Project Summary

Summary:
This project looks at a variety of data in order to identify the data that may be correlated to overall airline delays. By comparing airline delays to arrival delays, departure delays, specific states, airports, and months in the year, as well as air time and travel distance, the project works to identify any characteristics that appear to have a strong relationship with airline delay within this data set. After analyzing the data, it becomes clear that a variety of factors appear to be correlated with airline delay, including arrival and departure delays, longer distances, and greater airtime. However, the factors with the largest influence on higher than average airline delays seem to be the month of the year and the location of flights. Although this insight cannot be explicitly drawn from the available data, it appears that the human factor (perhaps surrounding holidays and travel as well as location) may be the greatest influence on larger than average airline delays.

Design Choices:
The choices for this design were primarily made to keep things simple and easy to use. When available, the Tableau palette specifically designed to be colorblind friendly has been utilized. When that palette was not an available choice, the colors chosen are oranges and blues, which are colorblind-friendly choices. Values that span a range of numbers are encoded in a range of colors, specifically orange:blue diverging. In order to make numerical differences more obvious in state maps, the colors were stepped in five steps, with the colors reversed so that higher numbers (in terms of delay minutes) were orange, while low delay numbers were blue. This is because people are generally familiar with colors closer to red being the less-desirable figures. Axis titles were made larger and rendered in bold for better legibility, while chart titles were removed from dashboards to reduce clutter. The most user-friendly design for month and airport filters seemed to be single-value dropdown menus, while utilizing a highlighter for states wound up being the most user-friendly tool.

Project Link :

https://github.com/22pallavi/flightdelay

Resources:
flights.csv

Udacity Data Foundations (Business Analytics) Nanodegree Program: https://www.udacity.com/course/business-analytics-nanodegree--nd098

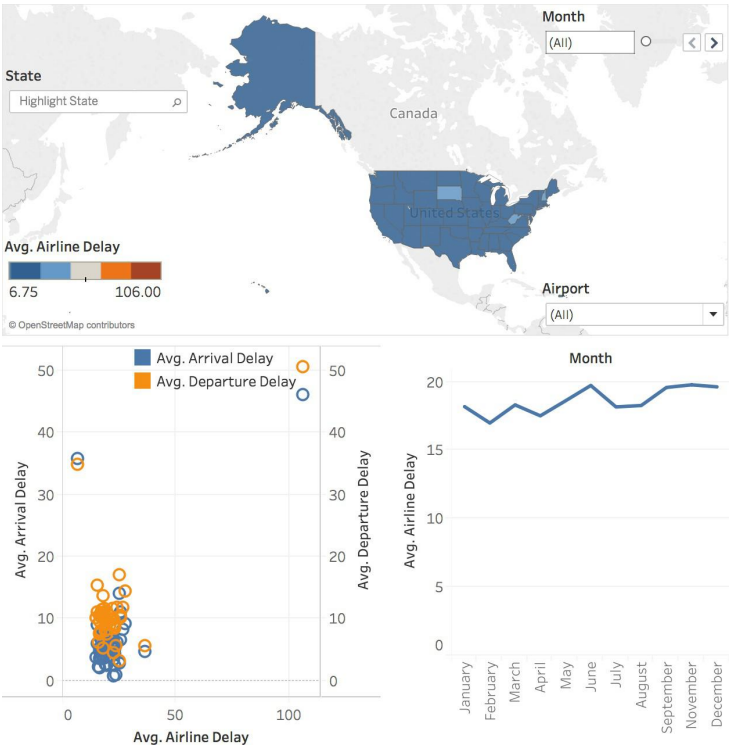
Video Walkthrough: https://www.youtube.com/watch?v=9xqHA732LMA&feature=youtu.be

