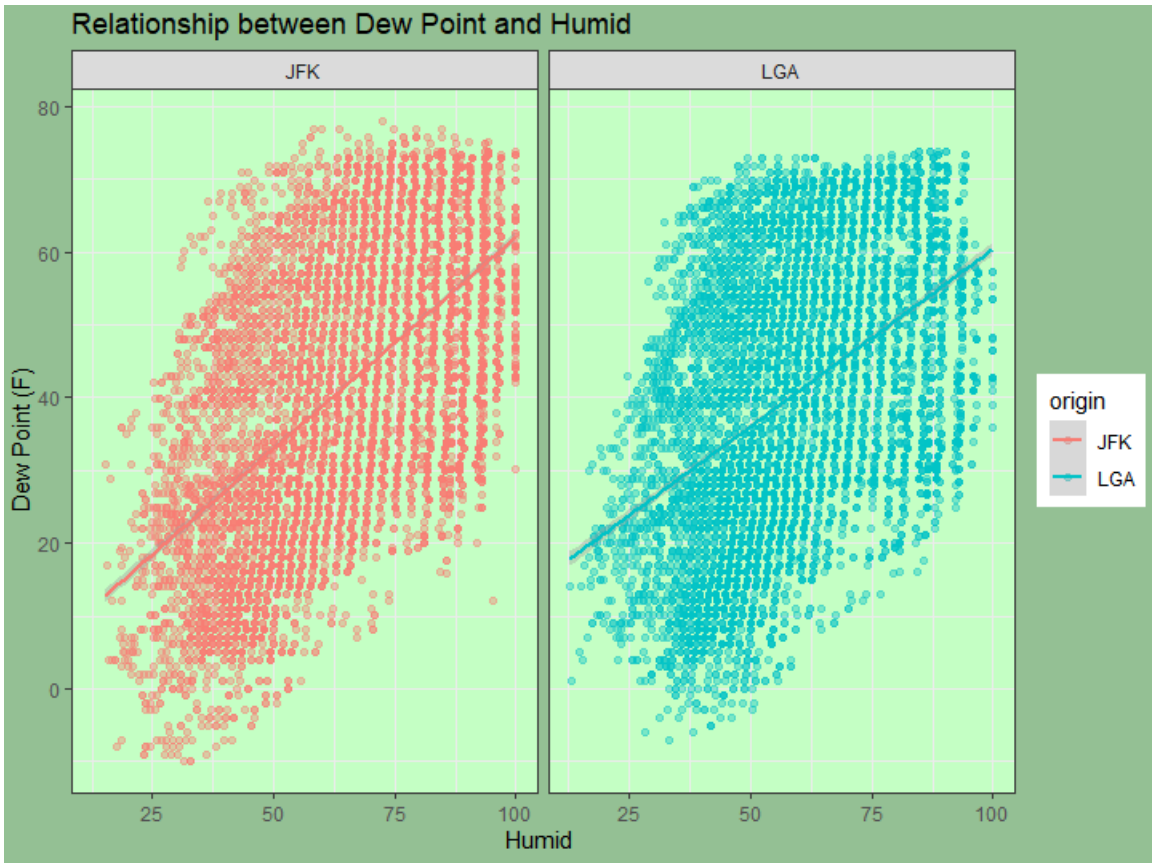

Figure 11: Dew Point and Humid

The scatterplot above is showing the relationship between dew point and relative humidity. Based on the graph, we can know that when the dew point is increasing, the relative humidity is also increasing. According to a research website [CITATION NationalWeatherService \l 17417], The dew point is the temperature at which the air needs to be cooled (at constant pressure) to reach a relative humidity (RH) of 100 percent. At this point, the air is unable to carry more water in the form of gas. If the air were to be cooled even more, water vapour would have to come out of the atmosphere in liquid form, usually as fog or precipitation. The higher the point of dew rises, the greater the amount of air moisture. This directly affects how outside it can feel comfortable." Relative humidity can be deceptive several times. For instance, a temperature of 30 and a dew point of 30 will give you 100 percent relative humidity, but a temperature of 80 and a dew point of 60 will produce 50 percent relative humidity. On an 80-degree day with 50 percent relative humidity, it will feel much more 'humid' than on a 30-degree day with 100 percent relative humidity. This is because the dew point is higher. As a result, tourist may not have a outdoor activity due to the muggier situation causes by dew point.

2.6 Analysis 6 Relationship between Wind Gust and Wind Speed

```
#Analysis 6 Wind Gust against Wind Speed
data%>%summarise (wind_speed = max(wind_speed,na.rm = TRUE))
data%>%summarise (wind_speed = min(wind_speed,na.rm = TRUE))
data%>%summarise (wind_speed = mean(wind_speed,na.rm = TRUE))

ggplot(data, aes(x = wind_gust, y = wind_speed)) +
  geom_point(alpha = 0.4)+ geom_hline(yintercept = 42.57886)+
  geom_hline(yintercept = 11.04554)+
  geom_hline(yintercept = 0)+
  labs(title = 'Relationship between Wind Gust and Wind Speed',x = 'Wind Gust (MPH)', y = 'Wind Speed (MPH)')+
  theme_bw() + theme(panel.background = element_rect(fill = "darkseagreen1"), plot.background = element_rect(fill = "darkseagreen"))
```

```
> data%>%summarise (wind_speed = max(wind_speed,na.rm = TRUE))
  wind_speed
1   42.57886
> data%>%summarise (wind_speed = min(wind_speed,na.rm = TRUE))
  wind_speed
1          0
> data%>%summarise (wind_speed = mean(wind_speed,na.rm = TRUE))
  wind_speed
1   11.04554
```

Figure 12: Wind Gust and Wind Speed (Code)

For the analysis 6, I am using ggplot function to plot the scatterplot diagram. First of all, I am summarising the mean, maximum, and minimum value of wind speed to find out the value from the dataset. The fourth line of the code is means that the graph is form with the wind speed and wind gust from the dataset. The X-axis is the wind gust and Y-axis is the wind speed. The fifth line is the type of graph that I choose. I am using ‘alpha=0.4’ to make the point in the graph more lightly. I am also using geom_hline method to make the graph more visualize. The eighth line is I am naming the scatterplot diagram with ‘Relationship between Wind Gust and Wind Speed’, x-axis represent ‘Wind Gust (MPH)’ and y-axis represent ‘Wind Speed (MPH)’ and the code for last line is the design of the graph’s background.

