Name: Jonathan Lawrence

Date: 12/18/2019

Assignment: Project Task 1 - Dashboard

**What you did and why you made the decisions you did?**

For this dashboard, I considered several different supplemental datasets that I could use. Ultimately, I narrowed them down to two datasets. The elimination process considered factors like determining whether or not the data was in a relevant timeframe and whether or not it answered my hypothesis. After that, I performed some EDA to learn about the data, and from that I identified several formatting issues that I had to resolve before plotting. This included transforming variables into useable formats and eliminating blank variables.

**Why did you choose the visualizations you did?**

I chose a treemap to display the number of incidents per airline for three reasons. First, I felt that it was easier to look at a lot of squares than a lot of bars. Second, the treemap neatly organized the data in order of largest to smallest in common English reading order (top-down, and left-right). Third, and most importantly, the treemap shaded the airlines so that the darker ones were easily distinguished as the airlines with a larger number of incidents. I didn’t want to use a pie chart for this because the slices would have been too difficult to distinguish. This is where the color shading of the treemap really helped.

I decided to make a second plot for the number of incidents per airline. The treemap showed everything I needed to know, but it was harder to determine which airlines were on the lower end due to size constraints with the dashboard. I didn’t want to remove the treemap because the data was still very useful, but I decided to add a bar plot to better represent some of the lower-count airlines.

I chose a stacked barplot to represent the number of fatal accidents per airline because it showed the total number of fatal accidents, while also distinguishing between accidents that occurred between the years 1985 - 1999 and 2000 – 2014.

I chose a continuous line plot to represent the number of airline passengers per month per year because it outlined a clear increase/decrease between timeframes.

I chose to use a pie chart to show the comparison between the number of incidents from 1985 - 1999 and 2000 – 2014. A pie chart was useful here because it is easily apparent which time period had the most incidents.

Lastly, I made a packed bubbles plot to represent the sentiment analysis that was pulled from Twitter. I felt this best represented the sentiments because a bar plot looked too similar to the other bar plots when placed on the dashboard. In this manner, the sentiment analysis is much easier to spot on the page. This was purely a visual decision.

**How do you plan to present to your internal team?**

I plan to present this dashboard to my internal team. I’ll go through each graph and explain the findings.

**What were your findings?**

I noticed there was a trend in the number of airline passengers. Every year it seemed to rise from the previous year. Naturally, the more passengers that need to fly, requires more flights, which in turn means more likelihood of an airline incident occurring somewhere in the world.

I also noticed that the number of fatal accidents for each airline was lower in 2000 – 2014 than in 1985-1999.

I believe a lot of the negativity surrounding flying is due to media attention. A sentiment analysis taken from Twitter regarding airline tweets was overall largely negative. People’s reactions to airlines could correlate to negativity that gives off the impression that airlines are not safe to fly.

The terrorist attack on 9/11/2001 caused a dip in the number of passengers flying for the following few years after. The number of passengers slowly climbed back up within those next few years, but after the attack, it reached the lowest point it had been at in several years. It is highly likely that the public reaction to this attack hurt the credibility of airlines, causing people to believe that flying was no longer safe.

There is a significantly larger number of incidents for Aeroflot than any other airline. This suggests that a good portion of the incidents reported might be a result of the handling or ownership of that airline rather than a safety concern for flying in general.