INDIVIDUAL ASSIGNMENT

TECHNOLOGY PARK MALAYSIA

CT127-3-2-PFDA

PROGRAMMING FOR DATA ANALYSIS

UC2F2008CS

HAND OUT DATE: 27 JULY 2020

HAND IN DATE: 22 November 2020

WEIGHTAGE: 50%

Name	Hong Jun Khuan
TP Number	TP050852
Intake Code	UC2F2008CS
Assignment Topic	Hourly Weather Data
Hand Out Date	22 September 2020
Hand In Date	23 November 2020
Weightage	50%

Table of Contents

Introduction

Based on this project, analysis is form with a dataset given form weather station located at JFK and LGA. The dataset contains different type of data that have been collected based on the weather in an hour, a day, month, and year. This dataset can be useful as it helps to assists the airline gain information on the location of airport. Each analysis that have been provided below is based on the dataset that have been collected with additional of facts that helps to enhance the analysis.

Data Analysis

```
#import data
filepath ='C:/Users/user/Desktop/PFDA Individual Assignment/PFDA Individual Assignment/4. Hourly weather data.csv'
data = read.csv(filepath)
library(ggplot2)
library(dplyr)
library(RColorBrewer)
```

Figure 1 import data

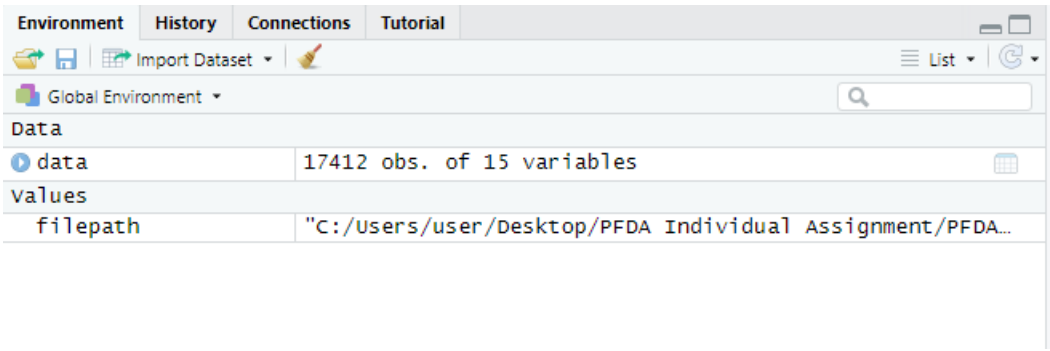

Figure 2 data stored

As for the first step to start to get different information from the data given. Importing data into the global environment to make sure the data from the csv file is stored. Follow by, loading the library(ggplot2) is to load for different graph to view the data. Load(**dplyr)** is a new package which provides a set of tools for efficiently manipulating datasets in R. While library (RcolorBrewer) is used for changing the color to make it more interesting in the graph view.

```
#summary of data set
summary(data)
```

Figure 3 summary of data

```
> #summary of data set
> 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 4 data of summary

The summary of data set is to view the summary that have been collected from the 4. Hourly weather data.csv. The data set will be shown once in a summary from the csv file.

Analysis example 1

```
#Analysis Example 1
ggplot(data = data, mapping = aes(x =visib, fill = origin)) +
  geom_histogram() +
  labs(title = 'Histogram of Visiblility',x = 'Visible (Miles)')+
  facet_wrap(~month) + theme_bw() + theme(plot.title = element_text(hjust = 0.5, face = "bold", colour = "black"),
                                          panel.background = element_rect(fill = "linen"), plot.background = element_rect(fill = "ivory3"))
data[which.max(data$visib),]
data[which.min(data$visib),]
```

Figure 5 snippet code 1

To start on the graph, ggplot is used while the "aes" is used to get the data from the dataset. Then, the geom_histogram() is used for indicating the type of graph. With the labs, title is shown to make sure the graph is filled with suitable title. The facet_wrap() is to specify the data that is used to measure the visibility. Theme is also used to make the background of table. Function of data is used to find the maximum and minimum of the month to make a better comparison from both airports.

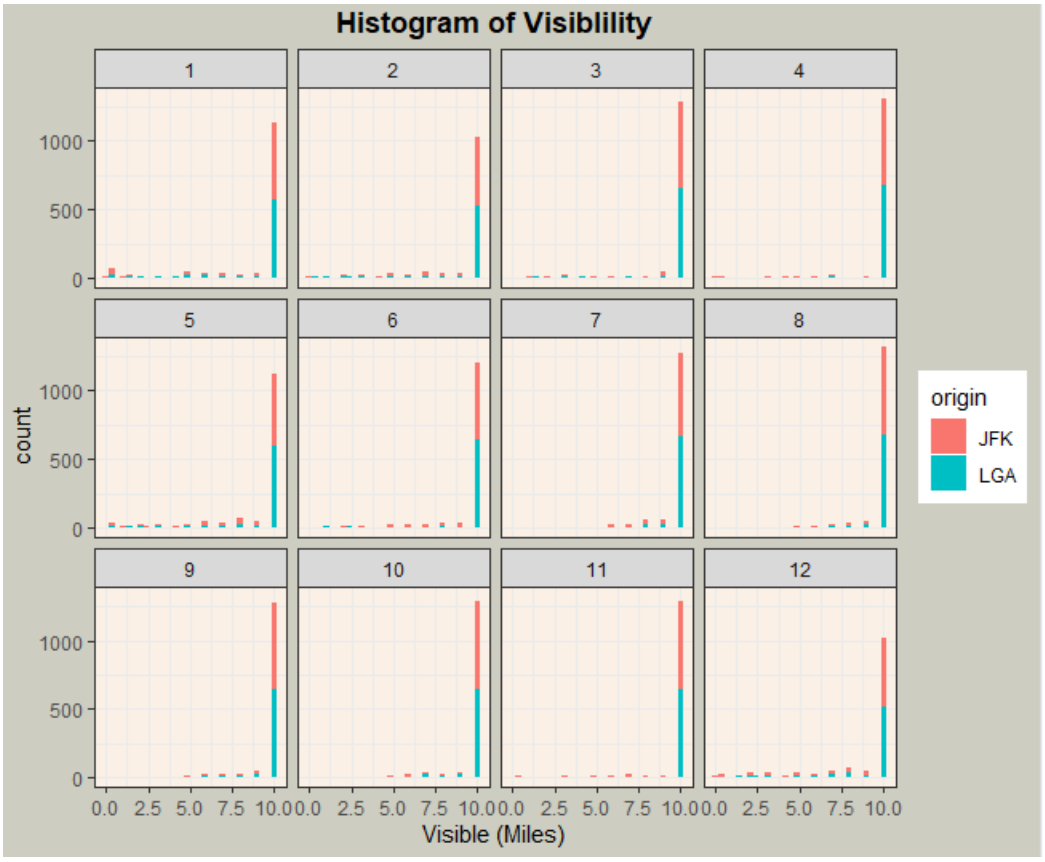

Figure 6 graph analysis 1

```
> data[which.max(data$visib),]
  origin year month day hour  temp  dewp humid wind_dir wind_speed wind_gust precip pressure visib        time_hour
1    JFK 2013     1   1    1 39.02 26.06 59.37      260   12.65858        NA      0   1012.6    10 01/01/2013 01:00
> data[which.min(data$visib),]
    origin year month day hour temp  dewp humid wind_dir wind_speed wind_gust precip pressure visib        time_hour
699    JFK 2013     1  30    4 46.4 46.04   100      170    8.05546        NA      0       NA     0 30/01/2013 04:00
```

Figure 7 max & min data

Based on the analysis example 1, this shows the visibility of the weather data based on each month. The data that have been shown the airport that have a best visibility is on the month of January 1st, it has a visible month of 10 Miles on that day. As a contrast, on the same month of day 30th the visibility on that day is on a count of 0 Miles. Therefore, on January 1st the aircraft of JFK can work as usual but on the day of January 30th the aircraft may delay due to bad weather condition. As an example, the minimum visibility requirement for airplane to land and fly is one mile forward.

Analysis example 2