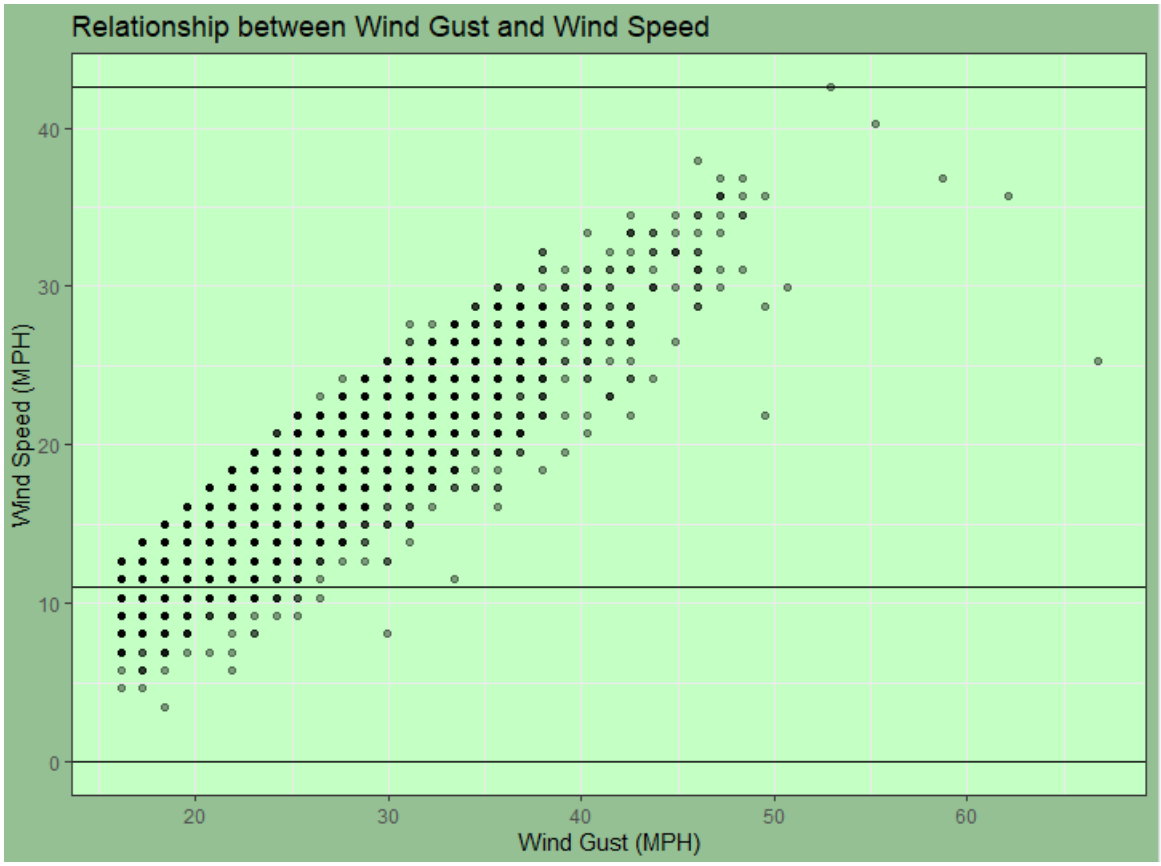

Figure 13: Wind Gust and Wind Speed

The scatterplot diagram above is showing the relationship between wind gust (MPH) and wind speed (MPH) of two different airport in year 2013. From the smooth line showing in the scatterplot, we can know that the higher wind speed, the stronger wind gust. A brief rise in wind speed is a wind gust. Basically, gusts occur while the wind is blowing. As the wind speed increases, gusts are more apparent. This reason is if the wind speed increases, the force applied by the wind will also increases quickly. Flight delay may be occurred is the speed of the wind gust is too strong because it will cause some accident.

2.7 Analysis 7 Relationship between Pressure and Temperature

```
#Analysis 7 Pressure against Temperature
ggplot(data, aes(x = temp, y = pressure, color = origin, na.rm = TRUE)) +
  geom_point(alpha = 0.4)+ geom_smooth(method = "lm") +
  labs(title = 'Relationship between Pressure and Temperature',x = 'Temperature ', y = 'Pressure (Milibars)')+
  theme_bw() + theme(panel.background = element_rect(fill = "darkseagreen1"), plot.background = element_rect(fill = "darkseagreen"))
```

Figure 14: Pressure and Temperature (Code)

For the analysis 7, I am using ggplot function to plot the scatterplot diagram. The first line of the code is means that the graph is form with the temperature and pressure from the dataset. The X-axis is the temperature and Y-axis is the pressure. The colour is also using to show the two different airports. Remove not available function also including after fetch the data from dataset. The second line is the type of graph that I choose. I am using 'alpha=0.4' to make the point in the graph more lightly. I am also using geom_smooth method to make the graph more visualize. The third line is I am naming the scatterplot diagram with 'Relationship between Pressure and Temperature', x-axis represent 'Temperature' and y-axis represent 'Pressure (Millibars)' and the code for last line is the design of the graph's background.

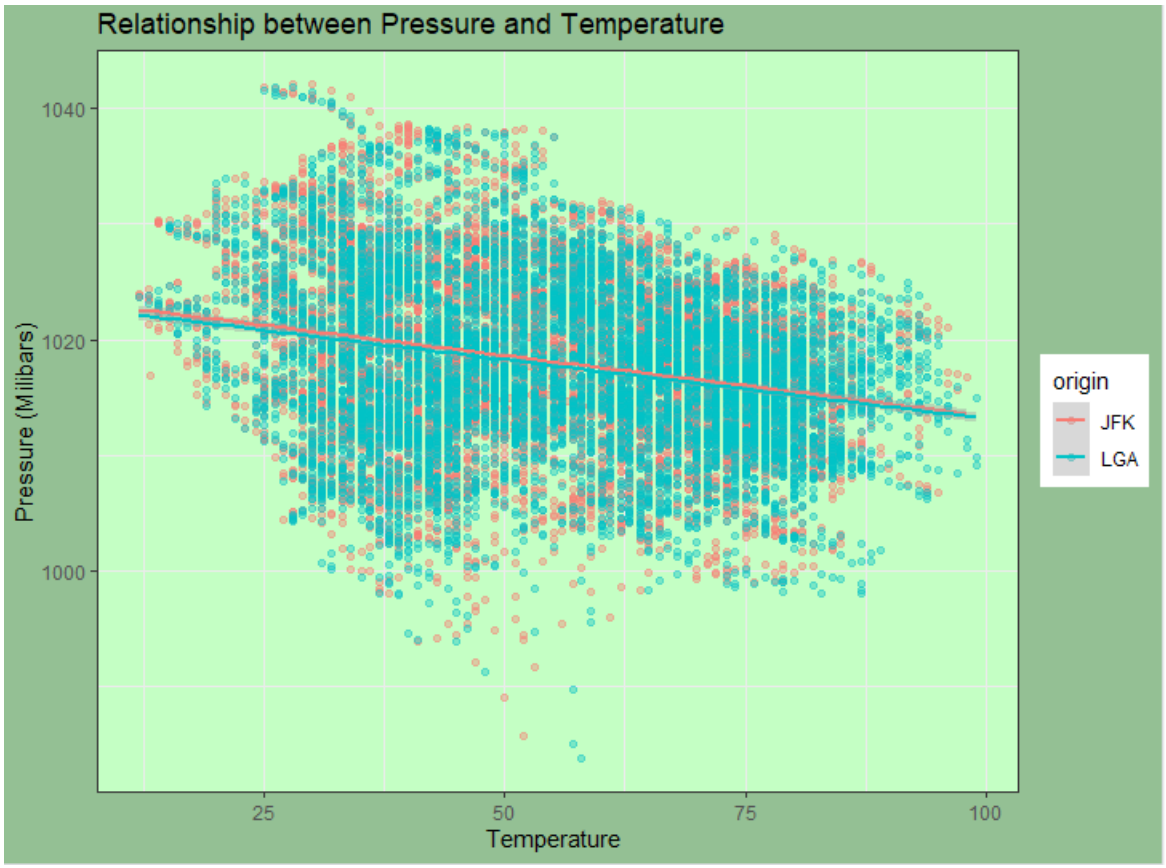

Figure 15: Pressure and Temperature

The scatterplot above is showing the relationship between pressure and temperature. Based on the scatterplot diagram above, we can see that when the temperature is cold, the pressure will be higher. Areas of high air pressure can be produced by extremely cold temperatures because cold air has greater density, and the concentration of molecules can increase the air pressure. A high-pressure system is called area of higher pressure, H, and typically has a denser air mass where the air temperature is cold. According to the researcher [CITATION CharlieandChris \l 17417] , a system of low pressure, or "low," is an environment in which the ambient pressure is lower than that of the area around it. High winds, warm weather, and atmospheric lifting are commonly correlated with highs. Because of this, clouds, snow, and other poor weather such as tropical storms and cyclones usually generate lows. Conversely, an area where the air pressure is greater than that of the surrounding area is a high-pressure system, or "high." Since more air fills the room left from the low pressure rises here. Subsidence often evaporates much of the water vapour in the atmosphere, so high pressure systems are typically connected to clear skies and calm weather. Therefore, relative to a low-pressure situation, an aircraft is more suited for flying in a high-pressure situation.

