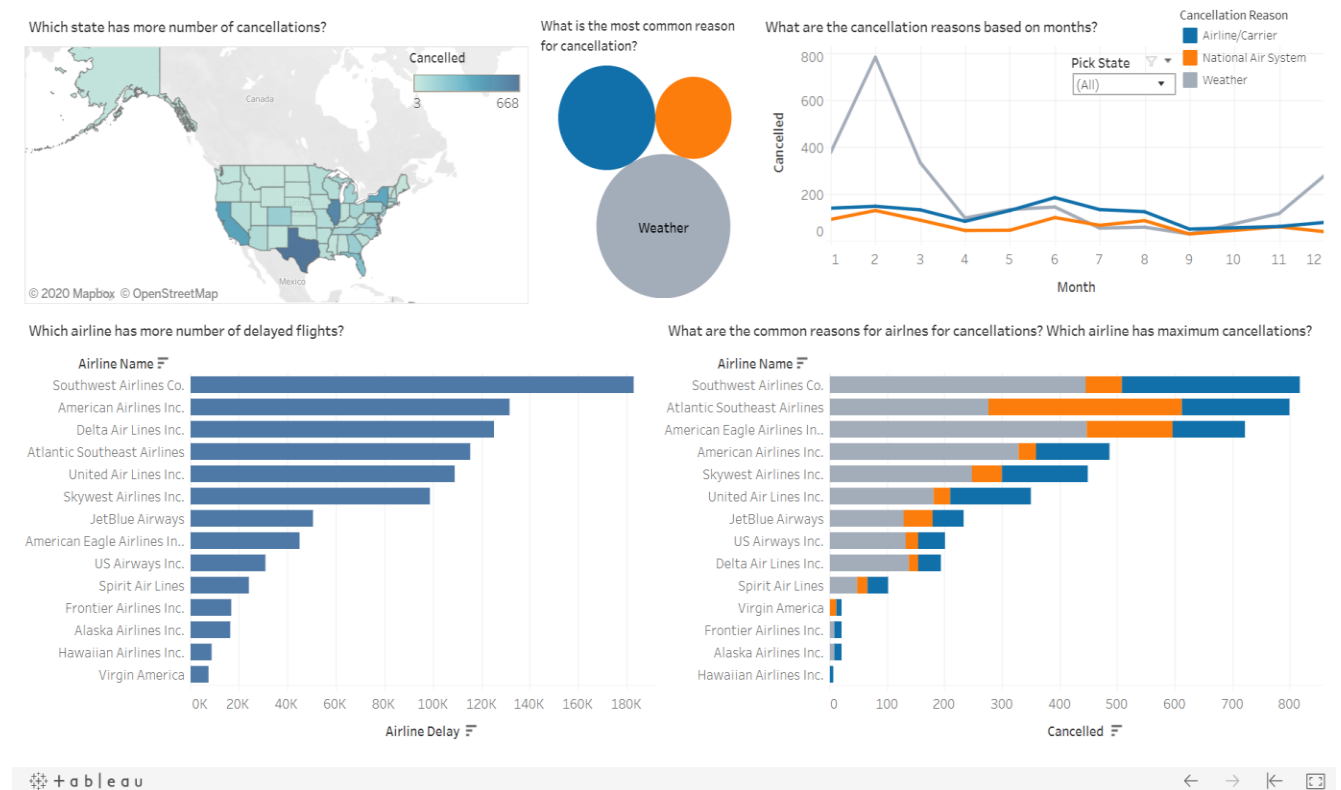


# Building data dashboard Tableau

## Dashboard for Kaggle's data: Flight Delays and Cancellations



Dashboard URL:

[https://public.tableau.com/profile/harsh.jain2653#!/vizhome/FlightDelaysandCancellationsDashboard\\_15930180145860/Dashboard1](https://public.tableau.com/profile/harsh.jain2653#!/vizhome/FlightDelaysandCancellationsDashboard_15930180145860/Dashboard1)

Data Source: Creation of the visualizations and dashboard needs Kaggle's dataset 'Flight Delays and Cancellations' which has three CSV files: Flights, Airplanes, and Airports.

**Dashboard Summary:** The dashboard incorporates five visuals, four of which are interactive. A user can dive deeper to see the state specific insights by picking a state from the dropdown filter named "Pick State". The "Color Blind" color palate is chosen for all the visuals and dashboard for better readability for all the viewers/users.