```
#Analysis Example 2
ggplot(data = data, mapping = aes(x = precip, fill = origin)) +
  geom_histogram(bins = 30, col = "black") +
  labs(title = 'Histogram of Precipitate in month September',x = 'Precipitate (Inch)')+ facet_wrap(~day)+
   theme_bw() + theme(plot.title = element_text(hjust = 0.5, face = "bold", colour = "black"),
                                    panel.background = element_rect(fill = "linen"), plot.background = element_rect(fill = "ivory3"))
```

Figure 8 snippet code 2

To start on the graph, ggplot is used while the "aes" is used to get the data from the dataset. Then, the geom_histogram() is used for indicating the type of graph. With the labs, title is shown to make sure the graph is filled with suitable title. The facet_wrap() is to specify the data that is used, facet_wrap(~day) is used to determine the data of 'day' for precipitation. Theme is also applied to make the background of table.

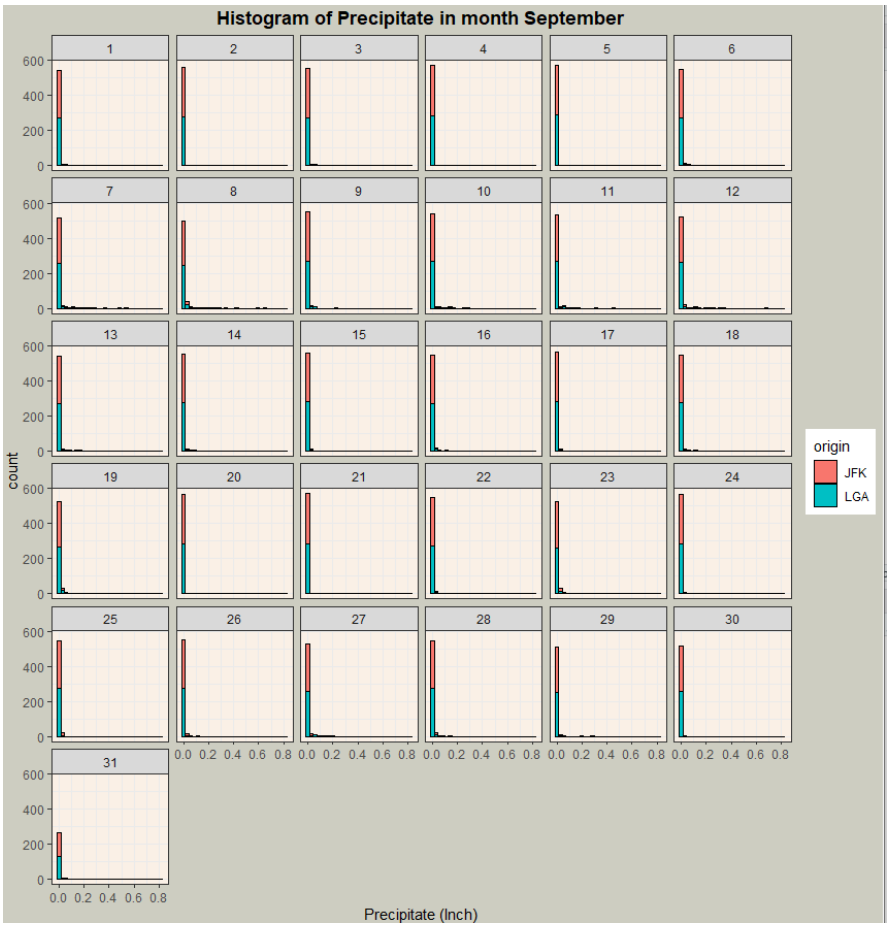

Figure 9 graph analysis 2

As for the analysis example 2, the histogram shows the precipitate of a certain month which is month September. (Climate - Atmospheric humidity and precipitation, 2020) It shows that the day of the month which have a higher precipitate or lower precipitate. Based on the histogram,

September 31 have a lower precipitation from both airport JFK and LGA weather station than the other dates. Hence this means that on the day of September 31, rainfall shortage could happen as the count drops below 400 inches.

Analysis example 3

```
#Analysis example 3
ggplot(data = data, mapping = aes(x = wind_speed, na.rm = TRUE, fill = origin )) +
  geom_histogram(bins = 30, col = "black") +
  labs(title = 'Histogram of Wind speed',x = 'Wind speed (MPH)') +
  facet_wrap(~origin)+ theme_bw() + theme(plot.title = element_text(hjust = 0.5, face = "bold", colour = "black"),
                                          panel.background = element_rect(fill = "linen"), plot.background = element_rect(fill = "ivory3"))
```

Figure 10 snippet code 3

To start on the graph, ggplot is used while the "aes" is used to get the data from the dataset. Then, the geom_histogram() is used for indicating the type of graph. With the labs, title is shown to make sure the graph is filled with suitable title. The facet_wrap() is to specify the data that is used, facet_wrap(~origin) is used to determine the data for wind speed. Theme is also applied to make the background of table.

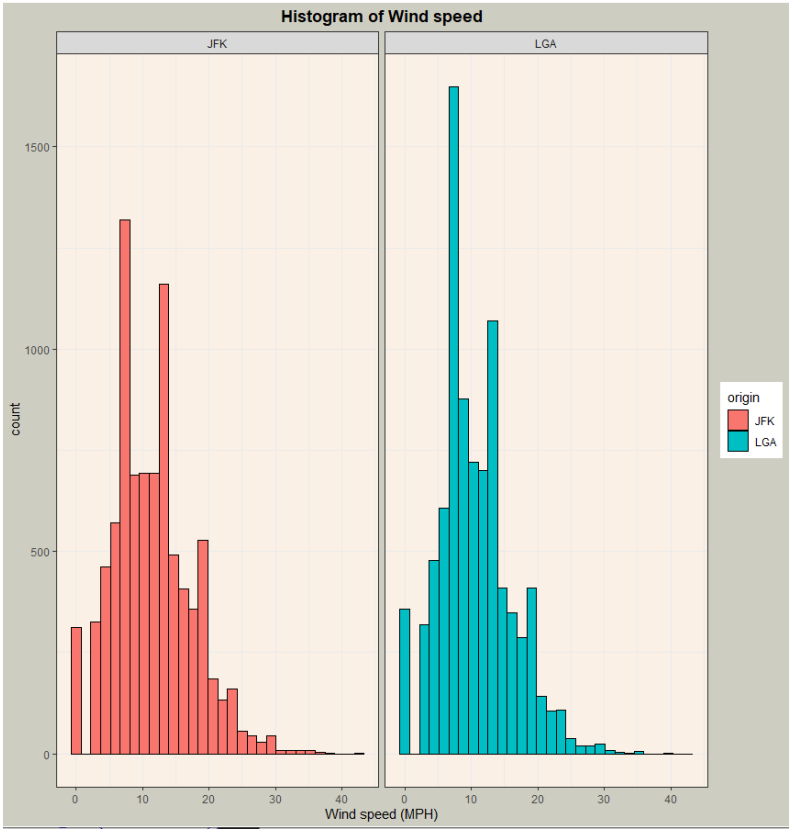

Figure 11 graph analysis 3

Based on the analysis example 3, the histogram shows the wind speed that have been detected from both airports. From the result, LGA airport have a higher count of 0 to 10 MPH wind speed than the JFK airport. This shows the airport of LGA occurs more windy day than JFK airport. (7 Reasons New TSA Temperature Checks Will Make Air Travel Worse - View from the Wing, 2020) Having a strong wind would increase the difficulty to let the aircraft lands as strong

surface wind could cause a turbulence to an aircraft which would makes the pilot job more difficult to land the aircraft. (How do wind conditions affect flight? | Finavia, 2020) Generally, the aircraft is not allowed to takeoff or landing when it is in crosswinds(wind above 30-35 kts) .As a safe wind speed for an aircraft to land or fly is below 30 to 35 kts (34 to 40 mph).

Analysis example 4

```
#Analysis example 4
ggplot(data = data, mapping = aes(x = temp, fill = origin)) +
  geom_histogram(bins = 30, col="black") +
  labs(title = 'Histogram of Temperature',x = 'Temperature (F)') +
  facet_wrap(~day) + theme_bw() + theme(plot.title = element_text(hjust = .5, face = "bold", colour = "black"),
                                        panel.background = element_rect(fill = "linen"), plot.background = element_rect(fill = "ivory3"))
```

Figure 12 snippet code 4

To start on the graph, ggplot is used while the "aes" is used to get the data from the dataset. Then, the geom_histogram() is used for indicating the type of graph. With the labs, title is shown to make sure the graph is filled with suitable title. The facet_wrap() is to specify the data that is used, facet_wrap(~day) is used to determine the data of 'day' for precipitation. Theme is also applied to make the background of table.

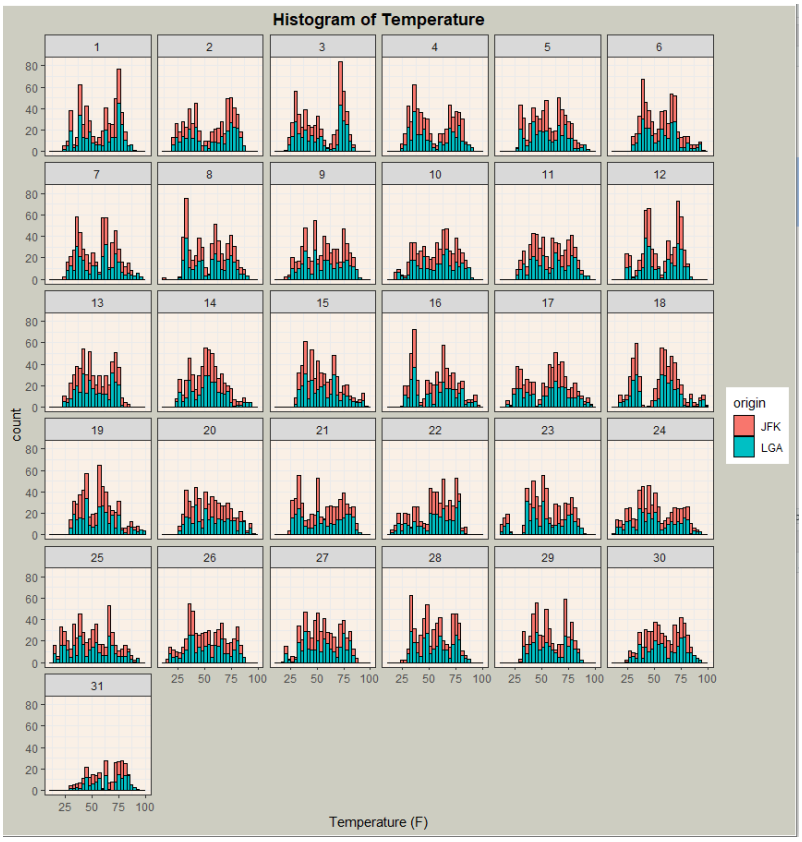

Figure 13 graph analysis 4

Based on the histogram diagram in analysis example 4, the histogram of temperature that is shown for days. From the histogram it shows that LGA airport have a lower temperature and count as a comparison with JFK airport. The JFK overall have a higher temperature and higher count, as an example, at 3rd of the particular month JFK airport had hit a temperature of 75(F)

with a count of 80. The temperature will affect how the aircraft fly as the lower the air density, the faster it is required for an aeroplane to fly to generate enough lift to take off.

Analysis example 5