2.8 Analysis 8 Relationship between Precipitate and Humid

```
#Analysis 8 Precipitate against Humid
ggplot(data, aes(x = humid, y = precip, color = origin)) +
  geom_point(alpha = 0.4) + geom_smooth(method = "lm") +
  labs(title = 'Relationship between Precipitate and Humid',x = 'Humid ', y = 'Precipitate (Inch)')+
  theme_bw() + theme(panel.background = element_rect(fill = "darkseagreen1"), plot.background = element_rect(fill = "darkseagreen"))
```

Figure 16: Precipitate and Humidity (Code)

For the analysis 8, I am using ggplot function to plot the scatterplot diagram. The first line of the code is means that the graph is form with the humidity and precipitate from the dataset. The X-axis is the humidity and Y-axis is the precipitate. The colour is also using to show the two different airports. The second line is the type of graph that I choose. I am using 'alpha=0.4' to make the point in the graph more lightly. I am also using geom_smooth method to make the graph more visualize. The third line is I am naming the scatterplot diagram with 'Relationship between Precipitate and Humidity', x-axis represent 'Humidity' and y-axis represent 'Precipitate (Inch)' and the code for last line is the design of the graph's background.

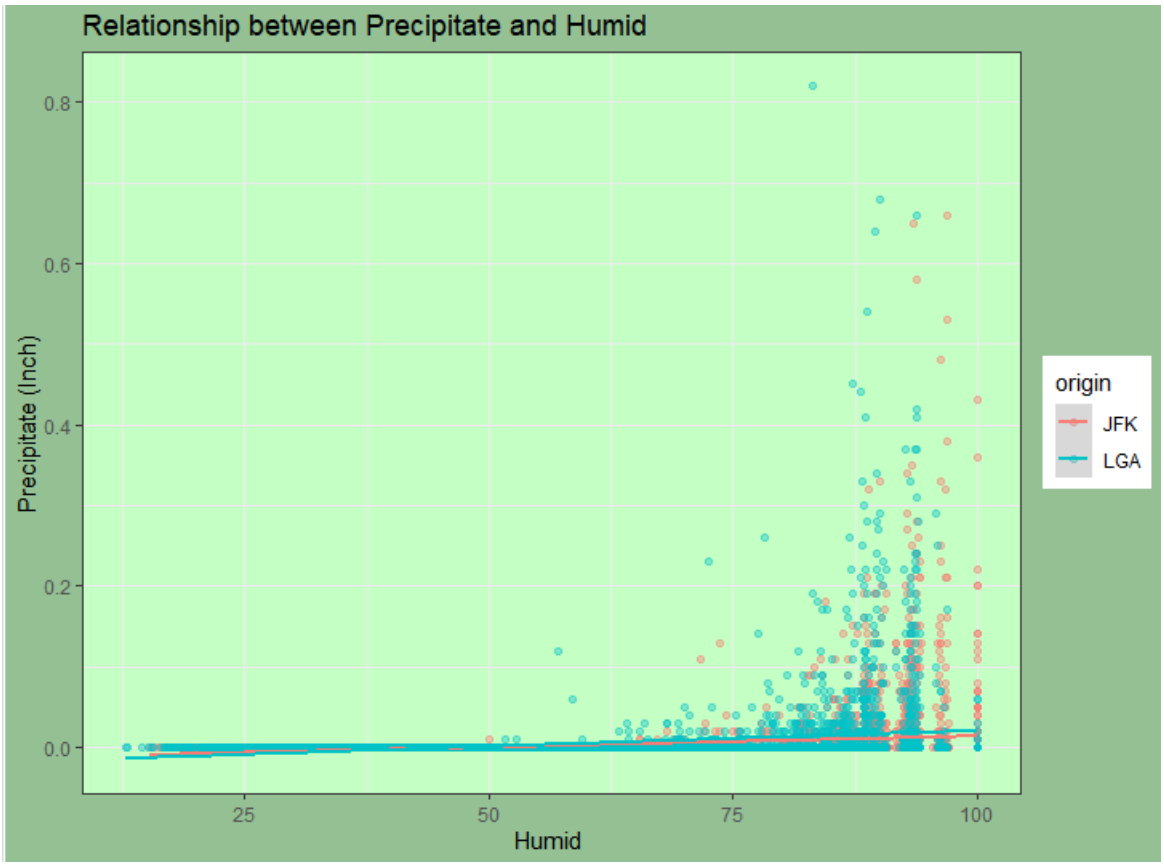

Figure 17: Precipitate and Humidity

The scatterplot diagram above is showing the relationship between precipitate and humidity. Based on the diagram above, we can see that when the humidity is increasing, the precipitate will be increase as well. As we know that the precipitation can causes raining, the humidity is 100% during raining, which is why the clouds cannot longer to carry water. During raining day, the relative humidity will increase because of the evaporation. Water vapour will not completely saturate the air where the rain is falling. The more it rains, however the higher the humidity will rise due to the continual drainage of the water by air. As the result, the flight may be delay or cancel if the precipitate and relative humidity goes high that can cause heavy rainfall.

2.9 Analysis 9 Relationship between Dew Point and Temperature

```
#Analysis 9 Dew Point against Temperature
ggplot(data, aes(x = temp, y = dewp, color = origin)) +
  geom_smooth(method = "lm") +
  labs(title = 'Relationship between Dew Point and Temperature',x = 'Temperature ', y = 'Dew Point (F)')+
  theme_bw() + theme(panel.background = element_rect(fill = "darkseagreen1"), plot.background = element_rect(fill = "darkseagreen"))
```

Figure 18: Dew Point and Temperature (Code)

For the analysis 9, I am using ggplot function to plot the scatterplot diagram. The first line of the code is means that the graph is form with the temperature and dew point from the dataset. The X-axis is the temperature and Y-axis is the dew point. The colour is also using to show the two different airports. The second line is the type of graph that I choose. I am using geom_smooth method to make the graph more visualize. The third line is I am naming the scatterplot diagram with 'Relationship between Dew Point and Temperature', x-axis represent 'Temperature' and y-axis represent 'Dew Point (F)' and the code for last line is the design of the graph's background.

