# flightAnalysisDoc

Kerin Grewal

May 2, 2018

Introduction

For this project I choose to analyze flights from six different cities to see if the weather directly impacts the delay time. My hope is that flight departure delays are dependent enough on the airport they are departing from, and not overly dependent on the airport which the plane is arriving from and the aiport that the flight will be landing at. If they are too dependent on these other factors then the coorelation between flights and weather may be low. My data came from developer.flightstats.com, and I uesd both Flight Status & Track by Airport (departures), and Weather APIs. The flight tracking information goes from about 9:45 am to 3:45 pm on Wednesday, April 25th.

Airports Used

Here are the airports I used. I selected them based off of their weather on April 25th, as I wanted to have a variety of temperatures and conditions. I used USA Today National Weather map to make my selections.
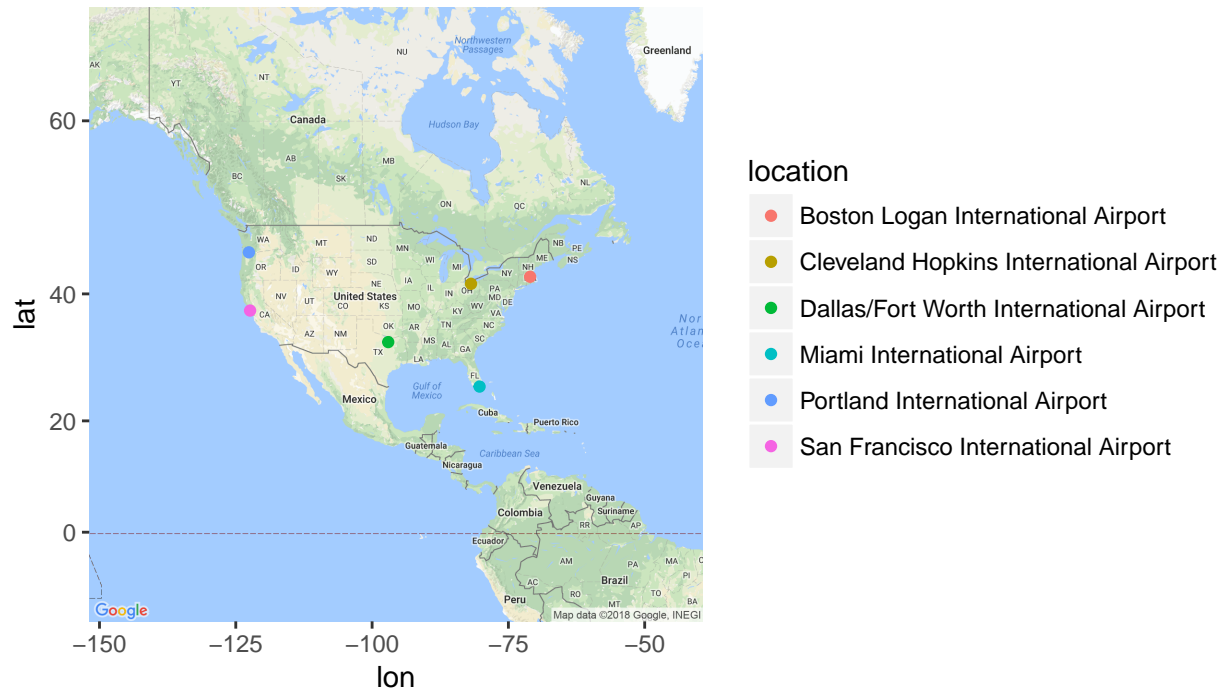
Using the Afternoon Forecast

The first data I used to examine the weather conditions was the afternoon forecast. Using text mining, I ranked each city's forecast as positive and negative. I then graphed the data, subtracting the negative words. I included a scaled down version of the delay minutes, to show how severe the average delay time was in contrast to the weather. The bars represent the weather, and points represent avg. delay.

1

Figure 1: Good Weather

Good Forecast Data by City

```
##   word city
## 1   10  BOS
## 2   10  CLE
## 3   10  DFW
## 4   10  SFO
## 5   15  BOS
## 6   15  DFW
```

These words have have been chosen from the weather forecasts to represent good weather. The numbers, e.g. 5 and 10 represent wind speed, and are included because these are low wind speeds.