```
#Analysis Example 5
ggplot(data = data, mapping = aes(x = humid, y = dewp, color = origin)) +
  +    geom_line(alpha = 0.2)+ stat_smooth(method = "lm") +
  +    labs(title = 'Scatter Plot of Dew Point against Humid',x = 'Humid ', y = 'Dew Point (F)')+
  +    theme_bw() + theme(plot.title = element_text(hjust = 0.5, face = "bold", colour = "black"),
                          +panel.background = element_rect(fill = "linen"), plot.background = element_rect(fill = "ivory3"))
```

Figure 14 snippet code 5

To start on the graph, ggplot is used while the "aes" is used to get the data from the dataset. Then, the geom_line() is used for indicating the type of graph. With the labs, title is shown to make sure the graph is filled with suitable title. The theme is also applied to make the background of table.

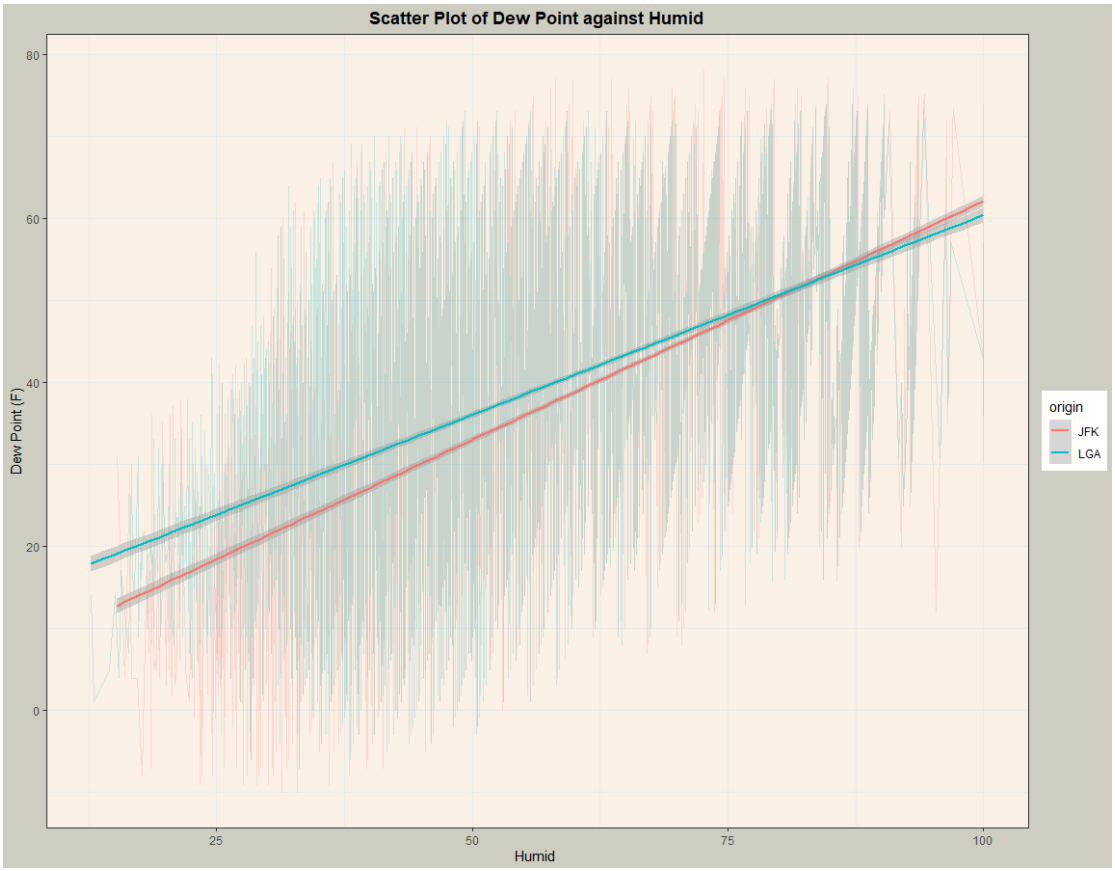

Figure 15 graph analysis 5

Based on the analysis example 5, scatter plot method is used to research the humidity of both airport weather. The scatter plot graph is drawing by using the geom_line function to draw. The humidity of JFK is higher than LGA, based on the diagram, the relationship between dew point against Humidity is the higher the humidity the higher the dew point (F). The dew point gives the pilots humidity information in relation to the temperature, which can affect visibility. (Aviation

Glossary - Relative-humidity, 2020) The humidity is high if the dew point is close to the temperature, which can create hazy conditions or even fog. In terms of aircraft efficiency, if combined with pressure and temperature, the dew point is also needed to determine the true altitude of density. A high dew point means an altitude of greater density, which decreases the efficiency of the aircraft.

Analysis example 6

```
#Analysis Example 6
ggplot(data = data, mapping = aes(x = temp, y = pressure, color = origin, na.rm = TRUE)) +
  geom_point(alpha = 0.2)+ stat_smooth(method = "lm") +
  labs(title = 'Scatter Plot of Pressure against Temperature',x = 'Temperature ', y = 'Pressure (Milibars)')+
  theme_bw() + theme(plot.title = element_text(hjust = 0.5, face = "bold", colour = "black"),
                     panel.background = element_rect(fill = "linen"), plot.background = element_rect(fill = "ivory3"))
```

Figure 16 snippet code 6

To start on the graph, ggplot is used while the "aes" is used to get the data from the dataset. Then, the geom_point() is used for indicating the type of graph. With the labs, title is shown to make sure the graph is filled with suitable title. The theme is also applied to make the background of table.

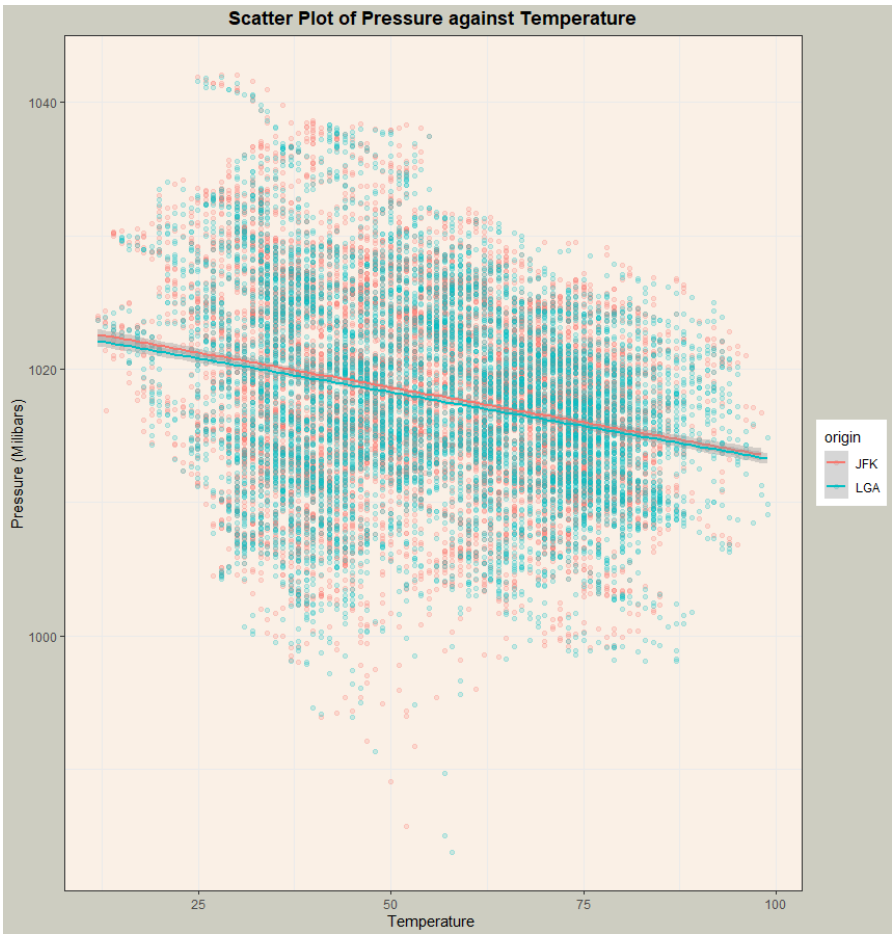

Figure 17 graph analysis 6

Based on the diagram, the scatter plot shows that the relationship of pressure against temperature. As the temperature rises, the air starts to rise upwards. It produces a low-pressure zone, but the air continues to sink downwards as the temperature decreases, producing a high-pressure

environment on the surface of the ground. Hence, the temperature and pressure are inversely connected to each other when the temperature increases the pressure increases. (Effects of Pressure & Density Altitude on Aircraft Performance, 2020) As an example, high air temperatures influence the dynamics of how aircraft fly, which means that on hot days, aircraft takeoff efficiency can be compromised. The amount of lift produced by an aeroplane wing is influenced by the air density. The lower the air density, the faster an aircraft must fly to create adequate lift to take off.

Analysis example 7