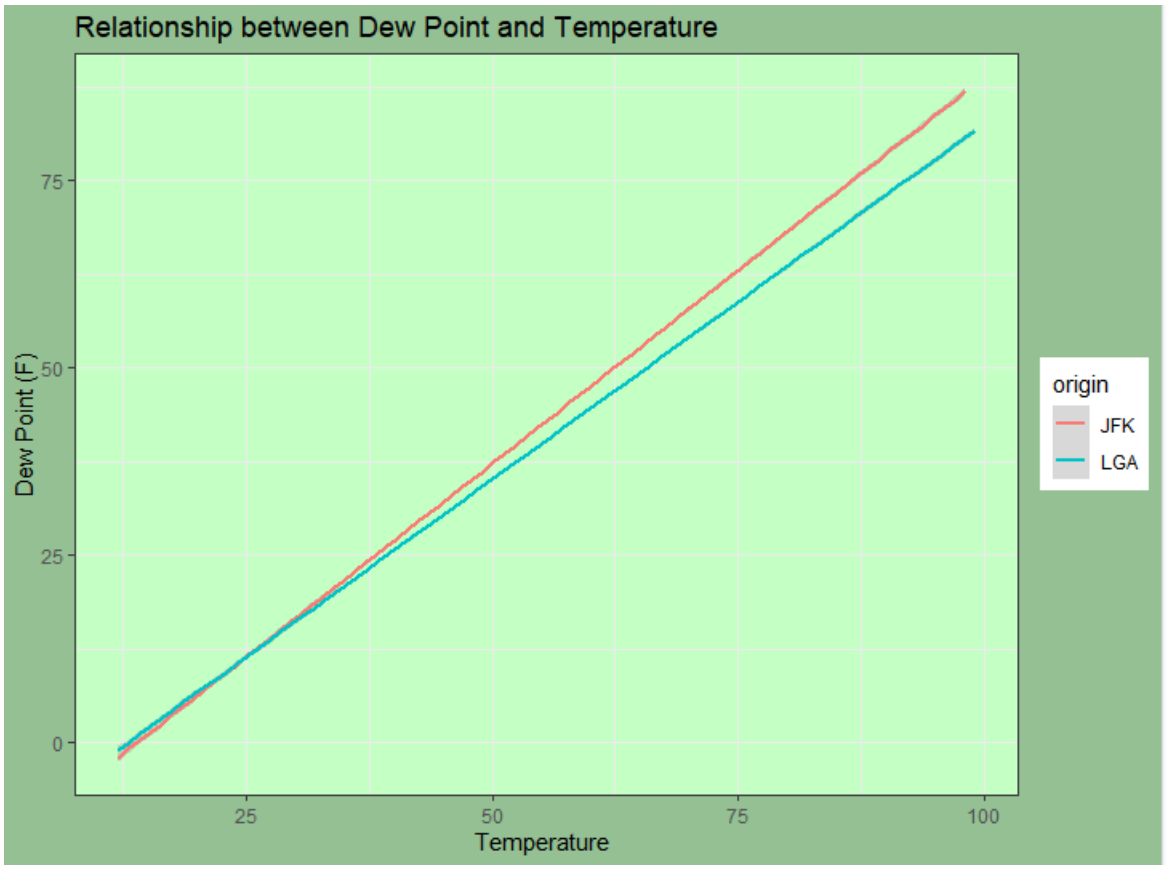

Figure 19: Dew Point and Temperature

The diagram above is showing the relationship between dew point and temperature. Based on the graph above, we can see that when the temperature is increasing, the dew point will also keep increasing. According to a website[CITATION WW2010 \y \l 17417], dew points

reflect the quantity of moisture in the air. The higher the dew points, the higher the moisture content of the air at a given temperature. The temperature of the dew point is known as the temperature to which the air will have to cool in order to reach saturation (at constant pressure and constant water vapour content). A state of saturation occurs when the air maintains the maximum amount of water vapour available at the current temperature and pressure. The air is said to be saturated when the dew point temperature and air temperature are equal. The temperature of the dew point is never bigger than the air temperature. Therefore, moisture must be extracted from the air once the air cools and this is done by condensation. The effect of this process is the creation of small droplets of water that can contribute to the production of fog, ice, clouds, or even precipitation. Therefore, the visibility of the airplane may be affected due to the dew point and temperature.

2.10 Analysis 10 Relationship between Precipitate and Month

```
#Analysis  10 Analysis between Precipitate and Month
data %>%
  filter(data$precip>0) %>%
  ggplot( aes(x = factor(month), y = precip)) +
  geom_boxplot() +
  labs(title = 'Relationship between Precipitate and Month',x = 'Month', y = 'Precipitate (Inch)')+
  theme_bw() + theme(panel.background = element_rect(fill = "darkseagreen1"), plot.background = element_rect(fill = "darkseagreen"))
```

Figure 20: Precipitate and Month (Code)

For the analysis 10, I am using ggplot function to plot the boxplot diagram. The first line and second line of the code is means that I am filter out the data which is precipitation bigger than 0 from dataset. The X-axis is the month and Y-axis is the precipitation. The fourth line is the type of graph that I choose. The fifth line is I am naming the boxplot diagram with 'Relationship between Precipitation and Month', x-axis represent 'Month' and y-axis represent 'Precipitate (Inch)' and the code for last line is the design of the graph's background.

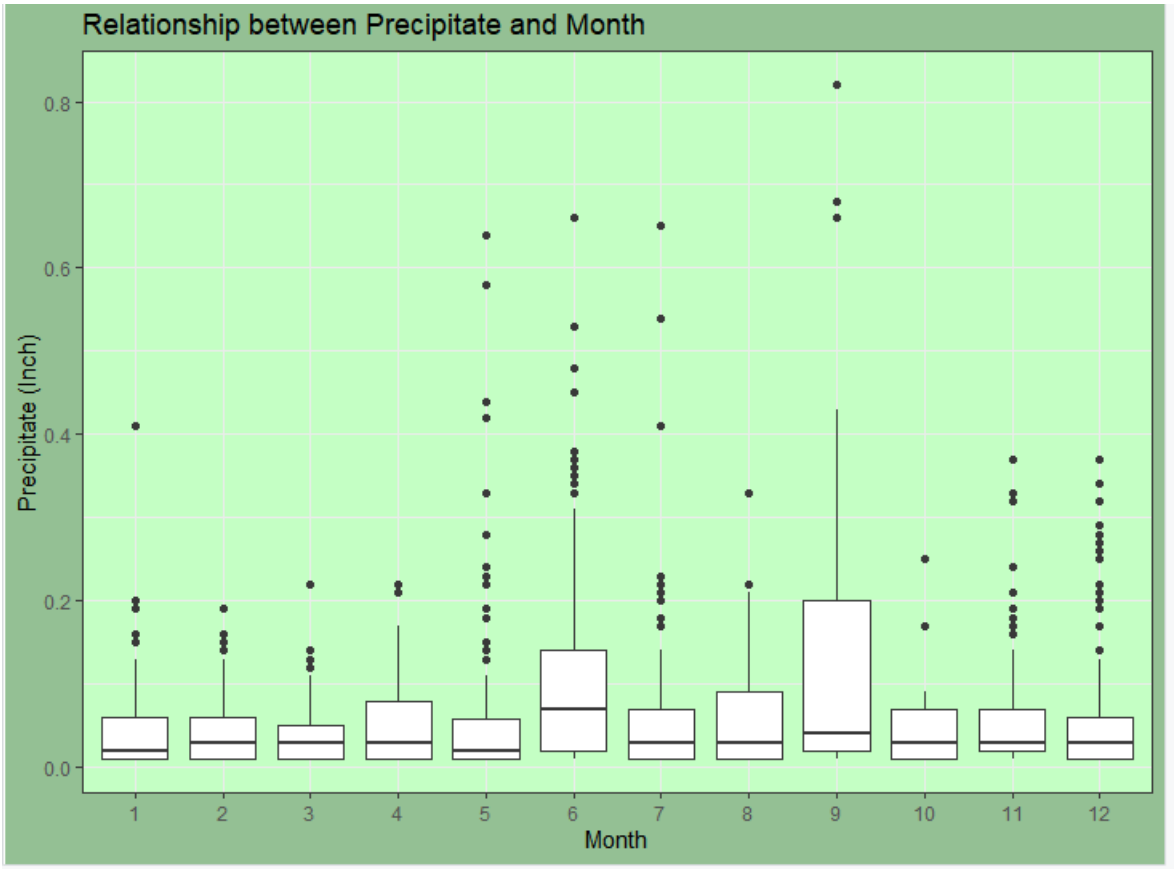

Figure 21: Precipitate and Month

The graph above is showing the relationship between month and precipitation. People are not encouraged to conduct an activity if the precipitation is serious. The reason is there may also be a detrimental effect on human activities, business and industry, agriculture, and the climate

from too much precipitation. For example, too much rain or snowmelt (water from melting snow) can lead to flooding at one time. In floodwaters, living organisms will drown, including crops. It is possible to wash away houses, businesses, even property.

2.11 Analysis 11 Relationship between Precipitate and Visibility

```
#Analysis 11 Analysis between Precipitate and visible
data %>%
  filter(data$precip>0) %>%
  ggplot(aes(x = factor(visib), y = precip)) +
  geom_boxplot() +
  labs(title = 'Relationship between Precipitate and Visible',x = 'Visible (Miles)', y = 'Precipitate (Inch)')+
  theme_bw() + theme(panel.background = element_rect(fill = "darkseagreen1"), plot.background = element_rect(fill = "darkseagreen"))
```

Figure 22: Precipitate and Visibility (Code)

For the analysis 11, I am using ggplot function to plot the boxplot diagram. The first line and second line of the code is means that I am filter out the data which is precipitation bigger than 0 from dataset. The X-axis is the visibility and Y-axis is the precipitation. The fourth line is the type of graph that I choose. The fifth line is I am naming the boxplot diagram with 'Relationship between Precipitation and Visible', x-axis represent 'Visible (Miles)' and y-axis represent 'Precipitate (Inch)' and the code for last line is the design of the graph's background.

