

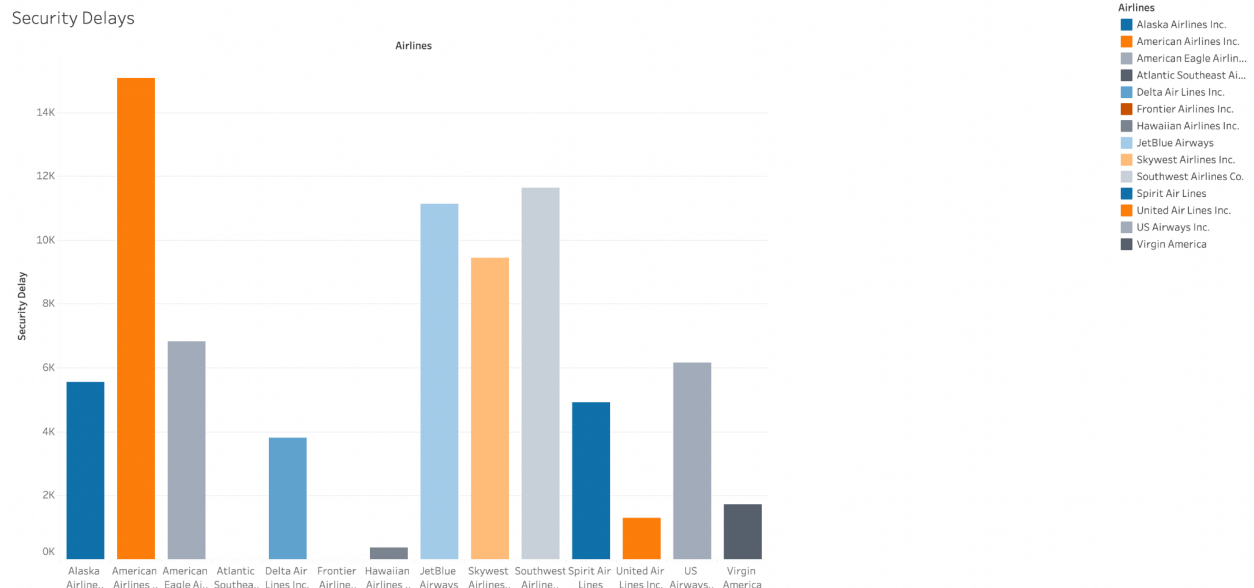
Data Visualization

With our 2015 Flight Data for the United States, let's investigate what causes delays. In order to understand what causes delays the first step is to look at our csv files provided by Kaggle. After exploration of the three csv files, there were several categories of delays ranging from weather to security that were noticed. Security is a very reasonable reason for a flight to be delayed or canceled. In order to get a picture of what is happening in regards to security delays, the first visualization that was created in order to compare the number of security delays with the different airlines in our dataset. A bar graph was chosen to represent this data because it clearly depicts the differences between the various categories which are the different airlines. The design choices made in this graph were firstly to have a legend to represent each color as a different airline, secondly to choose a color palette which was color blind friendly, and thirdly to ensure there was enough contrast between our foreground and background. The color palette used is called 'Color Blind', a palette specifically designed in Tableau for those who struggle with color blindness.

Our first insight for this sheet was that American Airlines had the most delays when it came to security threats, specifically, 15,084. While Atlantic Southeast and Frontier had 0 security related delays. Thanks to the first visualization, we now have a clearer picture of one specific type of delay. But what about other types of delays? The most common ones being departure and arrival delays.

Insight One:

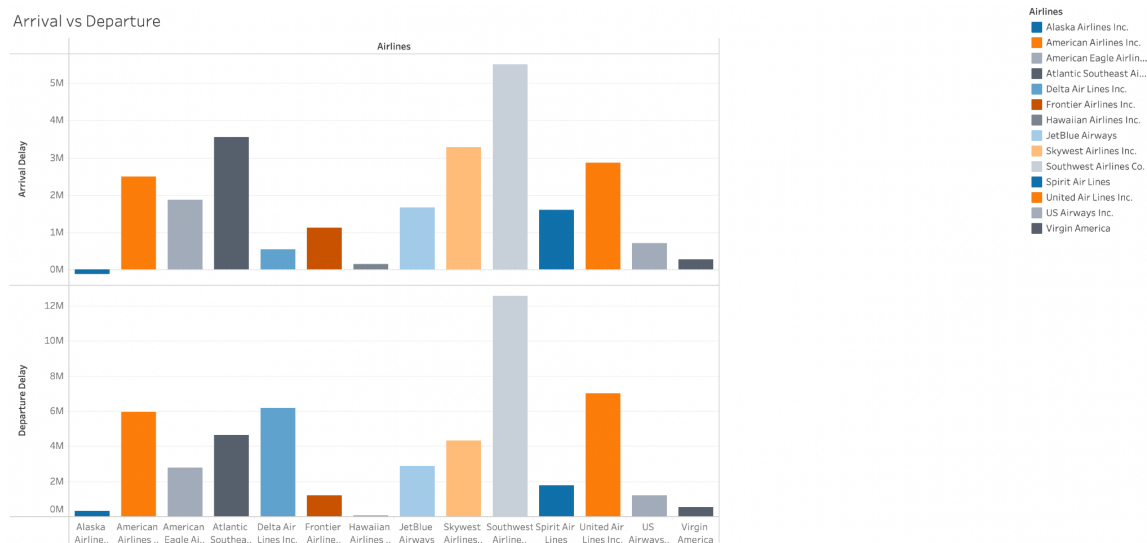
https://public.tableau.com/views/2015AirlineData-UdacitySubmission-InsightOne/SecurityDelays?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link



For our second visualization, we explored whether or not certain airlines had greater or fewer arrival and departure delays. Because we are working with different categories of airlines, the best choice was a bar graph once again. The color palette, foreground and background, and key were also implemented with the same design elements as our first sheet. The reason being, the color palette is color blind friendly, our background doesn't distract us from interpreting the graph, and our key gives us a clear idea of which airline we are referring to in each section of our chart. Following an inspection of the graph, the most interesting insight was that Alaska Airlines actually had a large number of flights arriving early compared to its competitors. This insight was drawn from the negative numbers shown in our graph. Specifically, 124,271 flights arrived earlier than expected rather than experiencing a delay in arrival time. While Southwest Airlines experienced more than 5.5 million arrival delays. Additionally, Southwest Airlines had the greatest number of departures delays compared to its competitors as well with more than 12.5 million departure delays. Now we have an understanding of which airlines struggle with maintaining their arrival and departure schedules but what influences these various delays? Our last visualization is a dashboard which ties our insights together making sense of why airlines experience delays and confirms a hypothesis that was formed from exploring our dataset. After seeing the first two visualizations, the question, does time of year have an effect on delays came up. After establishing a hypothesis that more than likely, yes depending on what time it is during a year there is an effect on whether airlines experience delays.

Insight Two:

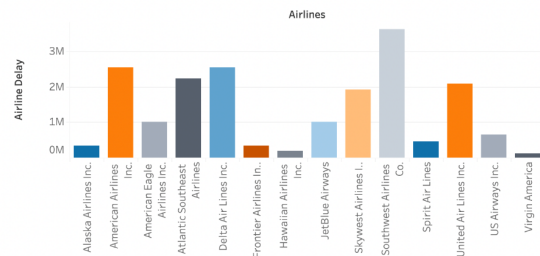
https://public.tableau.com/views/2015AirlineData-UdacitySubmission-InsightTwo_16632728599580/DeparturevArrival?:language=en-US&:display_count=n&:origin=viz_share_link



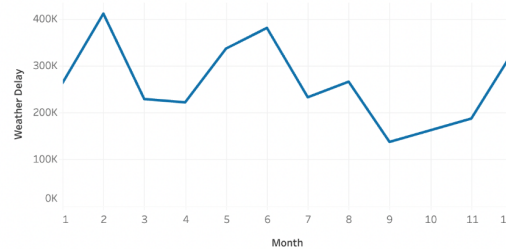
Insight Three (Dashboard):

https://public.tableau.com/shared/FZ5Z35KDQ?:display_count=n&:origin=viz_share_link