```
#Analysis Example 7
ggplot(data = data, mapping = aes(x = pressure, y = dewp, color = origin, na.rm = TRUE)) +
  geom_point(alpha = 0.2)+ stat_smooth(method = "loess",se=F) +
  labs(title = 'Scatter Plot of Pressure against Dewpoint',x = 'Pressure ', y = 'Dewpoint')+
  theme_bw() + theme(plot.title = element_text(hjust = 0.5, face = "bold", colour = "black"),
                     panel.background = element_rect(fill = "linen"), plot.background = element_rect(fill = "ivory3"))
```

Figure 18 snippet code 7

To start on the graph, ggplot is used while the "aes" is used to get the data from the dataset. Then, the geom_point() is used for indicating the type of graph. With the changes of method to "loess",se=F to fit a line to a scatter plot or time plot where the capacity to see a line of best fit interferes with noisy data values, scattered data points or poor interrelationships. With the labs, title is shown to make sure the graph is filled with suitable title. The theme is also applied to make the background of table.

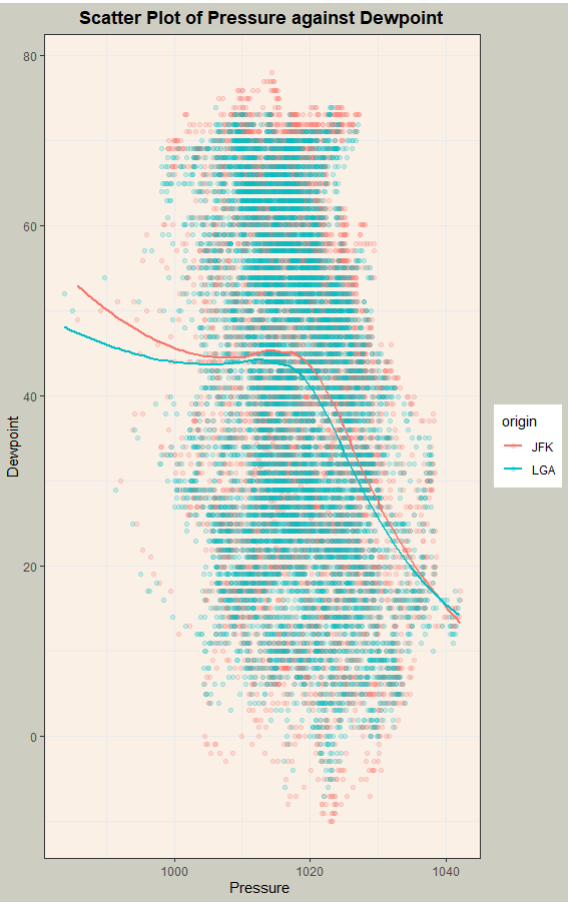

Figure 19 graph analysis 7

Based on the scatter plot of pressure against dew point, dew point is known as the amount of water in the air where the higher the dew points, the higher the moisture content at a given

temperature. (What is a pressure dew-point?, 2020) This means that if the pressure increases, the dew point also increases. As increment in dew point does provides many information to the pilot about the humidity of the weather of the day, the difference between the temperature of a place and its dew point shows how well the air will retain moisture at present. Meanwhile, the pressure is important as it determines how the altimeter of the plane works and it affects the atmosphere. To calculate how high your aircraft is above sea level.

Analysis example 8

```
#Analysis Example 8
ggplot(data = data, mapping = aes(x = temp, y = dewp, color = origin)) +
  geom_point(alpha = 0.2)+ stat_smooth(method = "lm",se=F) +
  labs(title = 'Scatter Plot of Dew Point against Temperature',x = 'Temperature(F) ', y = 'Dew Point')+
  theme_bw() + theme(plot.title = element_text(hjust = 0.5, face = "bold", colour = "black"),
                     panel.background = element_rect(fill = "linen"), plot.background = element_rect(fill = "ivory3"))
```

Figure 20 snippet code 8

To start on the graph, ggplot is used while the “aes” is used to get the data from the dataset. Then, the geom_point() is used for indicating the type of graph. With the changes of method to “lm”,se=F to fit a line to a scatter plot or time plot where the capacity to see a line of best fit interferes with noisy data values, scattered data points or poor interrelationships. With the labs, title is shown to make sure the graph is filled with suitable title. The theme is also applied to make the background of table.

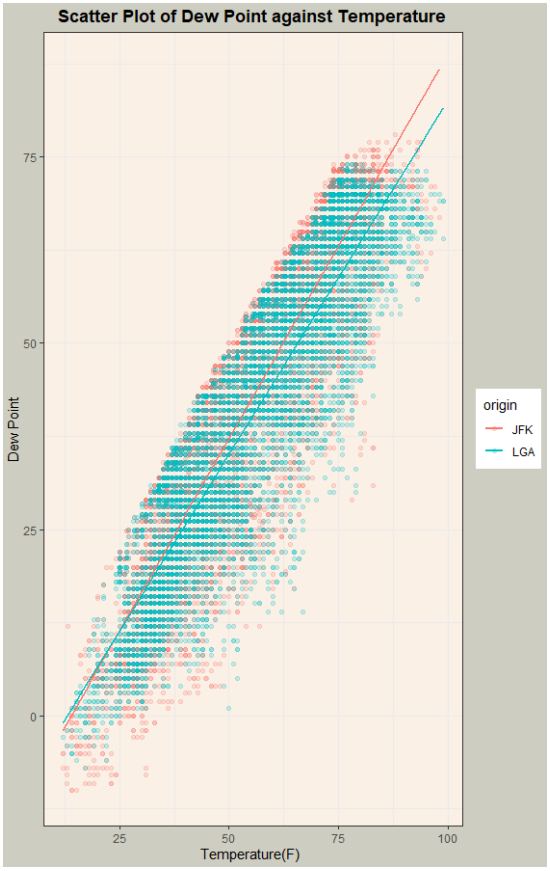

Figure 21 analysis graph 8

Based on the graph that have been shown on analysis 8, the scatter plot diagram is used to measure the Dew point against the temperature. The relationship that can be found based on the graph is the higher the temperature the higher the dew point. To calculate the dew point, is

dividing the total mass of water vapor by the volume of the air. As an example, the amount of moisture in the air is indicated by dew points. At a given temperature, the higher the dew points, the greater the moisture content of the air. As this would makes the passenger feel uncomfortable and sticky as the water molecules in the air is higher than air molecules.

Analysis example 9