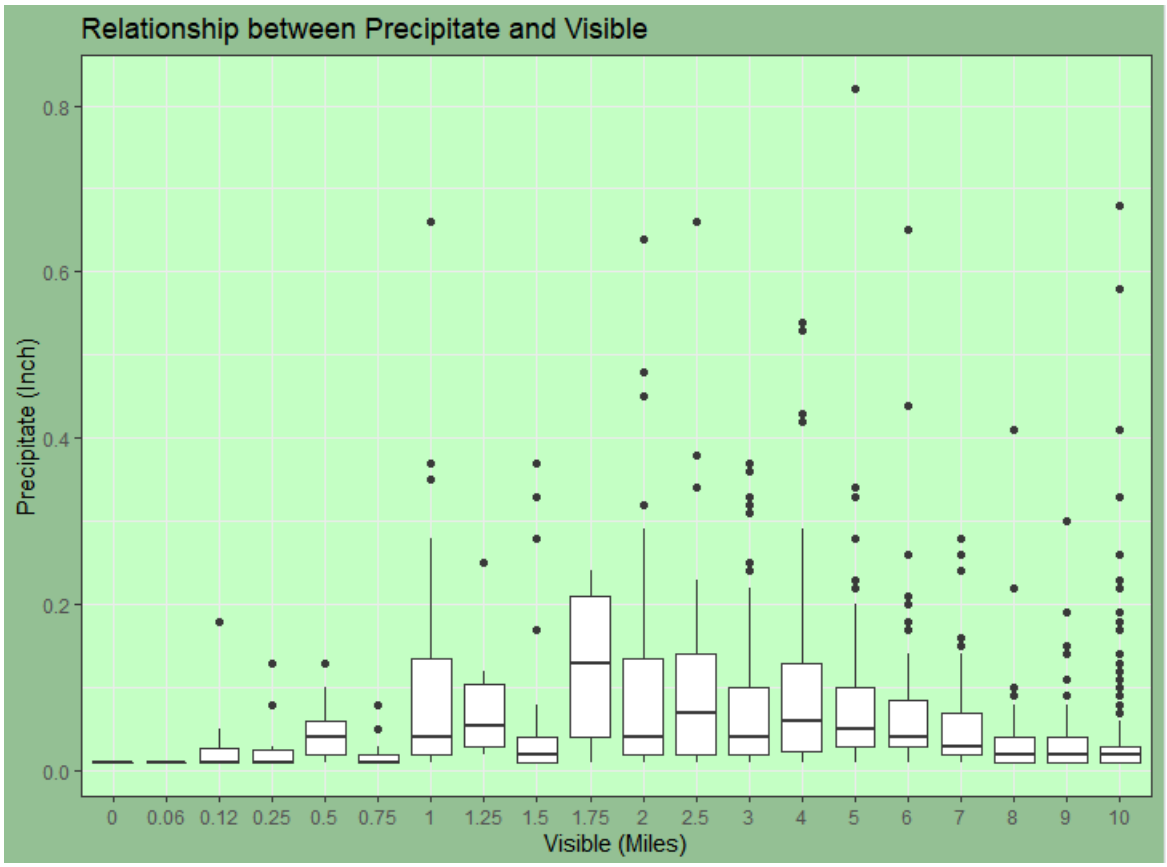

Figure 23: Precipitate and Visibility

The graph above is showing the relationship between precipitation and visibility. In all phases of flight, visibility is the most important factor especially when the flight needs to take-off or landing. However, the visibility will be reduced because of the reason precipitation

depending on intensity and droplet size. The light drizzle and snow that cause by precipitation can impede operations under the Visual Flight Rule (VFR). However, heavy rain, snow and thunderstorm activity will decrease visibility significantly. As a result, lowered visibility can result in loss of aircraft control when operating without instrument flight rule (IFR) certification.

2.12 Analysis 12 Relationship between Month and Wind Direction

```
#Analysis 12 Analysis between Month and Wind Direction
data %>%
  filter(data$wind_dir>130 & wind_dir<320) %>%
  ggplot( aes(x = factor(month), y = wind_dir)) +
  geom_boxplot() +
  labs(title = 'Relationship between Month and Wind Direction',x = 'Month', y = 'Wind Direction')+
  theme_bw() + theme(panel.background = element_rect(fill = "darkseagreen1"), plot.background = element_rect(fill = "darkseagreen"))
```

Figure 24: Month and Wind Direction (Code)

For the analysis 12, I am using ggplot function to plot the boxplot diagram. The first line and second line of the code is means that I am filter out the data which is wind direction bigger than 130 but smaller than 320 from dataset. The X-axis is the month and Y-axis is the wind direction. The fourth line is the type of graph that I choose. The fifth line is I am naming the boxplot diagram with 'Relationship between Month and Wind Direction', x-axis represent 'Month' and y-axis represent 'Wind Direction' and the code for last line is the design of the graph's background.

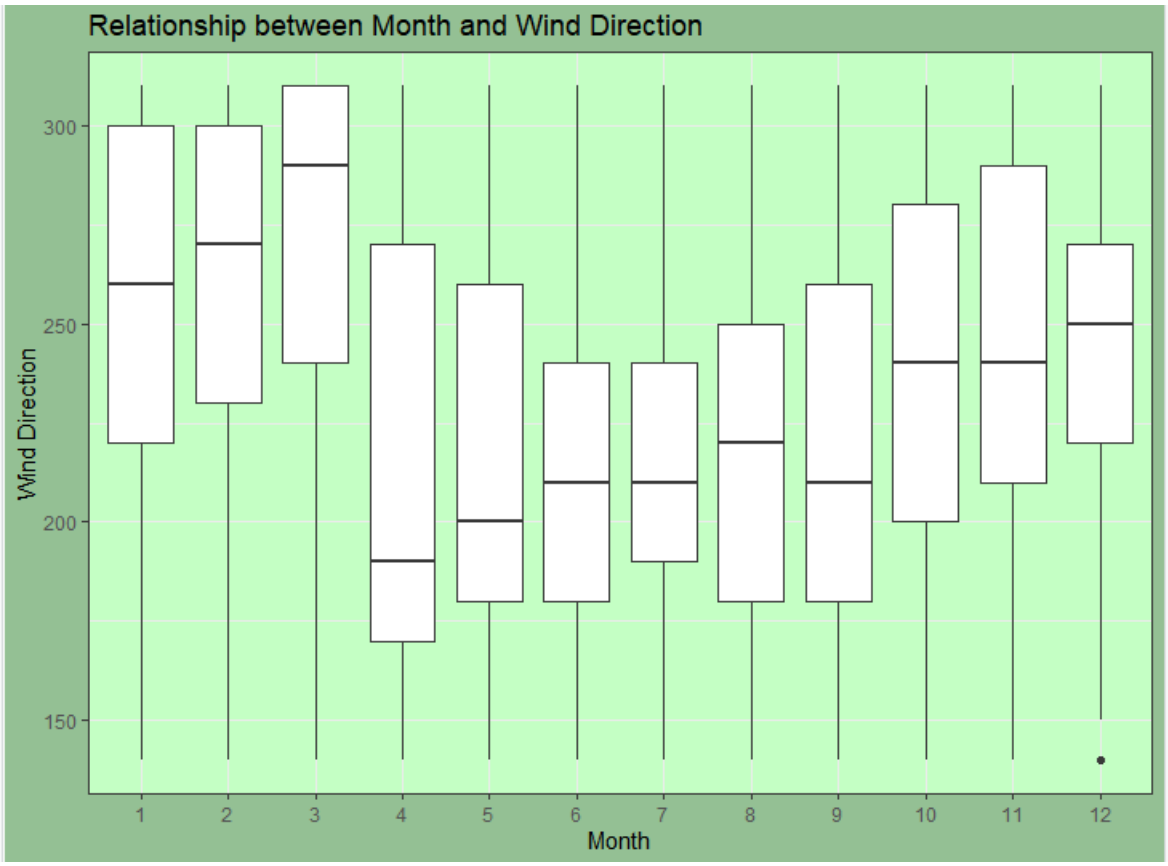

Figure 25: Month and Wind Direction

The graph above is showing the wind direction at LGA airport and JFK airport along 12 months. Wind direction can affect the flight directly. For example, headwind is wind blowing

towards the aircraft. As it boosts the lift, pilots tend to land and take off in the headwind. A lower ground velocity and a shorter run are required for the aircraft to become airborne in the headwind. Landing in the wind has the same advantages such as uses less runway, and touchdown ground speed is lower. Crosswinds and tailwinds are more complicated. Depending on the aircraft, it has maximum limits for the airport and the conditions on the runway. The aircraft will not attempt take-off or landing if winds surpass those limits. Therefore, the flight may be delay and affect airport or passenger’s schedule.

