BUILD DATA DASHBOARDS PROJECT

This project uses the Flight Delays and Cancellations data set provided by Udacity:
https://d17h27t6h515a5.cloudfront.net/topher/2017/December/5a3b1fad_flight-delays/flight-delays.zip

INSIGHT 1: FLIGHT CANCELLATIONS AND THEIR REASONS

- Link of the dashboard:

https://public.tableau.com/views/flight-cancellations/Dashboard1?:embed=y&:display_count=yes&publish=yes

- Summary: 4433 flights out of 274,964 were cancelled. This makes up for 1.61% of total flights. In an effort to decrease this number and increase customer satisfaction, I looked at the reasons for the cancellations, and the reason most flights were cancelled is extreme weather, followed by air carrier then national aviation system (NAS) faults.
- Design: Rather than use the original columns which would have been ambiguous to the user (0 and 1 for cancellation and A, B, C, D for cancellation reason), I opted to create two new columns in the dataset to keep the data-ink ratio high. I chose the horizontal bars to show the disparity between the two columns. I chose the bubbles because the categories were few and it lets the user see at a glance the most frequent reasons.
- Resources: This discussion on Kaggle (https://www.kaggle.com/usdot/flight-delays/discussion/35193) helped me find the metadata for the CANCELLATION_REASON column.

INSIGHT II: AIRLINE DELAYS

- Link of the dashboard:

https://public.tableau.com/views/airline-delay_0/Sheet1?:embed=y&:display_count=yes&publish=yes

- Summary: In 2015, Southwest Airlines Co. had the highest delay time at 182,670 minutes. On the other hand, Virgin America had the least delay time at 7,583 minutes.
- Design: I used treemaps to give the viewer two ways of reading the data: position and color. The viewer will first focus on the colors and can surmise some insights as the best and worst airlines in terms of delays, then they can use the position to get a better picture at the difference between each airline's delay. I also added a filter of months so the viewer can choose a specific month or a range of months and see which airlines had the most delays in that period.
- Resources: N/A

INSIGHT III: AIRPORTS WITH MOST CANCELLED FLIGHTS

- Link of the dashboard:

https://public.tableau.com/views/airport-cancellations/Dashboard1?:embed=y&:display_count=yes&publish=yes

- Summary: The three airports that have the most flight cancellations are: Chicago O'Hare International Airport, Dallas/Fort Worth International Airport and LaGuardia Airport, respectively. The map shows the airports with the twenty most flight cancellations, and it can be seen that the airports in the Eastern side of the country have more flight cancellations than their Western counterparts.
- Design: I chose the bar graph to sort the airports in terms of cancellations because it's the clearest chart to show this type of data. I chose that color for the bar graph because I felt it was a bit more interesting than the default blue. I first included the map to see if there was any relation between the location of the airports and the number of cancellations. The map felt too crowded so I limited it to 20 airports. I noticed that more than half of the twenty airports resided on the eastern side of the country, and since in the first insight we learned that the most frequent cause for cancellations is extreme weather, I thought this could be related since cities in the East coast tend to have a harsher weather. I chose the color-blind palette even though the number of countries is double the number of colors in the palette, to ensure that the viewer can clearly see the differences between the colors. I added the color legend so the viewer can quickly find the airport they're looking for.
- Resources: N/A