Before diving-in to create visuals for this dataset, we need to follow two steps:

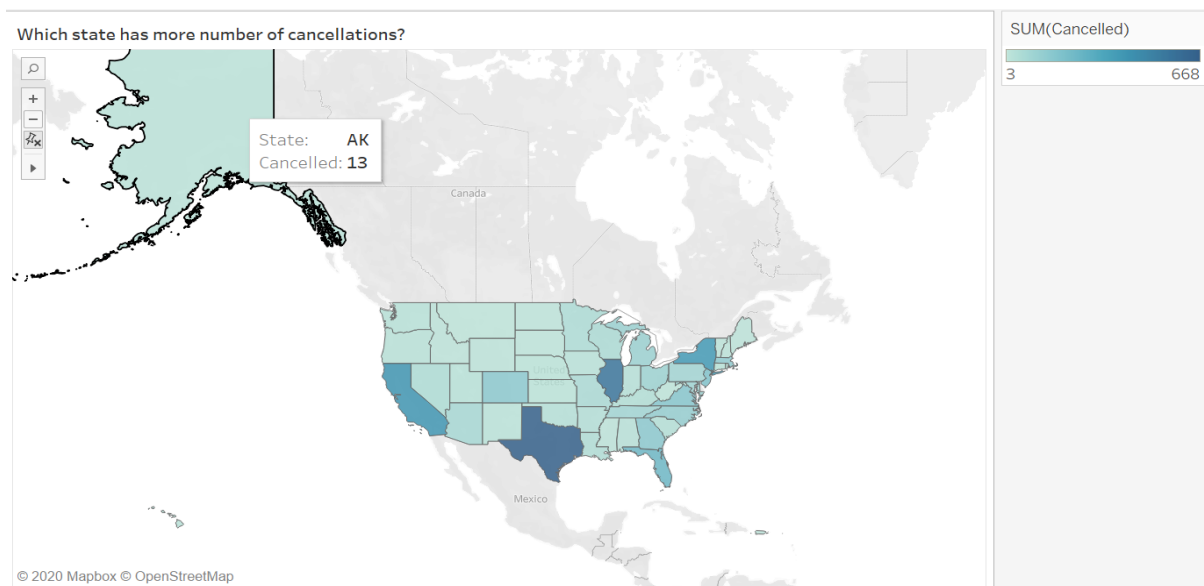
**Step 1:** To come up with final data set I used JOIN

1. Flights joined to Airplanes (join on IATA code with Airline)
2. Flights joined to Airports (join on IATA Code with Destination Airport)

**Step 2:** The dataset has “flight cancellation reason codes”, but we need to create a calculated field to show the reason and not the codes (like A, B, C as it is unclear). I renamed the dimension “Cancelled Reason” to “Cancellation Reason Codes” and created a new field called “Cancellation Reason Name” and used the following code:

```
IIF ([Cancellation Reason Code] = 'A' , 'Airline/Carrier',IIF ([Cancellation Reason Code] = 'B' ,  
'Weather',IIF ([Cancellation Reason Code] = 'C' , 'National Air System',IIF ([Cancellation Reason  
Code] = 'D' , 'Security', 'Other'))))))
```

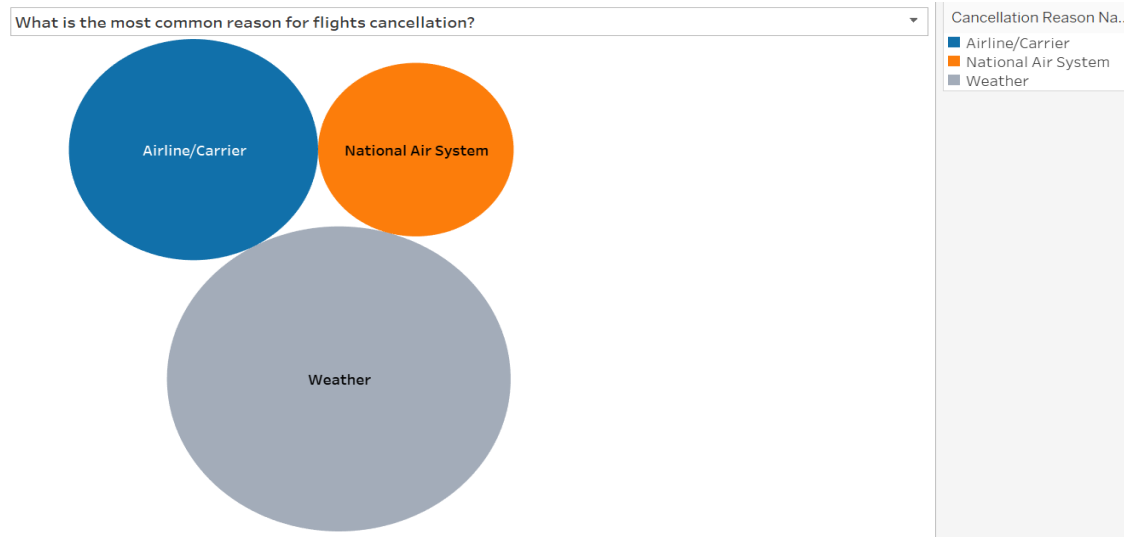
**Insight 1 :** Which state has more number of cancellations?