2.13 Analysis 13 Analysis between Temperature and Month

```
#Analysis 13 Analysis between Temperature and Month
data %>%
  filter(data$temp>50 & temp<86) %>%
  ggplot(aes(x = month, y = temp)) +
  geom_point() +
  labs(title = 'Relationship between Month and Temperature',x = 'Month', y = 'Temperature')+
  theme_bw() + theme(panel.background = element_rect(fill = "darkseagreen1"), plot.background = element_rect(fill = "darkseagreen"))
```

Figure 26: Month and Temperature (Code)

For the analysis 13, I am using ggplot function to plot the boxplot diagram. The first line and second line of the code is means that I am filter out the data which is temperature bigger than 50 but lower than 86 from dataset. The X-axis is the month and Y-axis is the temperature. The fourth line is the type of graph that I choose. The fifth line is I am naming the boxplot diagram with 'Relationship between Month and Temperature', x-axis represent 'Month' and y-axis represent 'Temperature' and the code for last line is the design of the graph's background.

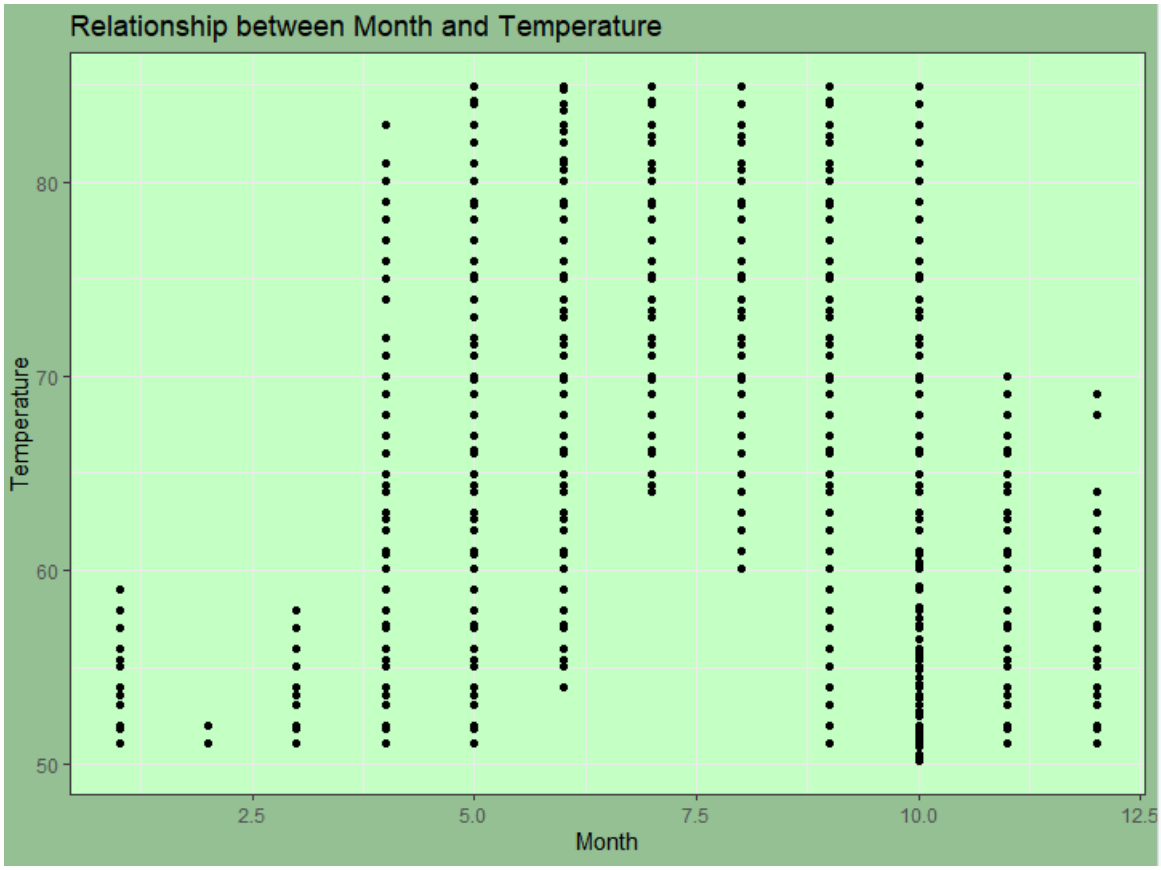

Figure 27: Month and Temperature

Based on my result, 50°F is equal to 10 degree Celsius and 86°F is equal to 30 degree Celsius. The graph above is showing the temperature of the month in the range of 50°F to 86°F. During this period, it is suitable to visit to the JFK and LGA airport which located at

New York City. The reason is the temperature between 10 degree Celsius to 30 degree Celsius is suitable to most of the person. It was not too cold or too hot. Most of the outdoor activity can be conducted and no need to worry about the weather.

2.14 Analysis 14 Analysis between Humidity and Temperature

```
#Analysis 14 Analysis between Humidity and temperature
ggplot(data, aes(x = humid, y = temp, color = origin, na.rm = TRUE)) +
  geom_point(alpha = 0.4)+ geom_smooth(method = "lm") +
  labs(title = 'Relationship between Humid and Temperature',x = 'Humid ', y = 'Temperature')+
  theme_bw() + theme(panel.background = element_rect(fill = "darkseagreen1"), plot.background = element_rect(fill = "darkseagreen"))
```

Figure 28: Humidity and Temperature (Code)

For the analysis 14, I am using ggplot function to plot the scatterplot diagram. The first line of the code is means that the graph is form with the humidity and temperature from the dataset. The X-axis is the humidity and Y-axis is the temperature. The colour is also using to show the two different airports. The second line is the type of graph that I choose. I am using 'alpha=0.4' to make the point in the graph more lightly. I am also using geom_smooth method to make the graph more visualize. The third line is I am naming the scatterplot diagram with 'Relationship between Humid and Temperature', x-axis represent 'Humid' and y-axis represent 'Temperature' and the code for last line is the design of the graph's background.