```
#Analysis Example 9
ggplot(data = data, mapping = aes(x = humid, y = precip, fill=origin)) +
  geom_point(alpha = 0.2) + stat_smooth(method = "lm") +
  labs(title = 'Scatter Plot of Precipitate against Humid in every month',x = 'Humid ', y = 'Precipitate (Inch)')+ facet_wrap(~month)+
  theme_bw() + theme(plot.title = element_text(hjust = 0.5, face = "bold", colour = "black"),
                     panel.background = element_rect(fill = "linen"), plot.background = element_rect(fill = "ivory3"))
```

Figure 22snippet code 9

To start on the graph, ggplot is used while the “aes” is used to get the data from the dataset. Then, the geom_point() is used for indicating the type of graph. The facet_wrap() is to specify the data that is used, facet_wrap(~month) is used to determine the data of ‘month’ for precipitation and humidity. With the labs, title is shown to make sure the graph is filled with suitable title. The theme is also applied to make the background of table.

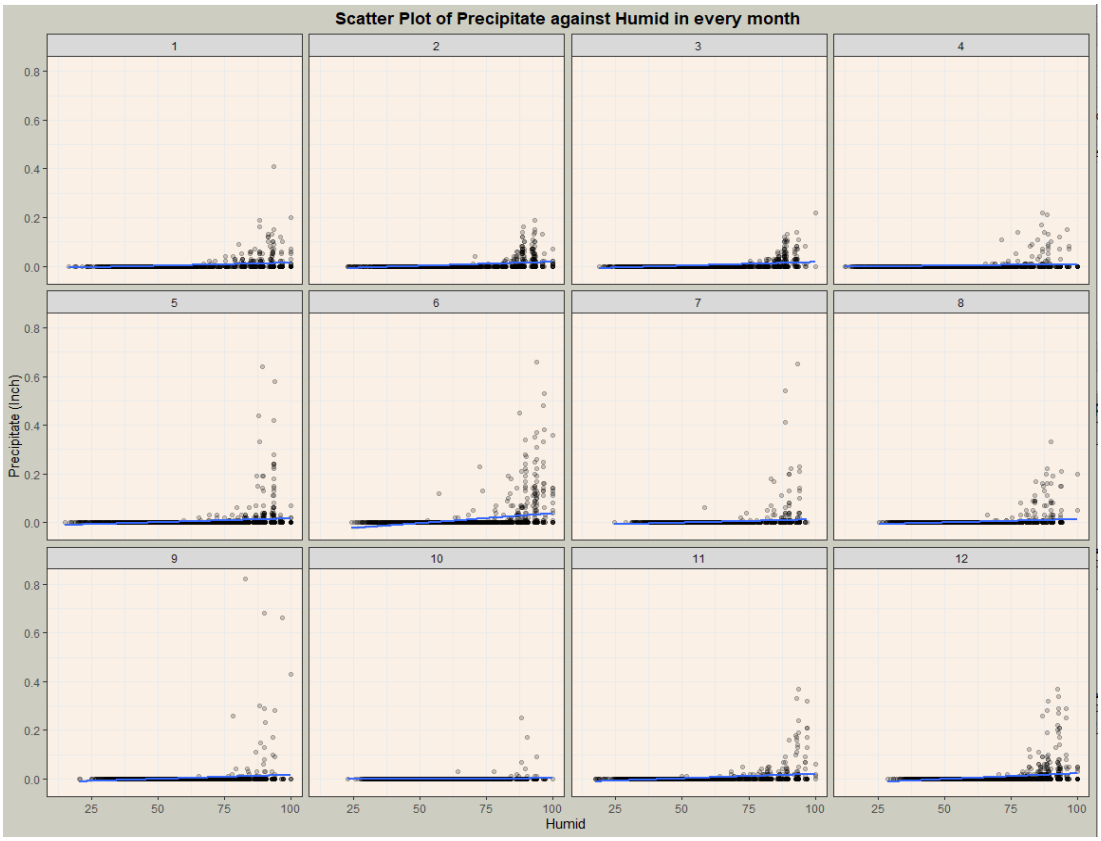

Figure 23 analysis graph 9

Based on the diagram, the scatter plot of precipitate against humid in every month is shown. As from the data that have been shown, the least humidity is in October while the highest humidity is in June. The relationship that is found in the diagram is the higher the humidity the higher the precipitate. An all-inclusive word denoting drizzle, rain, snow, ice pellets, hail, and ice crystals

are precipitation. Hence it could affect the safety of an aircraft as precipitation at the surface of the aerodrome is from rain but could be from snow at around 1000 ft above. Snow will impact aircraft 'in the circuit' or on approach under these circumstances. Snow must be fully cleared from an aircraft by sufficient means prior to take-off, typically by the application of a de-icing fluid. It is a fatal mistake to believe that during start and take-off, snow on the plane, or on the wing, will be removed by aerodynamic forces.

Analysis example 10

```
#Analysis Example 10
ggplot(data = data, mapping = aes(x = factor(month), y = temp, color = origin)) +
  geom_boxplot() +
  labs(title = 'Boxplot of Temperature against Month',x = 'Month', y = 'Temperature (F)')+
  theme_bw() + theme(plot.title = element_text(hjust = 0.5, face = "bold", colour = "black"),
                     panel.background = element_rect(fill = "linen"), plot.background = element_rect(fill = "ivory3"))
```

Figure 24 snippet code 10

To start on the graph, ggplot is used while the “aes” is used to get the data from the dataset. Factor of month is used in the code to find out the temperature of every month. Then, the geom_boxplot() is used for indicating the type of graph. With the labs, title is shown to make sure the graph is filled with suitable title. The theme is also applied to make the background of table.

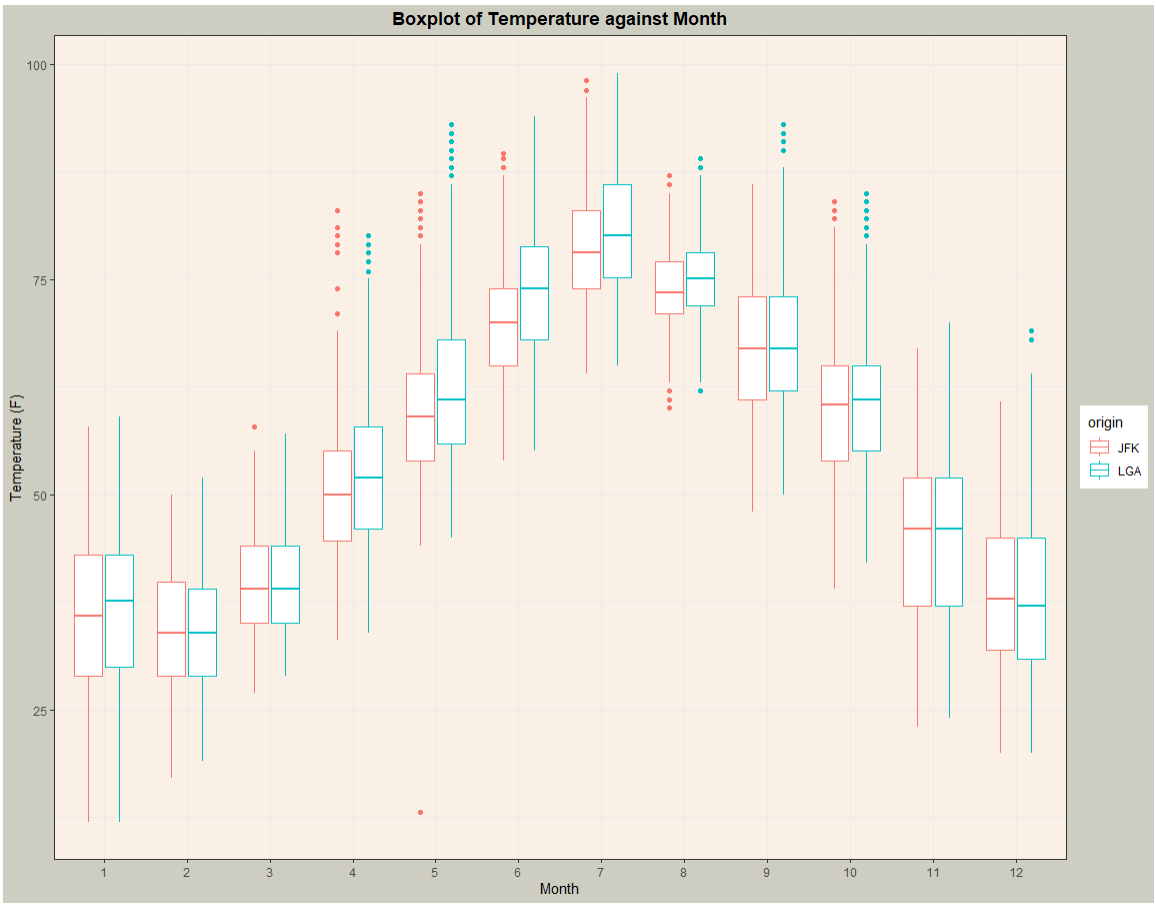

Figure 25 analysis graph 10

Based on the diagram, the box plot of temperature is shown by the graph. Based on the graph the relationship of the temperature against month the highest temperature is in July. The box plot diagram shows that the LGA airport have a higher temperature count. As an example, the

temperature on month January is at negative degrees, therefore the aircraft can have to be deiced often having to be night before with a fluid to avoid snow or ice build-up. Runways must be handled or ploughed. Lightning in the field prevents their job from being carried out by ground handlers and fuelers, regulations that employees are permitted staying outside only for brief periods of time if temperatures are too low.

Analysis example 11

