Author: Quang Luong

Project: Tableau Visualization

Viz links:

v1:
https://public.tableau.com/profile/quang.luong#!/vizhome/QuangLuong_UdacityDA_Term2Project/Story1

v2:
https://public.tableau.com/profile/quang.luong#!/vizhome/QuangLuong_UdacityDA_Term2Project_v2/AvgDelaysminbyCarriersbyMonths

v3:

https://public.tableau.com/profile/quang.luong#!/vizhome/QuangLuong_UdacityDA_Term2Project_v3/Str_FlightDelaysV3?publish=yes

Links to feedback:

https://study-hall.udacity.com/sg-346183-2559/rooms/community:nd002:346183-cohort-2559-project-948/community:thread-384738845-399905?contextType=room

Summary:

This is a dataset of airline flights collected for year 2008. This dataset was downloaded from: http://stat-computing.org/dataexpo/2009/the-data.html.

The 2008 dataset contains approximately 7 million records.

Here I'm trying to convey:

1. What are the different types of delays and in which months it is most prevalent?
2. Which months and which carriers had the most delays?
3. What are the airports that cause delays to the carriers?
4. If there is a departure delays, regardless of delay types, does the carrier make up for that delay in-flight?

Design:

The dataset is left-joined to the airport dataset by the origin and the iata code to show origin city.

- The dataset is again left-joined to the airport dataset to get the destination city
- The dataset is left-joined to the carrier dataset to show the airline/carrier

Version 1 (12/20/2018):

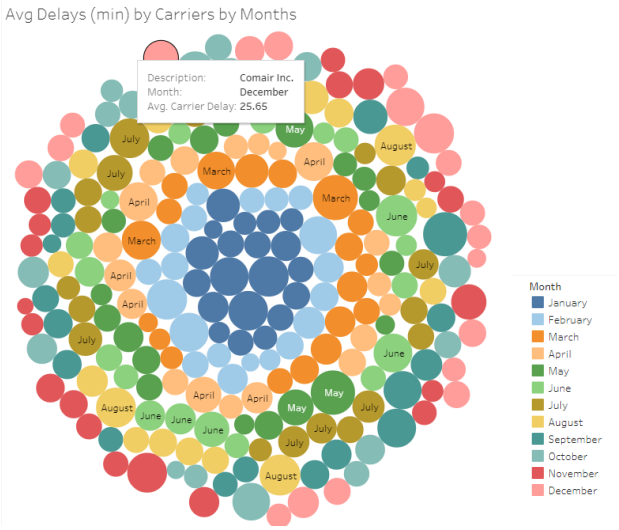
1. Dashboard 1 (dsb_DelaysByCarriers):
    a. Here I'm trying to show the various causes of delays and in which months it is the highest. Here I thought a Line graph would simply convey what I needed to show.
    b. I also tried to show which airlines encountered the most average departure delays by plotting it on a Highlight Table. I chose a Highlight Table because I thought it would be a bit more intuitive and to better display the 3 variables:
        i. Months
        ii. Airline carriers
        iii. Average minutes delayed.
2. Dashboard 2 (dsb_Airport_Vs_Carrier):
    a. Here I wanted to see, per airport, which Carriers has the highest average departure delays. I chose the Treemap for the airlines in hopes to convey by the block size and colors the Average departure delays.
    b. I also chose the Map to represent the airports more geographically, rather than a generic list of airport names, such as bar graph.
3. I put the dashboards into a Story to summarize and highlight some of my findings.

Feedback:

Version 2 (12/29/2018):

Patrick B.'s feedback on 1b above: Because the Highlight Table I used requires viewers to scroll horizontally to see all Carriers, he suggested that I use a Packed Bubbles graph.

1. Upon converting this to a Packed Bubbles it became very messy:

a. 

b. I then understood that someone could be viewing it on different monitor sizes, tablets, laptops, etc, and the need to scroll may not be a good idea.
    i. I pivoted the rows and columns; thus, the list of airlines now grows vertically, and the months are on the horizontal axis.
    ii. I also limited this to TOP 10.

Resources: list any sources you consulted to create your visualization

- https://interworks.com/blog/tmcconnell/2015/02/10/case-statements-vs-if-statements-tableau/
- https://www.travelmiles101.com/list-of-major-airline-hubs/