BUSINESS ANALYTICS NANODEGREE PROGRAM

Project: Build Data Dashboards

Data Set: 2015 Flight Delays and Cancellations

Insight 1

Which airlines have the more cancellations and for which reasons?

https://public.tableau.com/profile/juan.s.pinzon#!/vizhome/Insight1_CancelledflightsbyAirline_v2/Dashboard1?publish=yes

In 2015, Southwest Airlines had the more cancellations with 818 flights cancelled, that is 18.45% of the total cancelled. Atlantic Southeast airlines (18.05%) and American Eagle Airlines (16.31%) are ranked second and third respectively. Virgin America, Frontier Airlines, Alaska Airlines and Hawaiian Airlines have the less cancellation with less than 1% of the total flights cancelled.

This trend is maintained looking at the data during week days and during weekends, and there are not significant changes looking at the information each day of the week.

From the visualization, it appears that weather is the major reason for cancelled flights across all airlines in the United States. Due to the National Air System, there is a significant amount of cancelled flights for the Atlantic Southeast Airlines compared to other airlines. For the Airline/Carrier reason, Southwest Airlines has the more cancelled flights with 309.

I used a bar chart horizontal and vertical to visualize this data as it is the best way to look for insights. I used only one blue as using different colors would not give more information and maybe it could distract the audience.

Insight 2

Which destinations airports have the more delays and in what time of year?

https://public.tableau.com/profile/juan.s.pinzon#!/vizhome/Insight2_ArrivaldealysbyAirportandMonth_v2/Dashboard1?publish=yes

In general, the month of June of 2015 was the worst in terms of arrival delays. On the other hand, September was the best for the arrival delays.

By far, Chicago O'Hare International Airport had the most arrival delays with 98900 minutes (7.47% of the total) compared with the second one, Dallas/Forth Worth International Airport with 63052 minutes (4.77%). For the special case of this airport, June was the worst month in terms of arrival delays, as the general trend, but January was also a bad month. September was the best month also but, in this case, December had a lot fewer arrival delays for the Chicago O'Hare International Airport.

I used a line chart to a better view of the time line for this data. The bar chart is a good way to look at values for the different airports, and setting this chart as filter, it allows the viewer to use this to select and look for the time line for each airport. One color was also the best choice to encode data.

Insight 3

What is the relation between delays and cancellations for the different airlines?

https://public.tableau.com/profile/juan.s.pinzon#!/vizhome/Insight3_DelayscorrelationbyAirline_v2/Sheet1

After looking at arrival delays in the second insight, I started thinking about the relation between departure and arrival delays for the different airlines. From the visualization, it appears that the relationship between departure and arrival delays is positive and moderate. Airlines such as Southwest, Atlantic Southeast, Skywest, United and American Airlines have a higher departure delay than the average and a higher arrival delay, being Southwest the worst airline with the highest amount of departure delay (648419 minutes) and arrival delay (289992 minutes). The remaining airlines have lesser values of departure and arrival delays than the average, except for Delta Airlines that has a high departure delay and a low arrival delay.

I added the Cancelled measure to the size mark to look at the relationship between delays and cancellations. It appears that airlines with more delays also have a higher amount of cancelled flights.

I used a scatter plot as it is the best way to compare two quantitative variables. I added the average lines to each axis, so it is a tool to a better visualization and a way to get a better insight.