

Table of Contents:

Logan International Airport Domestic Flight Visualization

Mehmet Duman
01139404

As I mentioned in my proposal my aim is to answer a few questions by using the available data. I used the publicly available US Domestic Flight Dataset filtered for Boston Logan Airport. The Flight dataset for each year has around five million records. I used Python for filtering, data manipulation and data cleaning. I created json files with Python for JavaScript visualization. Here are the visualizations I have created so far with my data:

Figure 1: 'BOS-Logan Top 10 Popular Destinations'

I used this figure to show the answer for question "What is the most popular destinations from BOS or an origin to BOS by visualization in a geographic map?". I used Choropleth Map and D3 Brushing (Linked Highlighting) method to create this visualization to show the popular top ten destinations from Boston Logan Airport. I think this visualization is very efficient way to show the popular destination airports. I used Flight Dataset and Airport Data (airport.csv) to find airports' latitude and longitude information.

Figure 2: 'BOS-Logan Airport Day of Week Departure/Arrival Flight'

I also have Figure 2, which is a Heatmap showing BOS-Logan Airport Day of Week Departure/Arrival Flight. As I mentioned in my proposal by this graph I show the change of day of week amount of flight numbers in the years. This heatmap allows us to compare the days of the week. As it shows clearly, Saturday is the slowest day of the week by number of flights.

Figure 3: 'BOS-Logan Airport Monthly Departure/Arrival Flight' and Figure 5: 'BOS-Logan Monthly Line Chart'

First I used Figure 3: 'BOS-Logan Airport Monthly Departure/Arrival Flight' heatmap to

answer the question “What is the amount of yearly flight number and how it changed in years?”. However, even when I added monthly flight numbers for “mouseover function”, that heatmap was not good enough to show the correlation between twenty years’ data. I decided to go to different kind of visualization. I used Line Chart and created Figure 5: ‘BOS-Logan Monthly Line Chart’ and got much better visualization and clear correlation between years. You can see from the Line Chart, the flight number increases over the years. As we see from the chart, there is a sharp drop on September 2001, and as a result of 9/11. There are many other dramatic changes as we can see in the Line Chart and I am going to do more research for the reasons behind these changes.

I also used both Figure 3: ‘BOS-Logan Airport Monthly Departure/Arrival Flight’ and Figure 5: ‘BOS-Logan Monthly Line Chart’ to give answer to my other question “Flight number grouped monthly and compare with other years”.

Figure 4: ‘BOS-Logan Airport Hourly Departure/Arrival Flight’

To answer the question “What times is the airport busiest”, I used Heatmap and create Figure 4. From this graph we see that in almost all the years, morning busiest hour is 7-8am, and afternoon busiest hour is 5-6pm. The graph also shows us that there are very few flight between midnight to early morning. We also clearly see that during the years the total number of the flights increased, but the busiest hour pattern did not change.

I am still working on to answer my other questions “What is the on-time/delay performance by years?” and “What is the correlation between population growth and number of flights?”. I am planning to add a few more visualization to my final project.