Kaggle Flight Delay: https://www.kaggle.com/usdot/flight-delays/data

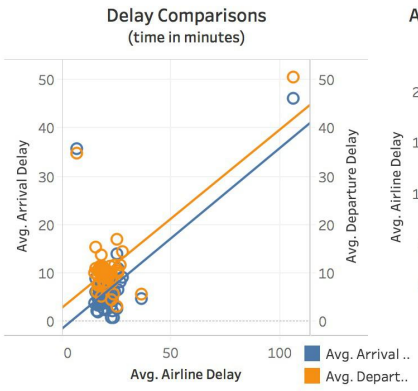
Kaggle flight delay discussion: https://www.kaggle.com/usdot/flight-delays/discussion/29308

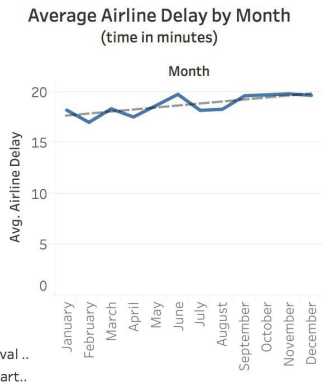
Tableau

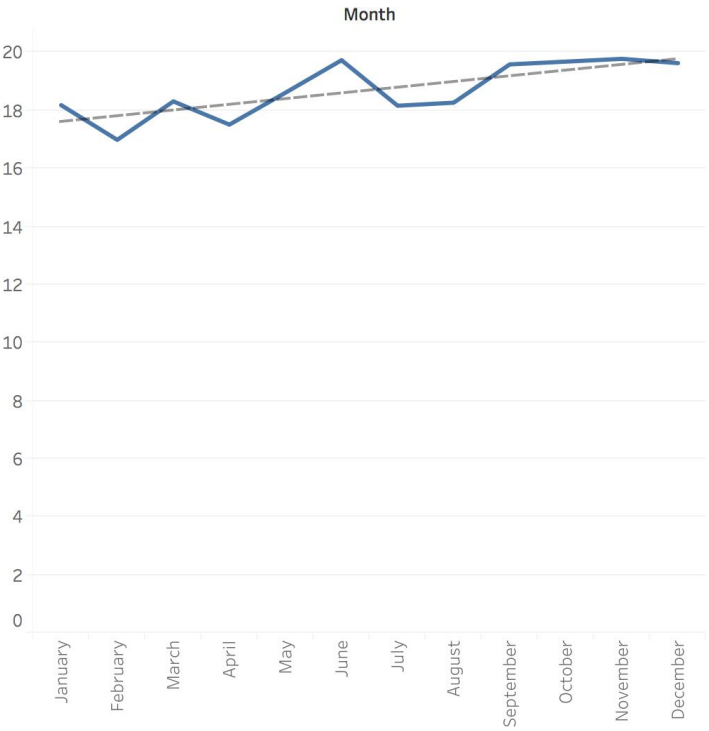
Flight Data Analysis



Project available on

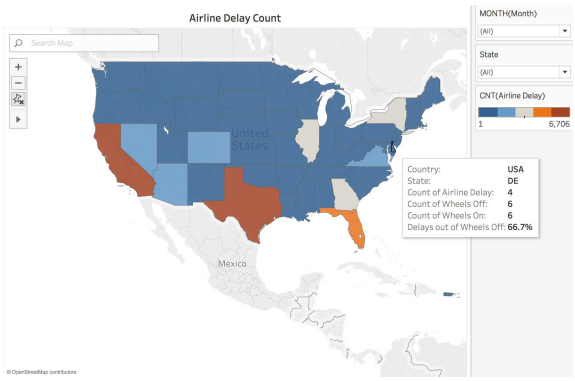


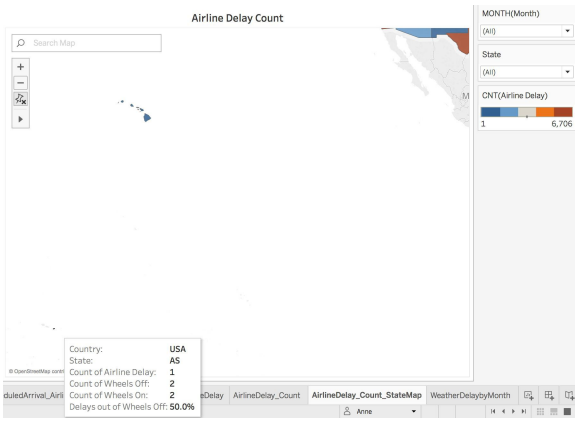


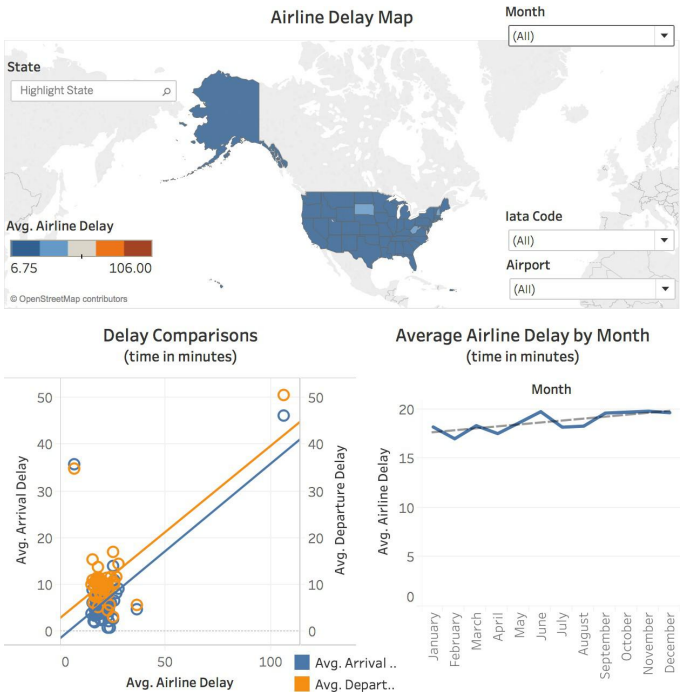


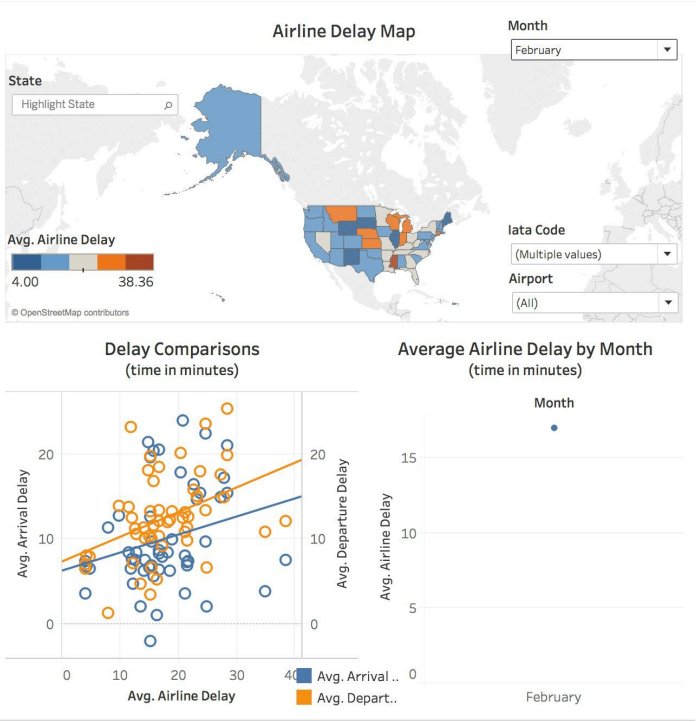
Average airline delay across all airports. (Time shown in minutes)

The dashboard on the left examines the relationship between airline delay and arrival delay as well as airline delay and departure delay across all airports. There appears to be a strong relationship between low arrival delays, departure delays, and airline delays, with a few outliers, specifically Delaware and Pago Pago. (However, it should be noted that there is only a count of four airline delays out of a count of six wheels-off (66.7%) reported in Delaware, and only one delay of the two wheels-off reported in Pago Pago (50%): see maps below.) Looking at the months in this year, while there is not a huge range of delays throughout the year, there are certain months that show higher and lower airline delays throughout the year, with peaks in June, November, and December. There is a clearly increasing trend of average airline delays as the year progresses, although the range of average delays is relatively small.