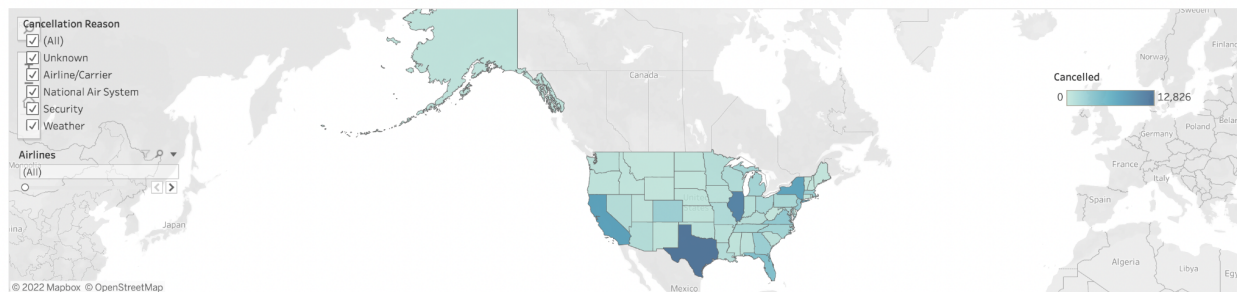
Delays by Airline



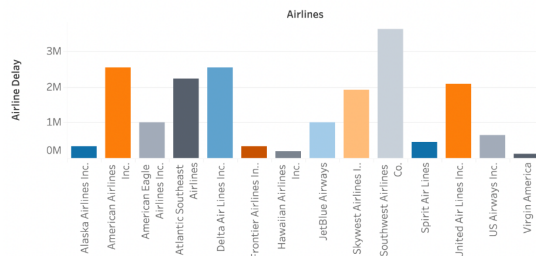
Weather Delays by Month



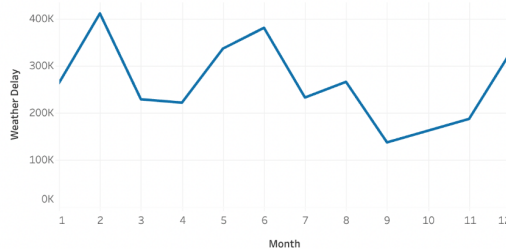
Cancellations by State and Airline



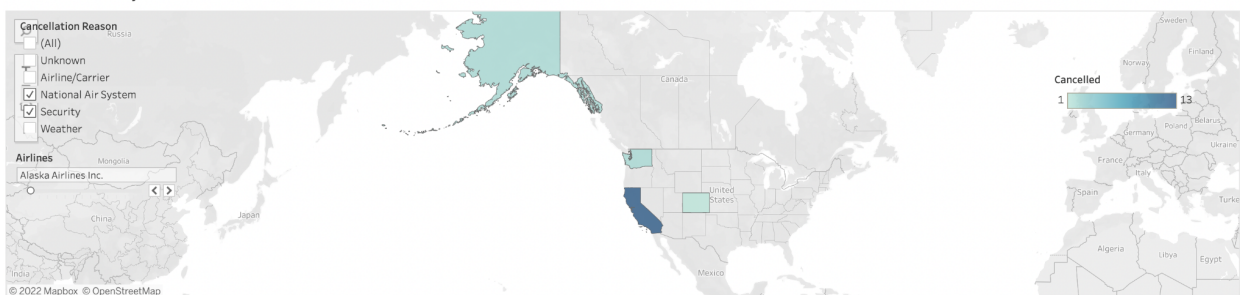
Delays by Airline



Weather Delays by Month



Cancellations by State and Airline



Our final visualization explored the concept of why delays occurred through three different visualizations in one dashboard and confirms our hypothesis. The first visualization contains the sum of airline delays and with different airlines in our dataset. The design choices were the

same as the first two visualizations in sheet one and sheet two. We want something that is color blind friendly, not distracting, and is easy to follow. The bar graph accomplishes all of those requirements. Our second visualization in the dashboard is a line graph which represents each month and helps us understand what time of year delays occur most often in terms of weather. This is the visualization which helped confirm our question the most. It seems from our visualization that the more extreme the weather is during certain months, the greater the chance of delay. Since our data is based in the United States, the colder months like January and December showed a spike in our delays. During the warmer months like June and August, there were spikes in the number of delays as well. Because the purpose of this visualization was showing our flight data over time, a line graph worked well. The design choices with this line graph carry the same sentiments as our other graphs, the blue color is not harsh and easy to follow, our background is not conflicting with our foreground, and our data is represented with 1 - 12 representing each month in a year. Originally, the data naturally graphs with 0 - 13 through Tableau, however, this doesn't make sense and those sections were empty and unnecessary. Lastly, our map of the United States gives us a representation of cancellations according to various reasons which we can filter like Airline, Weather, Security and so on. Additionally, our map can filter by airline giving us a picture of delays and cancellations specific parameters. The design choices for this map were to give a simple color palette ranging from different hues of blue. As the number of cancellations increases the color becomes darker. The filters are placed on the left and the color key is placed on the right in order to show their distinct functions to the user. Our main insight from this dashboard as stated earlier is that time of year affects delays. This insight brings greater clarity to our previous two insights as well.