```
#Analysis Example 11
data %>%
  filter(data$precip>0) %>%
  ggplot(mapping = aes(x = factor(hour), y = precip)) +
  geom_boxplot() +
  labs(title = 'Boxplot of Precipitate against Hour',x = 'Hour', y = 'Precipitate (Inch)')+
  theme_bw() + theme(plot.title = element_text(hjust = 0.5, face = "bold", colour = "black"),
                     panel.background = element_rect(fill = "linen"), plot.background = element_rect(fill = "ivory3"))
```

Figure 26 snippet code 11

To start on the graph, data is used %>% with the code magrittr (*November, 2014. Stefan Milton Bache*) which is a package with aims to reduce time taken and also improve readability and maintainability of code. Next ggplot is used while the "aes" is used to get the data from the dataset. Factor of hour is used in the code to find out the factor of hour and precipitation. Then the geom_boxplot() is used for indicating the type of graph. With the labs, title is shown to make sure the graph is filled with suitable title. The theme is also applied to make the background of table.

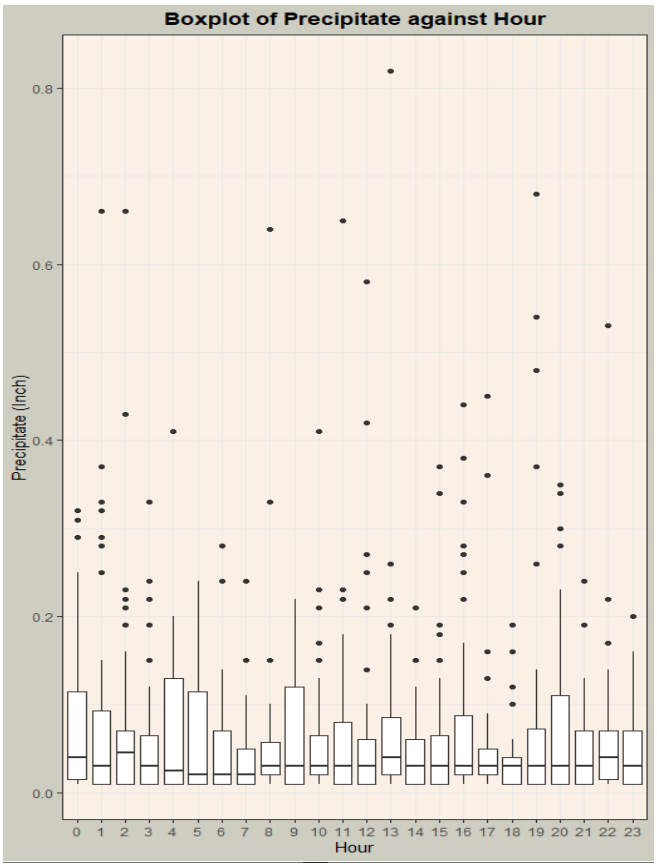

Figure 27 analysis graph 11

Based on the graph that have been shown, for 24 hour per day. It shows that the precipitate increases on 4 am, 5 am, 9am, based on the graph the highest of precipitate is at 9 am morning, as a contrast the least precipitate is on 6pm. Based on the information precipitation light rain will occurs of 0.10 inches per hour or less, normal rain will occurs at a rate of 0.11 to 0.30 inches per hour and heavy rain at a rain rate of 0.31 inches per hour or greater. Hence, it will affect the aviation as raining heavily visibility can be impaired which makes the flight unsafe to take off.

Analysis example 12

```
#Analysis Example 12
ggplot(data = data, mapping = aes(y = factor(visib), x = humid, color =origin)) +
  geom_boxplot() +
  labs(title = 'Boxplot of Humid against Visible',x = 'Humid', y = 'Visible (Miles)')+
  theme_bw() + theme(plot.title = element_text(hjust = 0.5, face = "bold", colour = "black"),
  panel.background = element_rect(fill = "linen"), plot.background = element_rect(fill = "ivory3"))
```

Figure 28 snippet code 12

To start on the graph, ggplot is used while the "aes" is used to get the data from the dataset. Then, the geom_point() is used for indicating the type of graph. With the changes of method to "lm",se=F to fit a line to a scatter plot or time plot where the capacity to see a line of best fit interferes with noisy data values, scattered data points or poor interrelationships. With the labs, title is shown to make sure the graph is filled with suitable title. The theme is also applied to make the background of table.

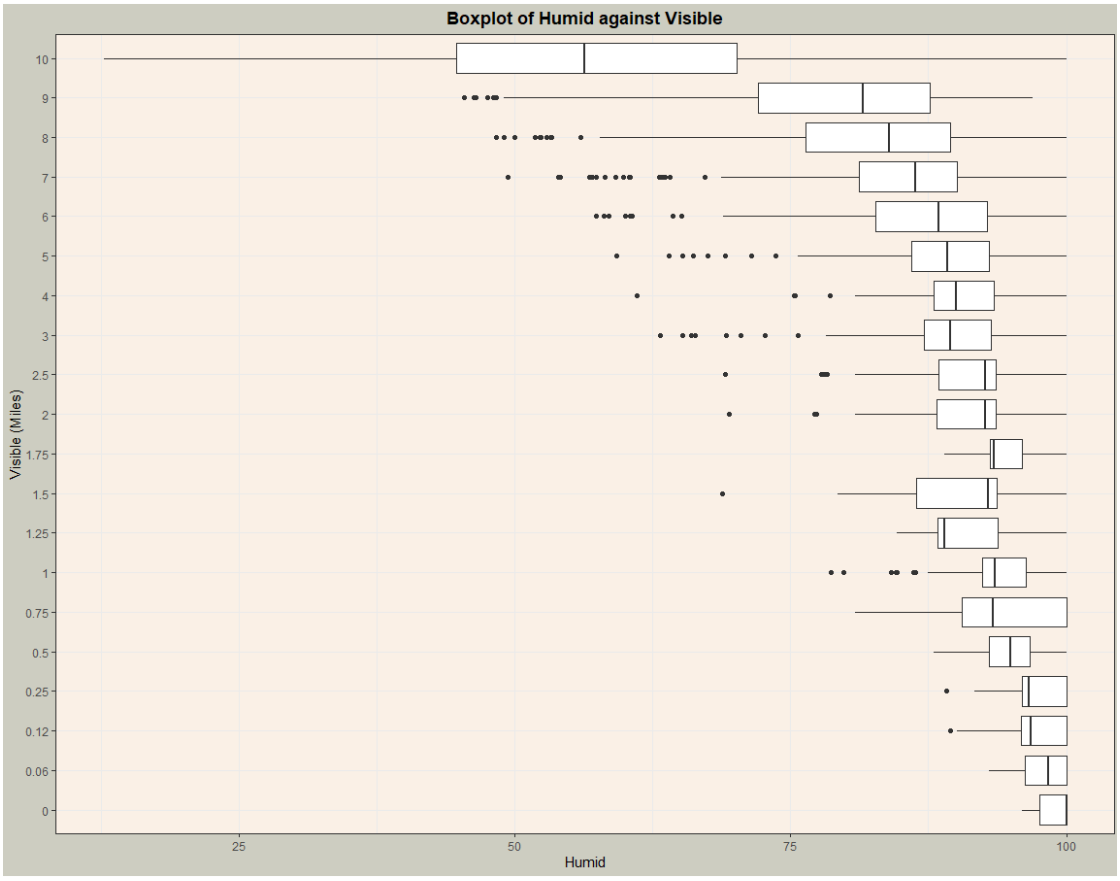

Figure 29 analysis graph 12

Based on the box plot of humidity against visible, when the humidity of the graph is at 50 to 75, the visible miles of is at 10 miles. The relationship that have been found is the lower the

humidity the higher the visible miles. When the humidity of weather is high, the temperature will decrease while low visibility will occur as fog may occur due to cold weather. Therefore, it is not suitable for aircraft to fly as it is low visibility on the runway of airport will be blur and unclear. As a contrast, if visibility is humidity is low then the temperature will be high and thus visibility will be good and clear and aircraft is suitable to fly.

Analysis example 13