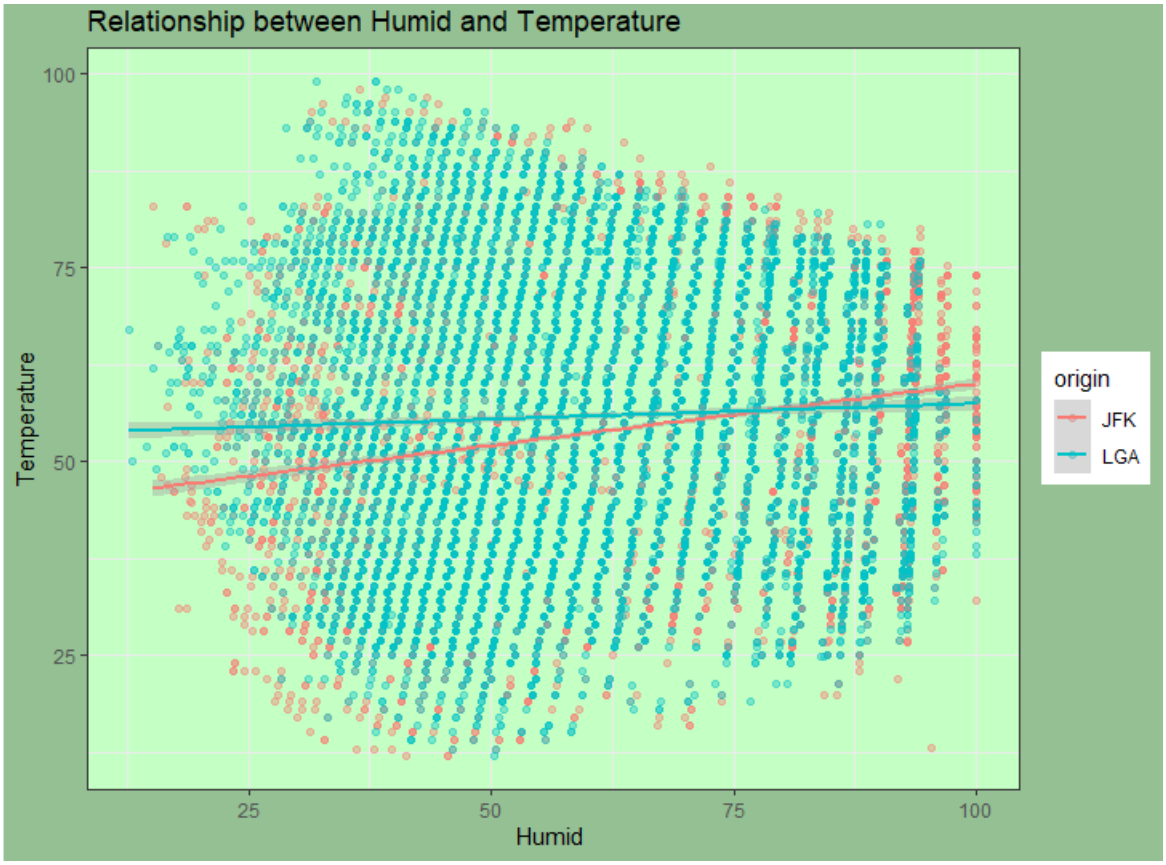

Figure 29: Humidity and Temperature

The scatterplot diagram above is showing the relationship between humidity and temperature. According to [CITATION JDianneDotson \l 17417], relative humidity is a proportion of the air's water vapour that varies as the temperature of the air changes For example, a completely saturated parcel of air does not carry any more water molecules at constant pressure, giving it a relative humidity of 100 percent. As the air temperature increases, the air can hold more water molecules and its relative humidity decreases. When temperatures drop, the relative humidity increases. When the dew point value is surpassed by the air temperature, high relative air humidity exists. Thus, temperature relates directly to the amount of moisture that the atmosphere can hold. Therefore, tourist who like summer can choose the time that high temperature with low relative humidity to travel.

3.0 Extra Feature 1

```
#Analysis 9 Dew Point against Temperature
ggplot(data, aes(x = temp, y = dewp, color = origin)) +
  geom_smooth(method = "lm") +
  labs(title = 'Relationship between Dew Point and Temperature',x = 'Temperature ', y = 'Dew Point (F)')+
  theme_bw() + theme(panel.background = element_rect(fill = "darkseagreen1"), plot.background = element_rect(fill = "darkseagreen"))
```

Figure 30: Geom_smooth (Code)

Based on the analysis 9, I am using geom_smooth as my extra feature to help me analyse the diagram. This method can help us analyse the trend of the graph clearly because it may be hard to view the point based on the points of the graph alone. There will be 2 regression line that showing the trend if we run this method.

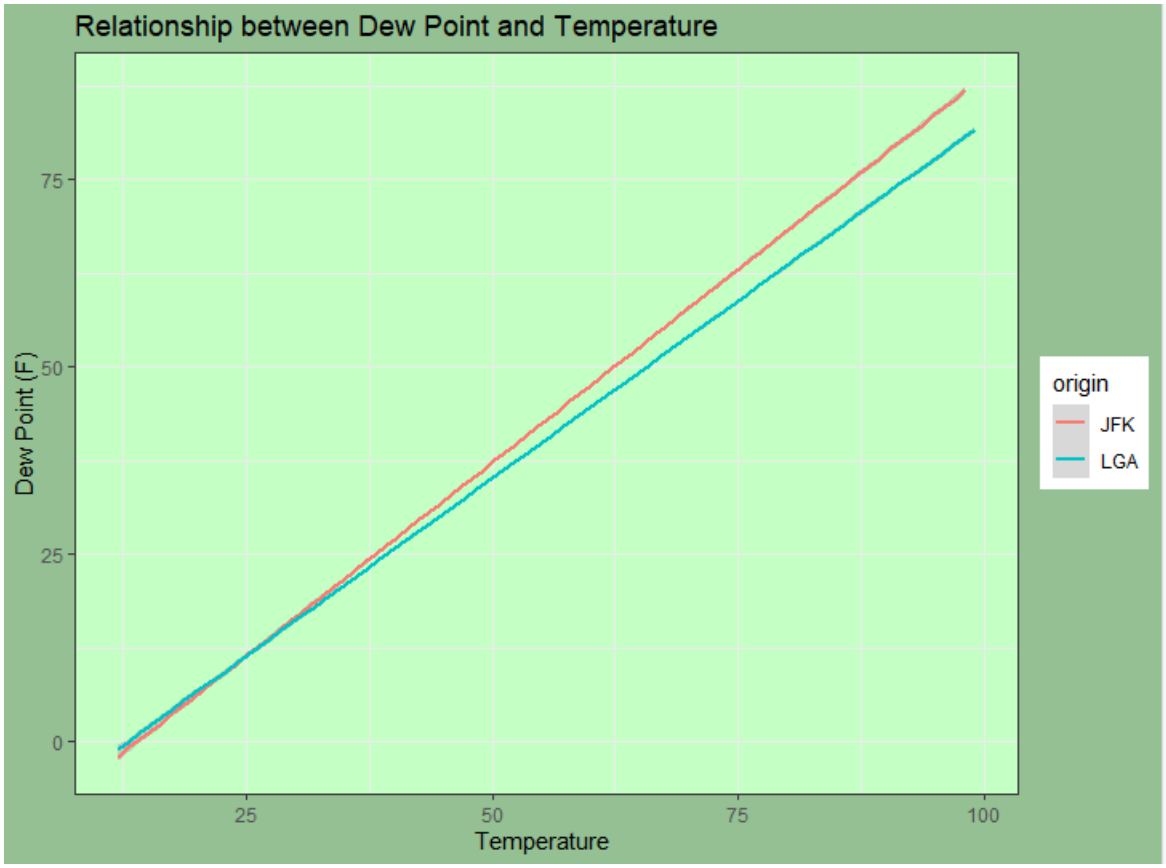

Figure 31: Geom_smooth

Based on the graph above, we can see that there are 2 regression line which represent two different airports are showing the relationship between dew point and temperature. After implementing this extra feature, we can clearly see that the dew point will increase when the temperature is increase. A going up trend is showing in the graph.

3.1 Extra Feature 2

```
#Analysis 6 Wind Gust against Wind Speed
data%>%summarise (wind_speed = max(wind_speed,na.rm = TRUE))
data%>%summarise (wind_speed = min(wind_speed,na.rm = TRUE))
data%>%summarise (wind_speed = mean(wind_speed,na.rm = TRUE))

ggplot(data, aes(x = wind_gust, y = wind_speed)) +
  geom_point(alpha = 0.4)+ geom_hline(yintercept = 42.57886)+
  geom_hline(yintercept = 11.04554)+
  geom_hline(yintercept = 0)+
  labs(title = 'Relationship between Wind Gust and Wind Speed',x = 'Wind Gust (MPH)', y = 'Wind Speed (MPH)')+
  theme_bw() + theme(panel.background = element_rect(fill = "darkseagreen1"), plot.background = element_rect(fill = "darkseagreen"))
```

```
> data%>%summarise (wind_speed = max(wind_speed,na.rm = TRUE))
  wind_speed
1   42.57886
> data%>%summarise (wind_speed = min(wind_speed,na.rm = TRUE))
  wind_speed
1          0
> data%>%summarise (wind_speed = mean(wind_speed,na.rm = TRUE))
  wind_speed
1   11.04554
```

Figure 32: Geom_hline (Code)

Based on the analysis 6, I am using geom_hline as my extra feature to help me analyse the diagram. There are 3 horizontal line in the graph above. The reason is I am implementing this extra feature three time. These three lines is representing maximum, minimum and the mean of the wind speed. We need to find out the three-value using summarise method before key in the value to the yintercept of hline. After finding out the values, we can implement it into our graph by key in the yintercept of hline.

