Data Visualisation: US Domestic Flights, 2013-2017

By Vincent Man

6th August 2018

Link to my Tableau Workbook
https://public.tableau.com/profile/vincent.man1586#!/vizhome/Udacityflightdata/Story?publish=yes

Summary

Delay has always been the most important indicator to measure performance especially for flights as time is money. Airlines may suffer consequences of paying penalty fines and losing their customer trust thereby affecting their brand image.
In this project, I downloaded the data from Transtats – Bureau of Transportation Statistics which monitors the on-time arrival performance of domestic flights in the US and this report is based on the flight data from January 2013 - December 2017.

Design

After downloading the data, I conducted exploratory data analysis on Tableau. Based on the Tableau collection of charts, there are 5 dashboards.
The first dashboard shows a bar chart of the total average delay for each cause of delay. At first there was only one colour for all delay categories but later advised from my interview to visualise each type of delay with a different colour.

The second dashboard illustrates a line chart which shows which month is the best time to travel in terms of the least number of delays, cancellations and diversions of flights. Initially I displayed the measure names separately as I originally thought it would be easier for viewers to understand. However, my interviewee got me thinking and it would make my analysis more efficient if I put the measure names altogether in order to see the pattern.

The third dashboard shows another line chart which is similar to the previous chart but in years. Initially I was only planning on showing both line charts in one dashboard however, after receiving my feedback from the interview it is better to display both charts separately in order to prevent confusion.

The fourth dashboard shows a map plot of which state has the highest average delay based on each airline. You can click and hover on the tree map plot below the map to choose the top airlines that has the most delays. At first, I kept the two plots on separate dashboards but later found out that I can link the tree map plot to the map plot which made it more interactive. Initially I did not display the two extra maps (Hawaii and Alaska) but later told by my interviewee to include them.

For the fifth dashboard I became curious to see a breakdown of different types of delay for airline carriers after creating the first dashboard (Types of Delay), so I decided to create a stacked bar chart. After creating the chart, I wanted to do more investigation and created the second stacked bar chart that shows the top 5 airports with a breakdown of different types of delay.

Feedback

In order to gain valuable feedbacks for my project, I decided to interview 3 people without any knowledge in Data Visualisation and the Flight dataset. After presenting them the charts I was given a lot of feedbacks.

First Interview

1. Why are there two line charts on this page? As I am confused as to what you were trying to show me (For line charts – Year and Month)
2. Why some axis has no title?
3. Can I highlight on multiple states for further analysis on the map plot?

Second Interview

1. Why the chart for 'Types of Delay' only have one colour? As I thought these delays are all the same.
2. What are you trying to show me regarding to the measure names? (Arrived Flights, All Delays, Diverted and Cancelled Flights line chart) and why are they shown separately?
3. How many cancellations and diversions have there been throughout the period?

Third Interview

1. I cannot see Hawaii and Alaska on the map that you are showing.
2. Why did you choose the top 5 airline carriers and airports?
3. Can I just look at data based on one particular year?

Resources

- Bureau of Transportation Statistics
- Tableau Documentation
- Stackoverflow
- Udacity Tableau