Jason Grenig
Udacity DFND Project 4 - Tableau Flight Data

Insight 1 - Departure Delay by Time of Day
Link:
https://public.tableau.com/profile/jason.grenig#!/vizhome/DepartureDelaybyTimeofDay/Dashboard1?publish=yes

Summary:
Scheduled departures have a relatively consistent distribution between 6:00am and 7:00pm, tapering off outside of those hours. This contrasts the average departure delays, which increase as the day moves on. Flights begin the day on-time, then delays gradually increase to around 50 minutes by 10:00pm. This is logical as delays compound over the course of the day and then reset overnight where significantly less flights occur.
I chose blue colors for the map to not affect colorblind individuals, since there are a lot of small objects to view. The purple of the bar & line charts is a split-complementary color.

Insight 2 - Most Departure Delays by Airline

Link:
https://public.tableau.com/profile/jason.grenig#!/vizhome/MostDepartureDelaysbyAirline/MostDepartureDelaysbyAirline?publish=yes

Summary:
I examined which airline has the most delays via rank, count, and proportion.
The data story begins looking at a bump chart of airlines ranked by number of departure delays by month. We can see Each airline is ranked by number of delays each month. Note the impact of the merger btwn US Airways & American in June.
I chose the Tableau 20 color palette so that similar colors would not be sitting right next to each other.

The second slide compares 2 histograms, each counting the duration of departure delays across each airline. Controlling delays is a constant challenge for airlines, and we can see how each performs. The average delay lasts 15 minutes. Hawaiian is the fastest with 25% of departure delays lasting over 15 minutes. Among major carriers, Southwest, Delta, and United all beat the industry average (ranging from a 43-47% chance of a delayed departure lasting less than 15 minutes. Good luck if your flight gets delayed on Spirit. They rank last with 61% of departure delays taking longer than 15 minutes to resolve.
I kept the Tableau 20 color palette so that similar colors would not be sitting right next to each other in the bars.

The third slide tracks the proportion of flights delayed for airlines, and we can filter by airport. United is the only airline to have over half of its flights delayed across all airports reported
I chose blue bars to stick with the Tableau 20 color scheme.

Insight 3 - Number of Airlines and Departures by Airport and State

Link:
https://public.tableau.com/profile/jason.grenig#!/vizhome/NumberofAirlinesandDeparturesbyAIrportandState/Dashboard1?publish=yes

Summary:
This dashboard shows by state, the number of airlines and airports operating there, as well as the number of scheduled domestic departures for the 2015 year.
By looking at the map, we can get a look at which states are the most/least served by airlines. California and New York are served by all 14 airlines, followed by Texas, Illinois, and Florida with 13 airlines each. However California is the most served, when looking at the 33,331 scheduled departures in 2015. New York compares with only 12,707. Even including Newark airport into the New York total adds roughly 5,300 departures, which is still less than Texas, Florida, and Illinois. Delaware, by contrast, only had 6 departures for the entire 2015 year, all serviced by Frontier Airlines.

We can also learn how large metro areas distribute their total number of domestic departures. For example, Atlanta has all flights in 1 major airport, whereas Chicago Midway handles about 20% of the departure volume for the Chicago metro area. Greater Los Angeles is the most fragmented, with 5 airports sharing roughly 14,400 departures in 2015.
Again, I chose blue for the map to not affect colorblind individuals, since there are a lot of small objects to view. The histogram is in a peachy color to set it apart from the map.

Data Cleaning changes made:
Flights.csv:

- Changed file name to flights_cleaned1.csv
- Removed columns “Unnamed: 0” and “Unnamed: 0.1”
- Split columns “Departure Delay” and “Arrival Delay” into 2 separate columns each - “DEPARTURE_DELAY_EARLY”, “DEPARTURE_DELAY_LATE”, “ARRIVAL_DELAY_EARLY”, “ARRIVAL_DELAY_LATE”. This is to separate flights that actually left early from the late flights. Then I removed the “0” values from each column, so they wouldn’t get picked up in aggregations (since there are also blank values due to cancelled flights).

Resources:
Creating the bump chart-
https://www.analyticsvidhya.com/blog/2018/03/tableau-for-advanced-users-easy-expertise-in-data-visualisation/?utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+AnalyticsVidhya+%28Analytics+Vidhya%29