Description: Through this visualization, it can be concluded that the highest number of cancellations had been in Texas (#668), followed by Illinois (#563). I chose shaded map to visualize this data because with a shaded map we can easily compare cancellations in all the states simultaneously. If a state is colored in a dark shade it has more cancellations as compare to the state that is colored in a lighter shade. A legend has been added to the visual that indicates the color encoding in the map. The darker the region, the more the number of cancellations. Also, the tooltip shows the state code and the number of cancellations.

URL:[https://public.tableau.com/profile/harsh.jain2653#!/vizhome/FlightDelaysandCancellationDashboard\\_15930180145860/CancellationbyStates](https://public.tableau.com/profile/harsh.jain2653#!/vizhome/FlightDelaysandCancellationDashboard_15930180145860/CancellationbyStates)

## Insight 2: What is the most common reason for flights cancellation?

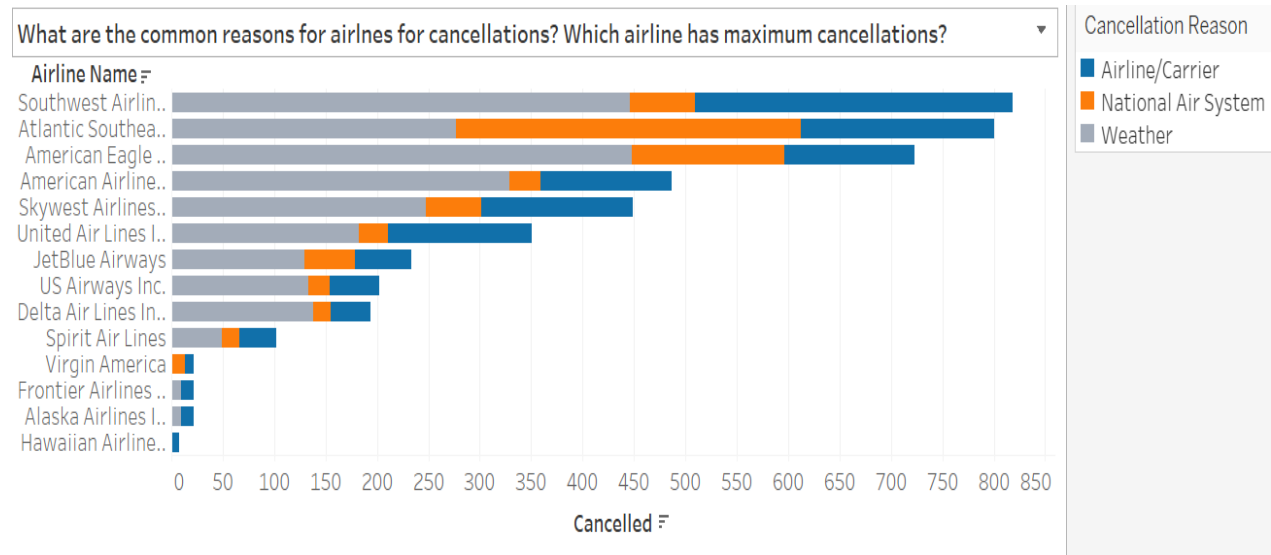


Description: This visual shows that the reason for most of cancellations is weather. The legend indicated the color encoding for the cancellation reasons, this will be useful in the next visuals. 'Color-Blind' color palate is chosen for color encoding for all the visuals to promote better readability for all the people (including the once who have color-blindness) as the colors are not too bright so they are not distracting. Bubble plot is chosen for this visualization as we can compare the major reason with the size of the bubble and compare which reason is the most common and least common by looking at the size of the visual.

Also, this plot is made interactive in the dashboard. So, we can pick a state from the drop-down and see the results for a specific state or we can choose to see the results for all the states. This can be achieved by picking up an option from the "Pick State" drop-down filter.

