In this report we will explore the records of nearly a million flights in the year of 2015. This report aims to determine how many flights are late, which airports are responsible for the most late flights, and which airlines are the latest. Additionally we will investigate how departure and arrival delays relate, flight punctuality by state, and the distribution of late and non-late flights over the months of the year. Finally,we will discuss the slicers and drill-down capabilities of the digital report.
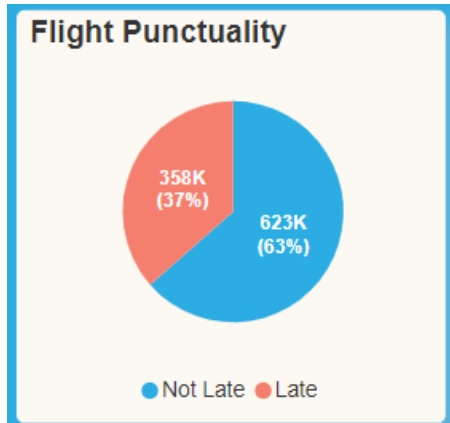

Figure 1: Pie chart of flight punctuality in 2015

According to our data set, 37% of all flights, in the year of 2015, had late arrivals. As can be seen in Figure 1, this accounts for 358 thousand flights with 63% (623 thousand flights) being on-time or earlier. In the digital report, this visualization has drill down capabilities, allowing the user to filter the entire report for either late or non-late flight data.

In Figure 2, we see the number of late (and non-late) flights out of each airport. Displayed are the top 9 in this metric; others can be seen in the digital report through use of the scroll bar. From this, we can see that the airport responsible for most late flights is the Hartsfield-Jackson Atlanta International Airport with almost 20 thousand late flights. If new strategies to reduce late flights are developed, implementing them here could be the most effective in reducing the overall number of late flights.

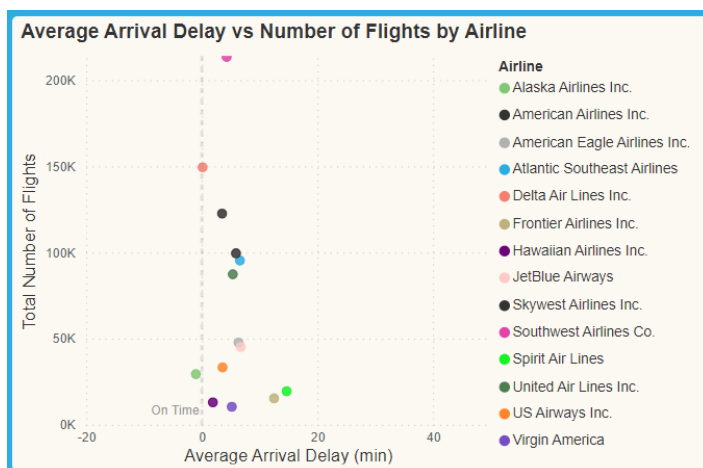Figure 2: Stacked bar chart of the number of late flights by airport.

Figure 3: Scatterplot of the average arrival delay versus the total number of flights by airline

Now, we look at the airlines. Figure 3 shows the average arrival delay for flights with each airline plotted against the *total* number of flights with that airline. Looking at this scatter plot, we can see that Spirit Airlines is the latest airline with an average delay of around 14½ minutes (as can be revealed in the digital report through the tooltip).
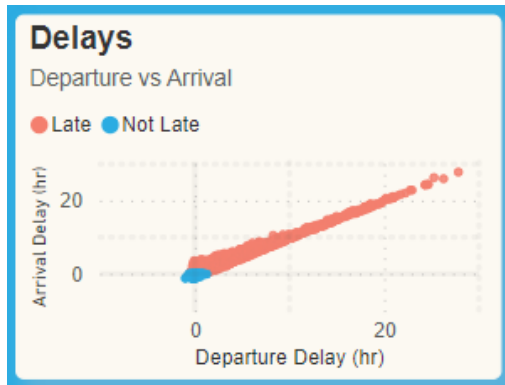
Figure 4: Scatterplot of departure delay vs arrival delay

