Udacity DAND – Data Story Telling Project Write-up

By: Jolene Yao

Project Links

Version 1:
https://public.tableau.com/profile/jolene2805#!/vizhome/ChicagoORDAnalysisforFutureTrips/Story1
Version 2:
https://public.tableau.com/profile/jolene2805#!/vizhome/ChicagoORDAnalysisforFutureTripsv2/Story1

Summary

The purpose of my data story is to perform an analysis of Chicago's biggest airport – O'hare. It starts with choosing a destination, then aircraft carrier, and finally the airfares. This process follows any process you would follow when planning trips and purchasing airplane tickets. Lastly, the year 2008 was chosen for in-depth analysis because full 2018 data was not available on all datasets.

Design

Andrea (friend)

- The BTS/RITA website does not provide statistics on airplane ticket prices for months, only quarters. I have added another storyboard onto my story for just quarters. The top airline carriers are not ones that travel just to Chicago. They travel to other airports within the US. The types of delays are given in the legend but I will change the names of the measures so that it conveys in the legend what delay it is rather than one word followed by "ct". Lastly, BTS/RITA website does not provide airfare ticket prices by aircraft carrier, it only provides ticket prices by airport.
- I have changed the story's caption so that it says "passengers" in it for the destinations visualization. The third graphic Andrea is not sure of is an interpretation issue. If you look at the axes, it specifically says years and then value with caption describing delays. I will make this more obvious by changing the legend names.

Corey (brother)

- The lines on the airfare prices look too close which is hard to read

Feedback

Andrea (friend)

- What months are the prices for airplane tickets higher? Are the top airline carriers in 2008 the main carriers who travel just to Chicago or in general? What kind of airplane delays are there? What airline carrier has the least expensive ticket prices? Which one has the most?
- The first graphic I am not sure if it explains the number of people or number of flights to Chicago. The third graphic I am not sure if I am looking at the number of delays each carrier had throughout the years.

Corey (brother)

- This is what the data has shown and I cannot change the dataset given to me. No changes will be reflected in version 2 of my story.

Resources

- Tableau.com tutorial articles
- Python notebook (for cleaning the datasets before importing into tableau)
- BTS/RITA website (https://www.transtats.bts.gov/OT_Delay/OT_DelayCause1.asp)
- Stack Overflow website (for referencing codes to clean datasets)

Data files

What is included in the zip file is 4 csv files that I have already cleaned using python notebook. I have also included the python notebook that showed how I cleaned each csv file. Lastly, the airport codes file was not included in the tableau workbook, but was used to rename all airport abbreviations to its full name so that the audience can get a better understanding of which airport it is.