Bad Forecast Data by City

```
##      word city
## 1    100  BOS
## 2     30  BOS
## 3     50s  BOS
## 4     50s  CLE
## 5 cooler  BOS
## 6 cooler  DFW
```

These words have have been chosen from the weather forecasts to represent bad weather. The numbers, e.g. 20 and 30 represent wind speed, and are included because these are high wind speeds.

Graph of Afternoon Weather Condition and Delay Time

The delay times have been scaled to appropriately fit the weather data. The bars represent the weather conditions and the points represent the delays.

General Weather Forecast and Delay Minutes
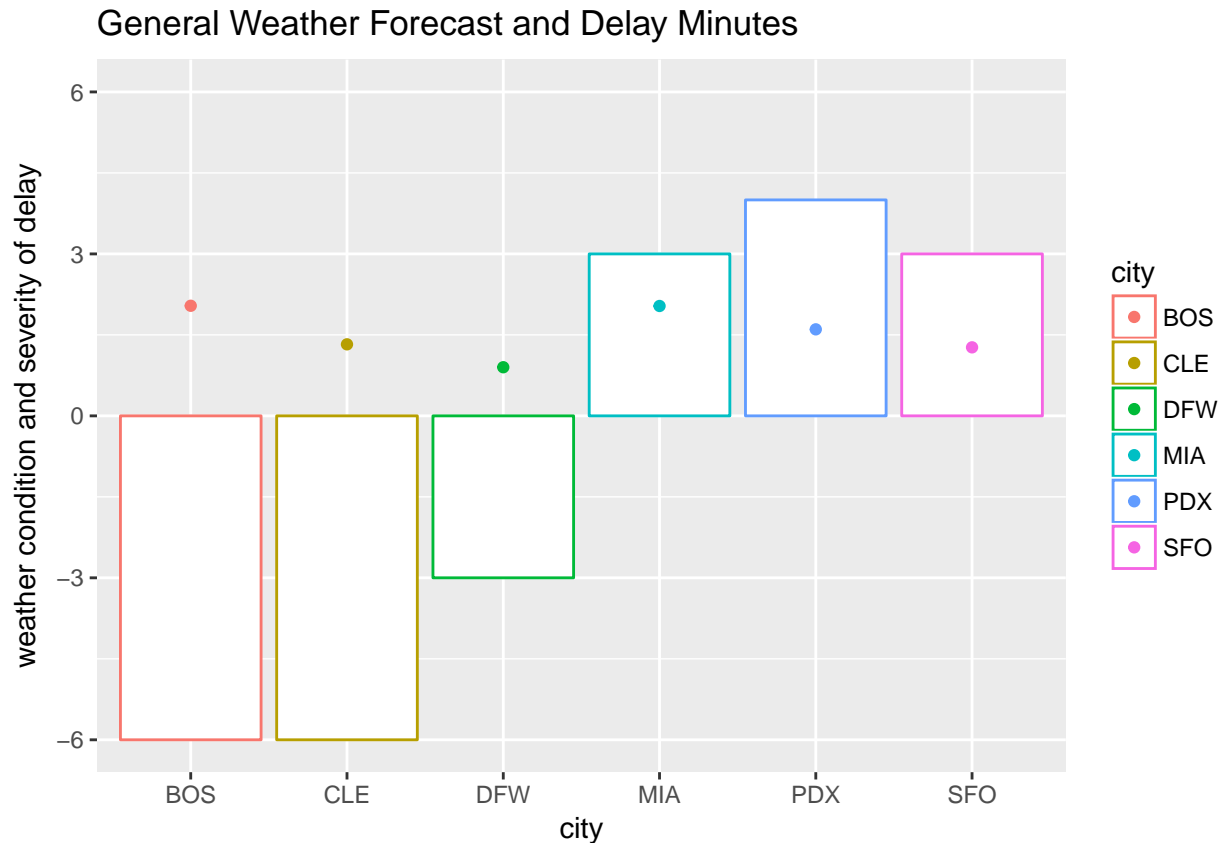
Week Long Weather

The afternoon weather patterns appeared to be matching up to the delay times for Boston, Cleveland, and Dallas, but moving on to Miami, Portland, and San Francisco the data is less accurate, leading me to believe that the first three accuracies may have been by chance.

I found that the data from the daily weather forecast was not enough. There are too few words to accurately describe the weather, and I decided that the typical weather patterns may be more indicative of airport delays than a specific afternoon forecast may be. This may have changed data, for example if the day when I took the flight delay data was sunny with light winds and the rest of the week ended up being rainy and windy.