Figure 5: Bar chart of the percent of flights in each state that were late

Supplementary to our visualizations so far are a few additional graphs giving further insight into our data set. In Figure 4, we can see a clear correlation between departure delays and arrival delays. This is not unexpected but it does reveal that the trend is nearly one to one. That is, the arrival delay is on average approximately the same as the departure delay. From this, we can hypothesize that the responsibility for arrival delays primarily rests on the departure end of a flight, rather than the arrival. Consequently, this report attributes "ownership" of each flight record to the airport of origin. More investigation into this phenomenon is advised.
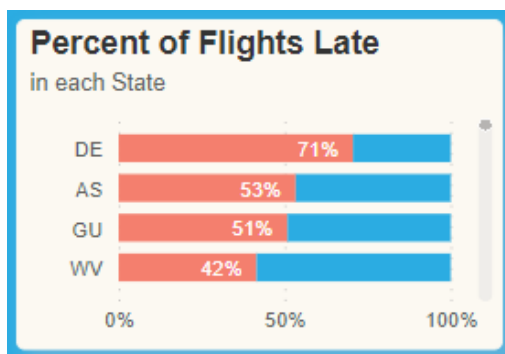
In Figures 5 and 6, respectively, we have the percent of flights that were late in each state/territory (complete with scroll bar) and the punctuality and number of flights by date. These, combined with the drill down capability of the flights by date line chart–which allows the user to display data for each individual month–reveal an apparent hole in the data. When drilling down into the month of October, the bar chart in Figure 5 depopulates (barring a single category dubbed "(Blank)". While this may, at first, look like no one in the country flew for the entire month of October, a little stroll into the data set, itself, reveals that the airport data was input incorrectly for

the month of October. Therefore, while we have the records for the flights that occurred in October, we do not have state/territory information until it can either be imputed or otherwise recovered. Such efforts would be advised should the need for this information arise.
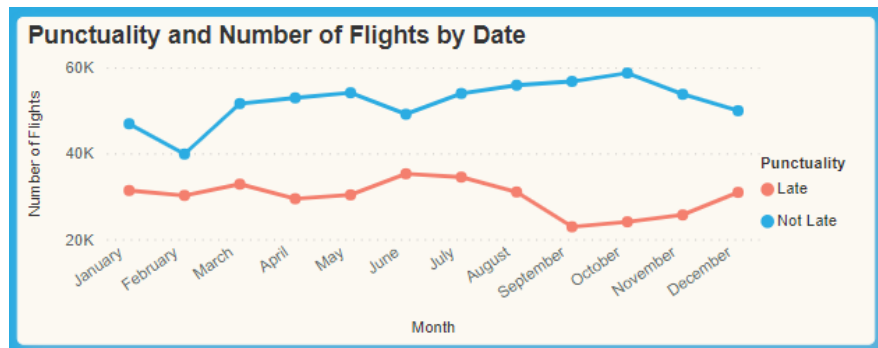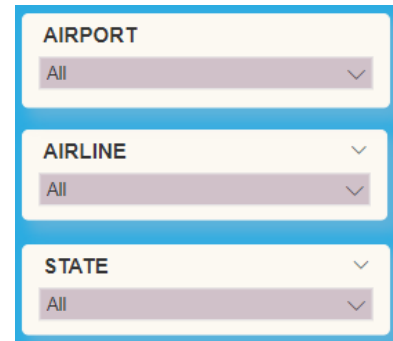

Figure 6: Line chart of the number of late and non-late flights in 2015 by date

Looking at the overall data in Figure 6, we see a distinct rise in late flights during the winter and summer months. This may be a result of increased demand during the associated holidays. Further analysis is necessary to determine how this issue can be assuaged.

Finally, in the bottom left of the digital report, we have slicers, as seen in Figure 7, which allow the user to filter the data in the report by airport, airline, and/or state by choosing options from the associated drop-down menu. This is useful for isolating specific airports, airlines, or states at which the user may want a closer look. Worth noting is the distinct lack of October data that appears in the flights by date line chart when slicing on any of the states or territories.

Figure 7: Slicers in the digital report for airport, airline, and state/territory

From this report, we found that 37% of the flights in 2015 were late. We determined that the airport responsible for the most late flights was the Hartsfield-Jackson Atlanta International Airport and the airline with the latest flights, on average, was Spirit Airlines. We also discovered an important hole in our data set, determined that more analysis on the late flights in the summer and winter months may be necessary, and discussed the function and usefulness of the slicers and drill-down feature provided in the digital report.