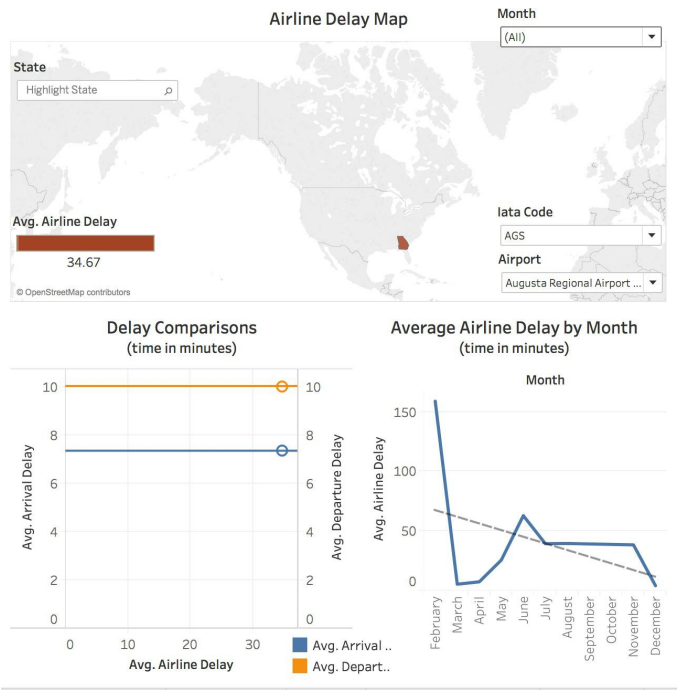

Airline Delay Count
Search Map
United States
Mexico
Country: USA
State: DE
Count of Airline Delay: 4
Count of Wheels Off: 6
Count of Wheels On: 6
Delays out of Wheels Off: 66.7%
MONTH(Month)
(All)
State
(All)
CNT(Airline Delay)
1
6,706


Airline Delay Count
Search Map
Country: USA
State: AS
Count of Airline Delay: 1
Count of Wheels Off: 2
Count of Wheels On: 2
Delays out of Wheels Off: 50.0%
MONTH(Month)
(All)
State
(All)
CNT(Airline Delay)
1
6,706
AirlineDelay_Count
AirlineDelay_Count_StateMap
WeatherDelaybyMonth
Anne