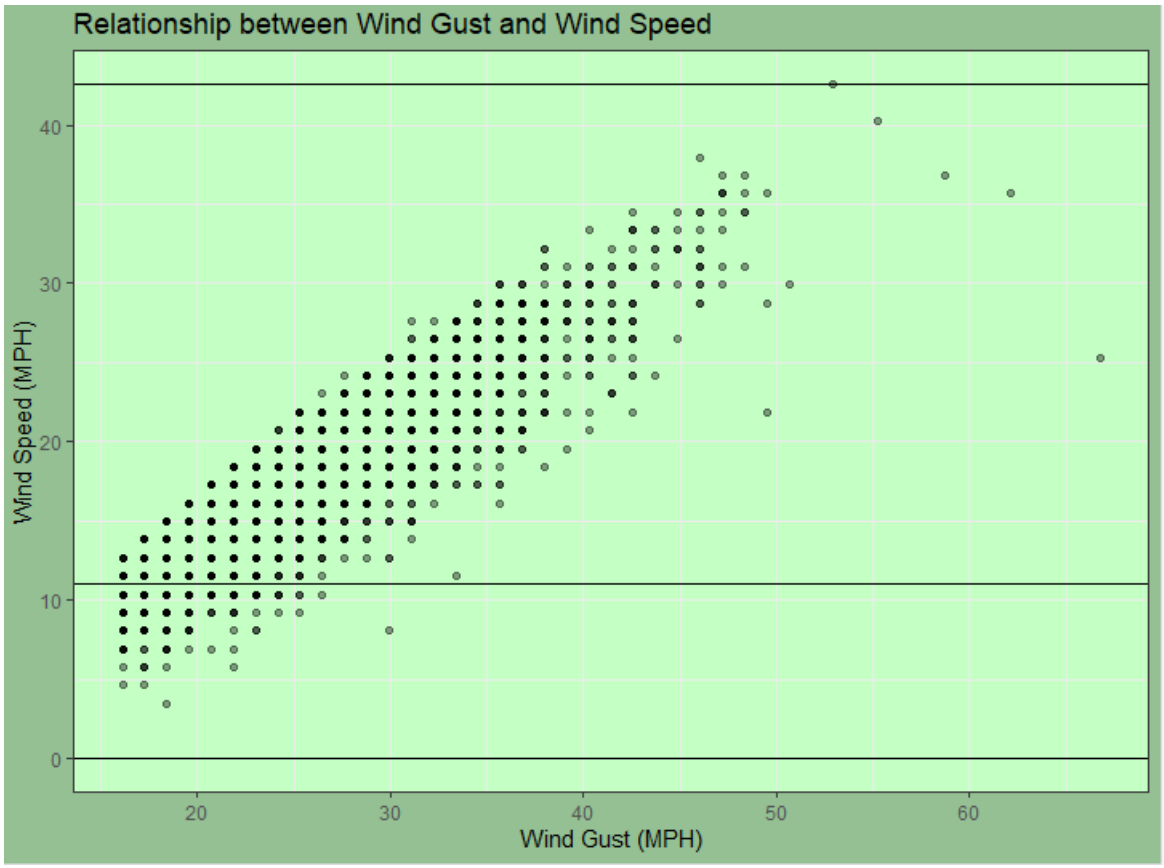

Figure 33: Geom_hline

Based on the graph above, we can see that there are three horizontal line which represent maximum, minimum and mean of the wind speed. The value of the horizontal line is very important to us to make an analysis. Throughout the value of this data, we can know that the faster the wind speed, the strongest the wind gust. By doing this, we can avoid the weather with strong wind gust to take-off or landing an aircraft so that the aircraft can be under control. Besides that, we can know the mean of the wind speed. An airport can use the mean of value to schedule a flight.

Conclusion

In the conclusion, it is very important to examine the dataset related to the hourly meteorological data for LaGuardia Airport (LGA) and John F. Kennedy International Airport (JFK). Through the dataset, a lot of analysis can be created. It aloud us making a strong decision by looking at the analysis. The reason is the study examines the dataset in depth and let the airport management and pilot can making its decision whether continue the flight as usual or delay the flight. Finally, as the wind speed, pressure, temperature, and everything else rises or decreases, the dataset is a very effective tool for humans to predict what will happen next. If each dataset is being well analysed, it will certainly safe a live from any unwanted accident.

References

1. Skyscanner US. 2020. *What Wind Speed Delays Flights? What You Need To Know | Skyscanner*. [online] Available at: <https://www.skyscanner.com/tips-and-inspiration/what-windspeed-delays-flights> [Accessed 21 November 2020].https://www.weathershack.com/static/ed-rain-measurement.html
2. Weather.gov. 2020. *Dew Point Vs Humidity*. [online] Available at: <https://www.weather.gov/arx/why_dewpoint_vs_humidity#:~:text=The%20dew%20point%20is%20the,(RH)%20of%20100%25.&text=For%20example%2C%20a%20temperature%20of,a%20relative%20humidity%20of%2050%25.> [Accessed 21 November 2020].https://www.theweatherprediction.com/habyhints3/963/#:~:text=A%20wind%20gust%20is%20a%20brief%20increase%20in%20wind%20speed.&text=The%20air%20closer%20to%20the,air%20is%20more%20smoothly%20flowing.
3. Sciencing. 2020. *How Does Temperature Affect Barometric Pressure?*. [online] Available at: <https://sciencing.com/temperature-affect-barometric-pressure-5013070.html> [Accessed 21 November 2020].https://www.gunsonpegs.com/articles/shooting-talk/how-air-pressure-affects-the-flight-of-game-birds
4. Smart Fog. 2020. *How Rain And Humidity Connected? | Smart Fog*. [online] Available at: <https://www.smartfog.com/how-rain-and-humidity-connected.html#:~:text=Connection%20of%20Rain%20and%20Humidity&text=When%20it%20rains%2C%20it%20will,humidity%20because%20of%20the%20evaporation.&text=On%20a%20larger%20scale%2C%20rain,relative%20humidity%20reduces%20through%20rain.> [Accessed 21 November 2020].
5. Ww2010.atmos.uiuc.edu. 2020. *Observed Dew Point Temperature:Indicates The Amount Of Moisture In The Air*. [online] Available at: <http://ww2010.atmos.uiuc.edu/(Gh)/guides/maps/sfcobs/dwp.rxml#:~:text=Dew%20points%20indicate%20the%20amount,air%20at%20a%20given%20temperature.&text=When%20the%20dew%20point%20temperature,GREATER%20than%20the%20air%20temperature.> [Accessed 21 November 2020].
6. Finavia.fi. 2020. *How Do Wind Conditions Affect Flight? | Finavia*. [online] Available at: <https://www.finavia.fi/en/newsroom/2017/how-do-wind-conditions-affect-

flight#:~:text=Headwind%20is%20wind%20blowing%20towards,the%20plane%20to%20become%20airborne.&text=If%20winds%20exceed%20those%20limits,not%20attempt%20takeoff%20or%20landing.> [Accessed 21 November 2020].

7. Sciencing. 2020. *How Temperature & Humidity Are Related*. [online] Available at: <https://sciencing.com/temperature-ampamp-humidity-related-7245642.html> [Accessed 21 November 2020].