```
#Analysis Example 13
ggplot(data = data, mapping = aes(x = factor(day), y = pressure, na.rm = TRUE)) +
  geom_boxplot() +
  labs(title = 'BoxPlot of Pressure against day',x = 'Day', y = 'Pressure (Milibars)')+
  theme_bw() + theme(plot.title = element_text(hjust = 0.5, face = "bold", colour = "black"),
                     panel.background = element_rect(fill = "linen"), plot.background = element_rect(fill = "ivory3"))
```

Figure 30 snippet code 13

To start on the graph, t ggplot is used while the "aes" is used to get the data from the dataset. Factor of day is used in the code to find out the factor of date and pressure. Then the geom_boxplot() is used for indicating the type of graph. With the labs, title is shown to make sure the graph is filled with suitable title. The theme is also applied to make the background of table.

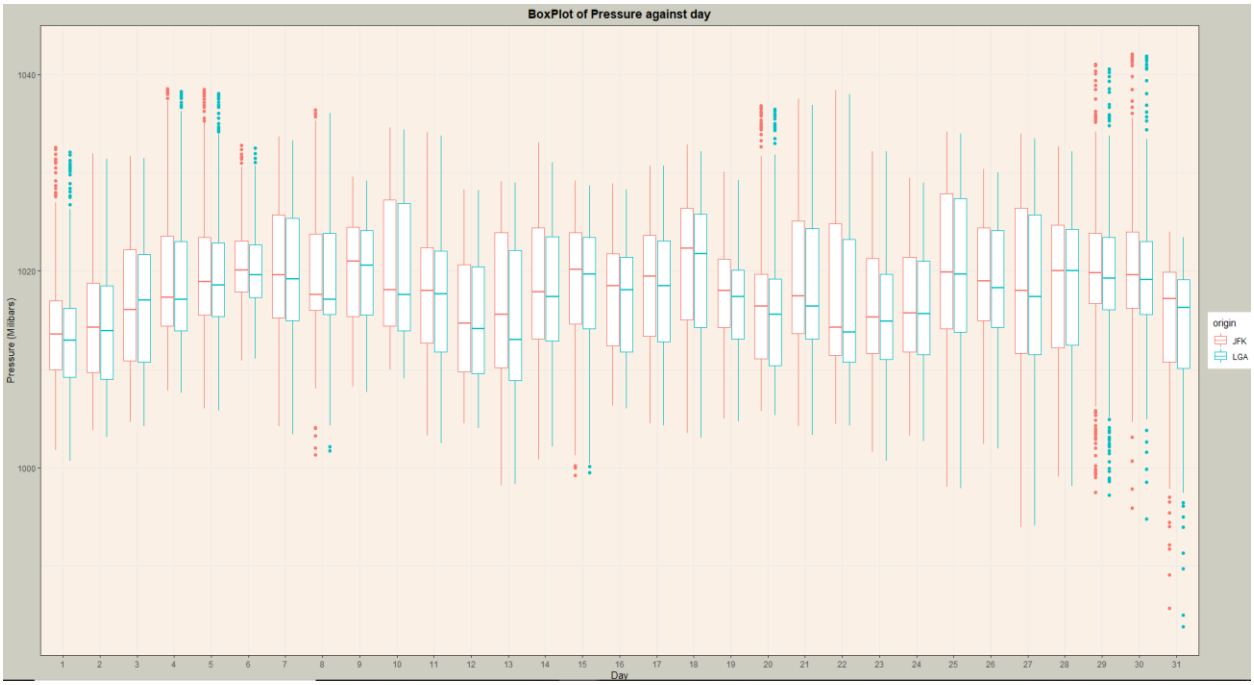

Figure 31 analysis graph 13

Based on the diagram, the box plot shows that the pressure that have been collected in a month. Barometer is a device that is used to measure the atmospheric pressure. As air pressure varies due to warmer and colder weather systems, it is used to forecast the weather. Atmospheric pressure is produced by the air that covers the Earth. The air is thinner when you go up into the mountains or ride high in an aeroplane and the noise is less. A rising barometer indicates that air pressure is increasing; a dropping barometer indicates that air pressure is decreasing. As an

example, the mercury or barometer reading increases when the air is dry. The barometer reading falls when the air is warm and wet. It normally means some form of storm or rainy weather is coming when the air pressure falls. It also means clear weather as it rises. There will be no immediate improvement in the weather if the barometer stays steady. Based on the result that have been found, on day 10^{th} of the month the millibars (mb) reading is above 1020 mb which shows that the weather on that day is cloudy and warm conditions. While on 1^{st} of the month the readings of barometer is around 1000mb to 1020mb which means on that day a rainy and stormy weather would occur.

Analysis example 14

```
#Analysis Example 14
 theme_set(theme_bw())
 ggplot(data = data, mapping = aes(x = factor(month), y = wind_gust, color = origin)) +
  geom_violin() +
  labs(title = 'Violin plot of Wind_Gust against Month',x = 'Month', y = 'Wind_Gust')+
  theme_bw() + theme(plot.title = element_text(hjust = 0.5, face = "bold", colour = "black"),
  panel.background = element_rect(fill = "linen"), plot.background = element_rect(fill = "ivory3"))
```

Figure 32 snippet code 14

To start on the graph, ggplot is used while the “aes” is used to get the data from the dataset. Factor of month is used in the code to find out the wind_gust of every month. Then, the geom_violin() is used for indicating the type of graph. With the labs, title is shown to make sure the graph is filled with suitable title. The theme is also applied to make the background of table.

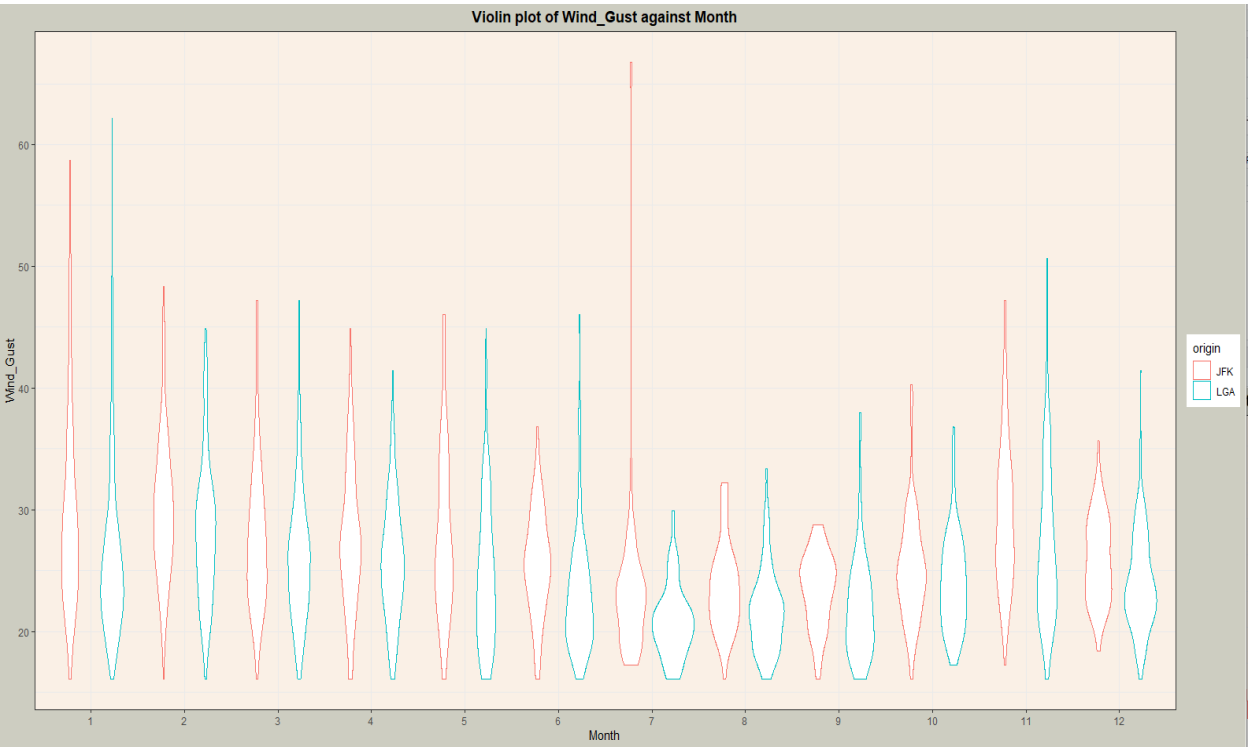

Figure 33 analysis graph 14

Based on the diagram, the violin plot of wind_gust is shown by the graph. A sharp, brief rise in the speed of the wind followed by a lull is a wind gust. Gusts are registered when the peak wind speed exceeds at least 18 mph and the difference in wind speed between the peaks and lulls is at least about 10 mph, according to National Weather Service observing experience. (What causes wind gusts?, 2020)The aircraft takes off and crosses a track of roaring winds. It undergoes a sharp increase in lift at this point, propelling it up quickly. The lift decreases when the aircraft

leaves the road, in turn. The characteristics required to produce gradients with extreme wind shear are generated as a consequence of the numerous changes in wind strength and direction in the vertical wind profile. As an example, the highest wind gust that is found is in January where both of the airport achieves the highest wind gust as a comparison with other months.

Analysis extra feature 1