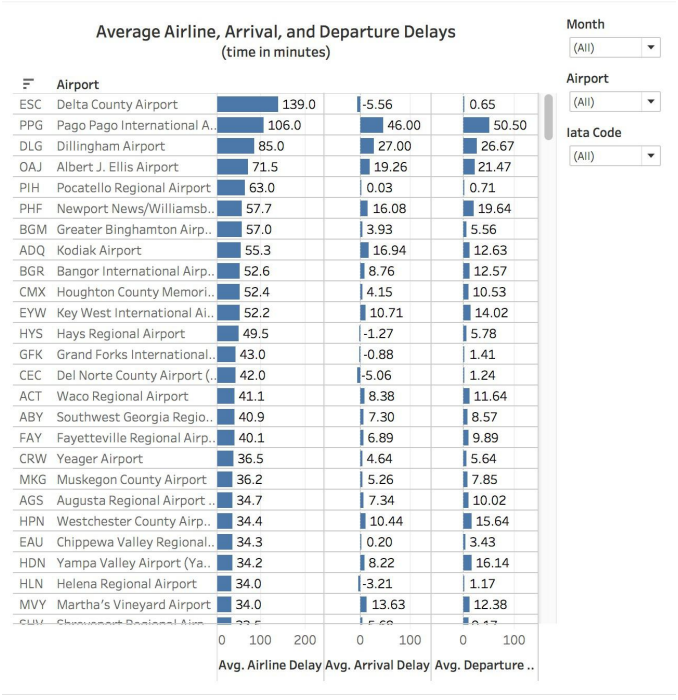


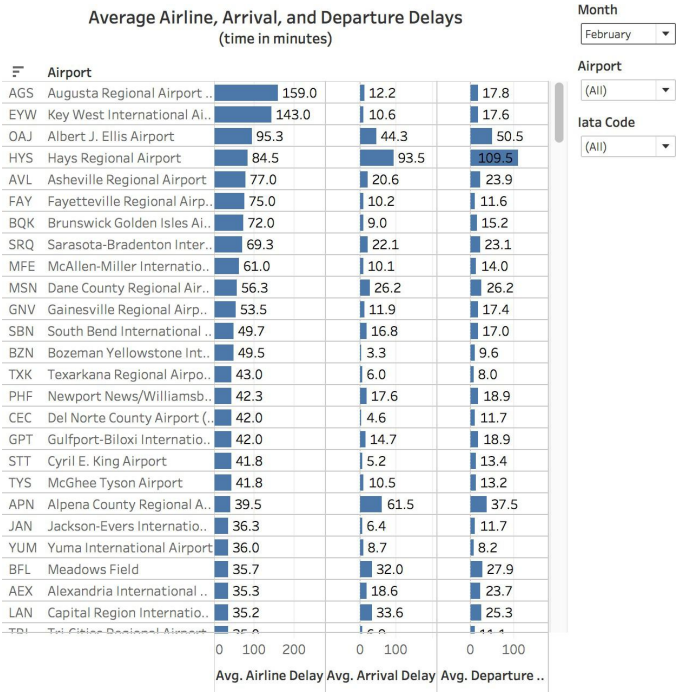


Average airline delay plotted against average arrival delay and average departure delay. (Time shown in minutes.)