Graph of Week-Long Weather Condition and Delay Time

The delay times have been scaled to appropriately fit the weather data. The bars represent the weather conditions and the points represent the delays.

Week–Long Weather Forecast and Delay Minutes
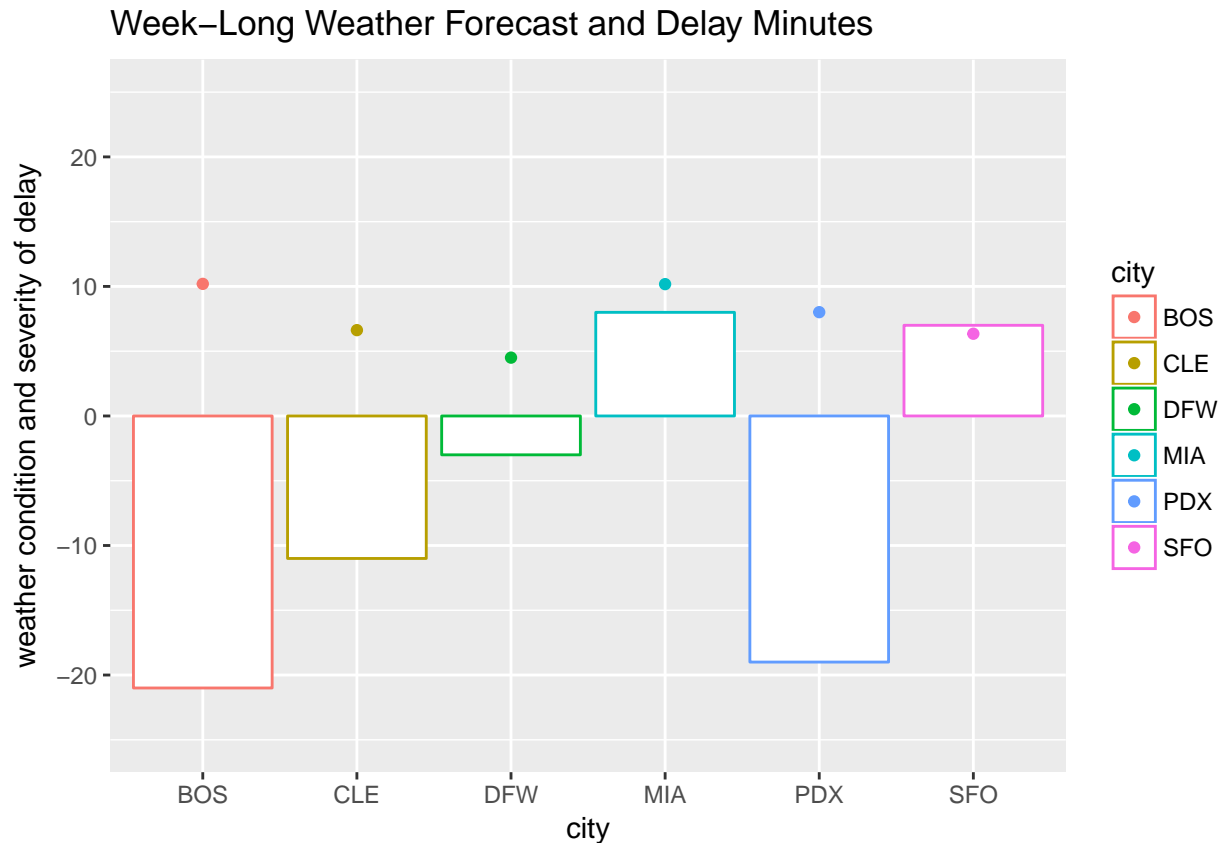
Conclusions

Based on My Graphs

The delay times in Boston, Cleveland, Dallas, and Portland show a coorelation to the weather. In San Francisco, there is a slight correlation. In Miami, there is no apparent coorelation. The long delays in Miami may have to do with some factor not related to weather, or conditions at other airports which lead to a backup in Miami.

Data

The limitations of my data was that the forecasts are relatively short, and that means there are not many words in each. Ranking the weather based on these words is therefore not completely accurate.

Since the flight tracks only give you a 6 hour window, I did not have a lot of data to use for flight delays and this may have made the delay data inaccurate.

Timing

I felt the timing was not ideal for this particular project, because in America in April there is very little severe weather, so the biggest hinderances will likely be rain and fog, whereas in January or February you may find more issues with snow, ice, and freezing temperatures.