```
#Extra Feature 1
ggplot(data = data) +
  geom_bar(mapping = aes(x = wind_dir, na.rm = TRUE, fill= origin)) +
  coord_polar(theta = "y") +
  labs(title = 'Polar Bar Plot of Wind Direction',x = 'Wind Direction ')+
  theme_bw() + theme(plot.title = element_text(hjust = 0.5, face = "bold", colour = "black"),
  panel.background = element_rect(fill = "linen"), plot.background = element_rect(fill = "ivory3"))
```

Figure 34 snippet code extra feature 1

To start on the graph, ggplot is used while the "aes" is used to get the data from the dataset. Factor of wind speed is used in the code to find out the of the year. Then, the coord_polar() is used for indicating the type of graph. With theta = "y" the circular graph is shown with direction stated around the graph. With the labs, title is shown to make sure the graph is filled with suitable title. The theme is also applied to make the background of table.

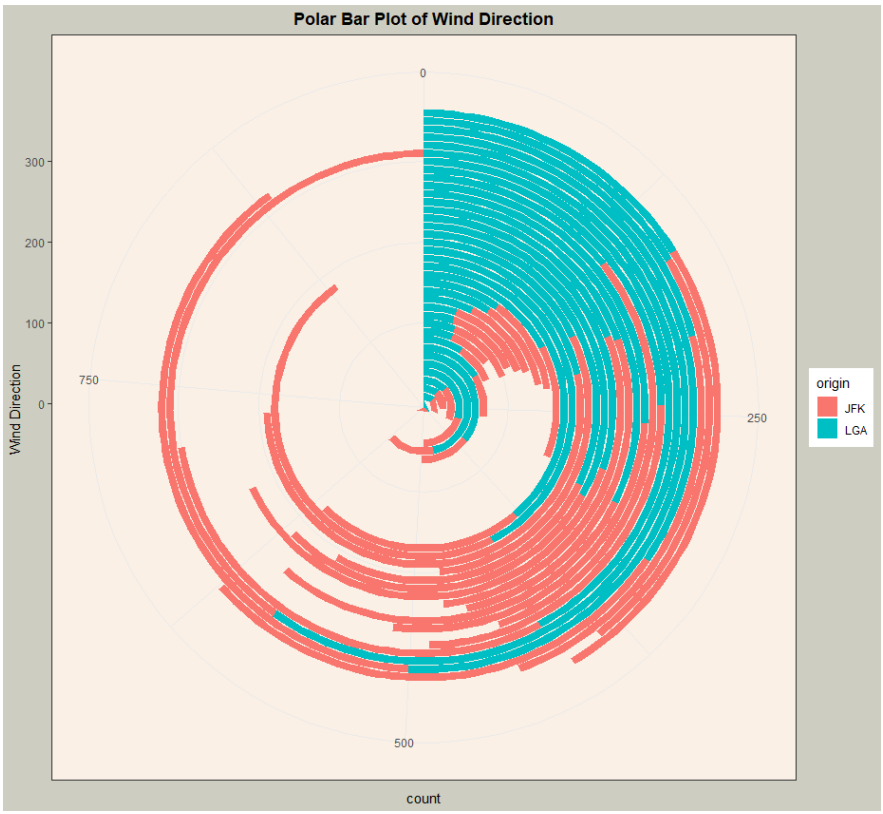

Figure 35 analysis graph extra feature 1

Based on the diagram the wind direction, shows that JFK airport and LGA airport wind direction which where the wind direction goes to. The wind direction of LGA is from north to southwest while the JFK airport detects a wind from northeast till northwest which shows a bigger count of

wind count. As an example, wind direction is important as it could affect the aircraft land and take off. The headwind is the aircraft's wind blowing. 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 advantage as at touchdown, it uses less runway, and ground speed is lower.

Analysis extra feature 2

```
#Extra feature 2
ggplot(data = data, mapping = aes(x = wind_speed, y = wind_gust, na.rm = TRUE, color = origin)) +
  geom_point(alpha = 0.2)+ stat_smooth(method = "lm") +
  geom_smooth(method="loess", se=F) +
  labs(title = 'Scatter Plot of Wind Gust against Wind Speed',x = 'Wind Speed (MPH)', y = 'Wind Gust (MPH)')+
  theme_bw() + theme(plot.title = element_text(hjust = 0.5, face = "bold", colour = "black"),
                     panel.background = element_rect(fill = "linen"), plot.background = element_rect(fill = "ivory3"))
```

Figure 36 snippet code extra feature 2

To start on the graph, ggplot is used while the "aes" is used to get the data from the dataset. Then, the geom_line() is used for indicating the type of graph. With the labs, title is shown to make sure the graph is filled with suitable title. The theme is also applied to make the background of table.

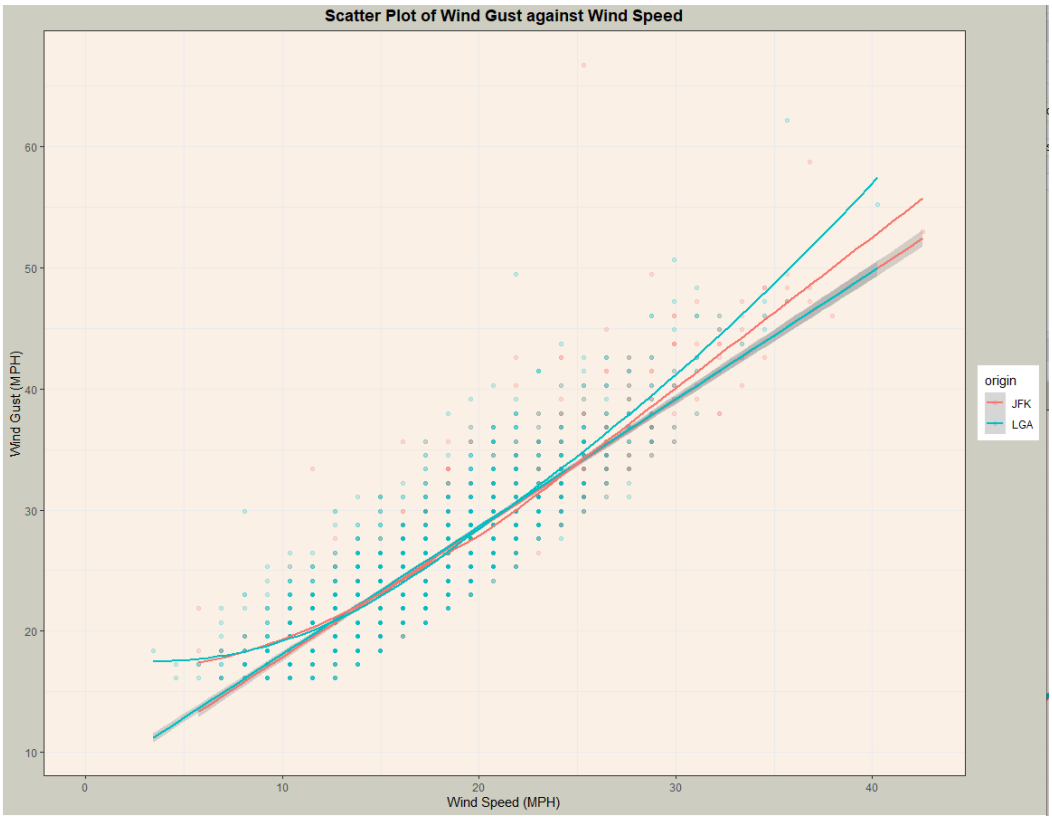

Figure 37 analysis graph extra feature 2

Based on the analysis example, the graph is shown to measure the wind gust against the wind speed. The relationship that is found based on the graph is the higher the wind speeds the higher the wind gust. Wind gust is known as a wind type that is faster than average wind speed by 18 miles per hour. Hence it is shown that JFK have a higher wind speed than LGA while LGA have

a lower wind speed than JFK. If the wind lasts at least 20 seconds, the average wind speed is surpassed by 10 or more knots, it is already considered to be a wind gust. This knowledge is revealed immediately after the average wind speed with the letter G (gusts), followed by the value of the gusts.

Conclusion

As a conclusion, based on the result that have been made and concluded from the dataset of LGA and JFK airport. Each airport shows different result on different aspect that have been concluded are the temperature, precipitate, humidity, wind speed, wind gus, wind direction, visibility and dew point. Based on this assignment I have gained many new knowledges about weather and how It will affect aircraft. I have gained more concepts on data analysis method through r studio.

References

- 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 19 November 2020].

- Wxguys.ssec.wisc.edu. 2020. *What Causes Wind Gusts?*. [online] Available at: <https://wxguys.ssec.wisc.edu/2012/01/09/what-causes-wind-gusts/#:~:text=A%20wind%20gust%20is%20a,at%20least%20about%2010%20mph.> [Accessed 19 November 2020].

- ANAC National Civil Aviation Agency - Brazil. 2020. *Wind Gusts*. [online] Available at: <https://www.anac.gov.br/en/safety/aeronautical-meteorology/conditions/wind-gusts#gusts> [Accessed 19 November 2020].

- Experimentalaircraft.info. 2020. *Effects Of Pressure & Density Altitude On Aircraft Performance*. [online] Available at: <https://www.experimentalaircraft.info/flight-planning/aircraft-performance-3.php#:~:text=A%20number%20of%20factors%20(altitude,in%20aircraft%20and%20engine%20performance.> [Accessed 19 November 2020].

- FST GmbH. 2020. *What Is A Pressure Dew-Point?*. [online] Available at: <https://www.fstweb.de/en/knowledge-tools/what-is-a-pressure-dew-point/> [Accessed 19 November 2020].

- Dictionary.dauntless-soft.com. 2020. *Aviation Glossary - Relative-Humidity*. [online] Available at: <http://dictionary.dauntless-soft.com/definitions/GroundSchoolFAA/relative-humidity#:~:text=The%20ratio%20of%20the%20amount,the%20prevailing%20temperature%20and%20pressure.&text=Since%20the%20latter%20amount%20is,both%20moisture%20content%20and%20temperature.> [Accessed 19 November 2020].

- Encyclopedia Britannica. 2020. *Climate - Atmospheric Humidity And Precipitation*. [online] Available at: <https://www.britannica.com/science/climate-meteorology/Atmospheric-humidity-and-precipitation> [Accessed 19 November 2020].

- View from the Wing. 2020. *7 Reasons New TSA Temperature Checks Will Make Air Travel Worse - View From The Wing*. [online] Available at: <https://viewfromthewing.com/7-reasons-new-tsa-temperature-checks-will-make-air-travel-worse/#:~:text=Passengers%20with%20a%20temperature%20reading,or%20higher%20would%20be%20flagged.> [Accessed 19 November 2020].