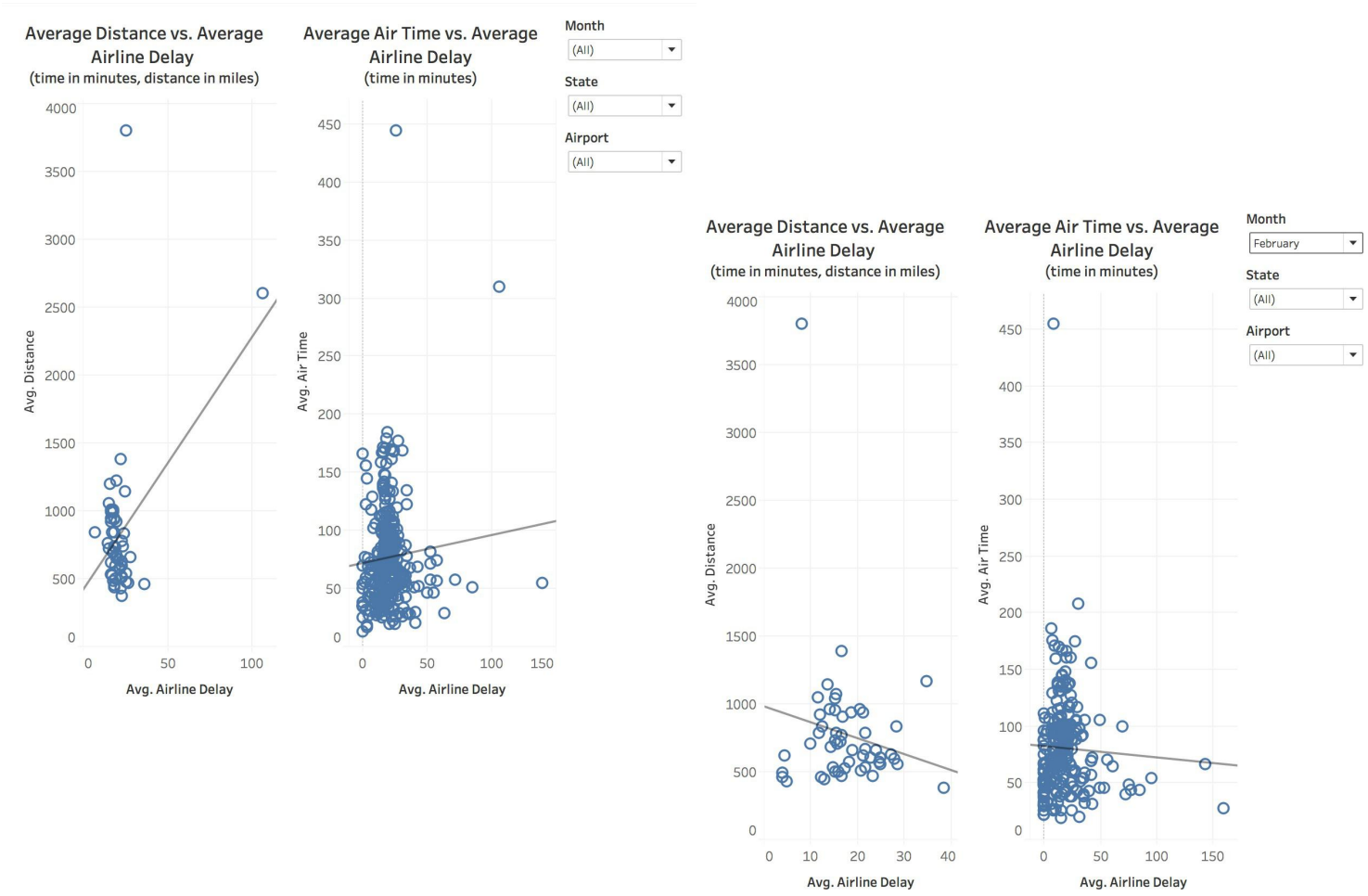
Looking more closely at the delays by month, it becomes clear that, although the appearance of a strong correlation between arrival, departure, and airline delays remains, the points become more spread out as the number of delays increases. The chart on the top right looks specifically at delays in the month of February, and it becomes clear that there are a number of states in this month with much higher than average delays as compared to the delays by state for the year as a whole. Looking specifically at Augusta Regional Airport (bottom left), February appears to be the month in which the number of airline, arrival, and departure delays are at an all-time high, followed by a drastic reduction in delays in March, with another increase in June, after which delays at Augusta decline.





Average airline delay compared to average arrival and departure delays in specific airports. (Time shown in minutes.)

This dashboard looks at the average airline, arrival, and departure delays in 2015 by airport. Across the year, Delta County Airport and Pago Pago International airport have the highest average airline delays, but it appears that while Pago Pago does have both higher than average arrival and departure delays, Delta County Airport has the highest airline delays but actually has low departure delays and negative arrival delays (which means that they arrived, as an average, earlier than expected). In the month of February (top right), Augusta Regional Airport has the highest average airline delays, followed by Key West International Airport. The chart in the bottom corner looks only at Augusta Regional Airport, which has the highest average airline delays in the month of February.



Average airline delay as compared to average distance and airline delay compared to average air time. (Distance shown in miles. Time shown in minutes)

This dashboard examines the relationship between airline delay and average distance, and also airline delay and average air time. Though there are outliers, it does appear that a low average air time is often related to a low average airline delay and a low average distance appears related to a low average airline delay. Looking at the month of February, there is a slightly greater spread in average airline delay, even while average distance remains low, but with fewer outliers. Looking specifically at Augusta Regional Airport in the month of February, the average distance is relatively low, but the average airline delay is very high. Average airtime is also extremely low while average airline delay is high.