URL: [https://public.tableau.com/profile/harsh.jain2653#!/vizhome/FlightDelaysandCancellationReasonsDashboard\\_15930180145860/FlightsCancellationReasons](https://public.tableau.com/profile/harsh.jain2653#!/vizhome/FlightDelaysandCancellationReasonsDashboard_15930180145860/FlightsCancellationReasons)

**Insight 3:** What are the common reasons for airlines for cancellations? Which airline has maximum cancellations?



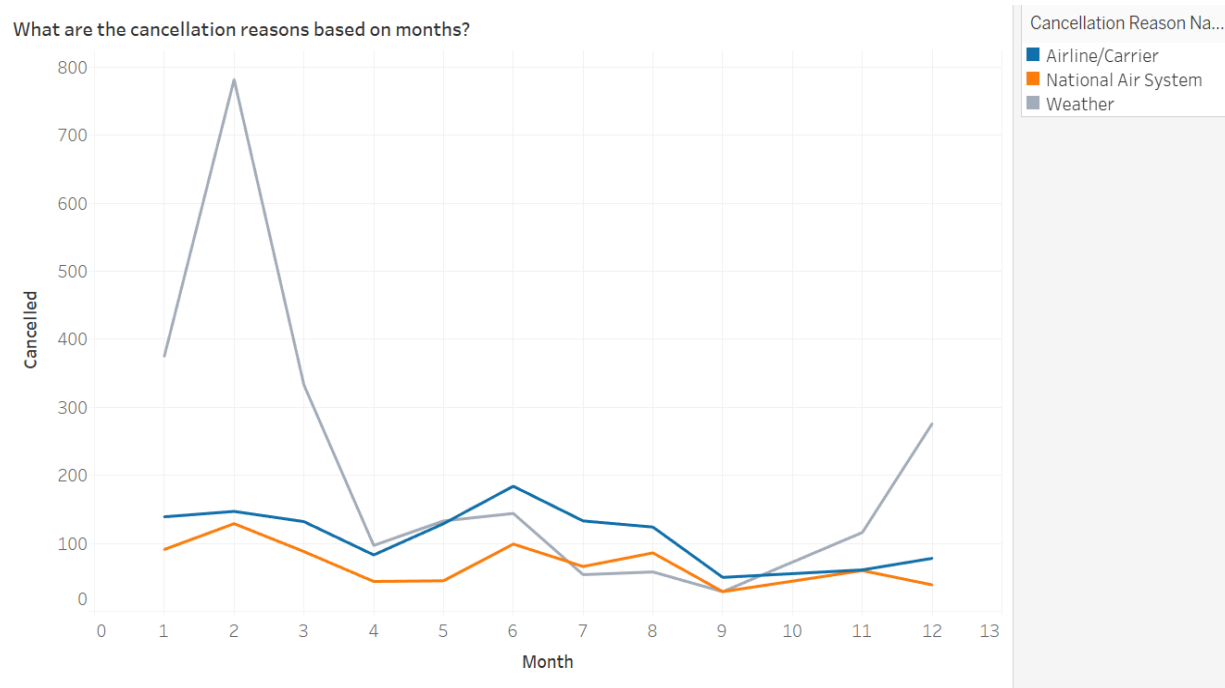
Description: From this visualization, we can determine that the airline that has the most number of cancellation is "Southwest Airline". Also, the major reason for cancellation is weather for all the airlines except "Virgin America" and "Hawaiian Airline".

I chose stacked bar chart for this visual it allows us to visualize parameters (2 dimensions and 1 measure) in one graph that allows us to get the insights we are looking for. This plot is also interactive in dashboard, we can pick a state from the drop-down and see the results for a specific state or we can choose to see the results for all the states. This can be achieved by picking up an option from the "Pick State" drop-down filter.

We can also use the legend and click on any reason and this visual will highlight the bars for selected reason.

URL: [https://public.tableau.com/profile/harsh.jain2653#!/vizhome/FlightDelaysandCancellationDashboard\\_15930180145860/CancellationReasonsbyAirlines](https://public.tableau.com/profile/harsh.jain2653#!/vizhome/FlightDelaysandCancellationDashboard_15930180145860/CancellationReasonsbyAirlines)

#### Insight 4: What are the cancellation reasons based on months?



Description: The cancellation reason seen in most of the months is 'Weather'. However, in the month of May 'Airline/Carrier' and 'Weather' both are equally main reasons for cancellations. Also, in the month of June, July, August, and September the major reason is 'Airline/Carrier'.

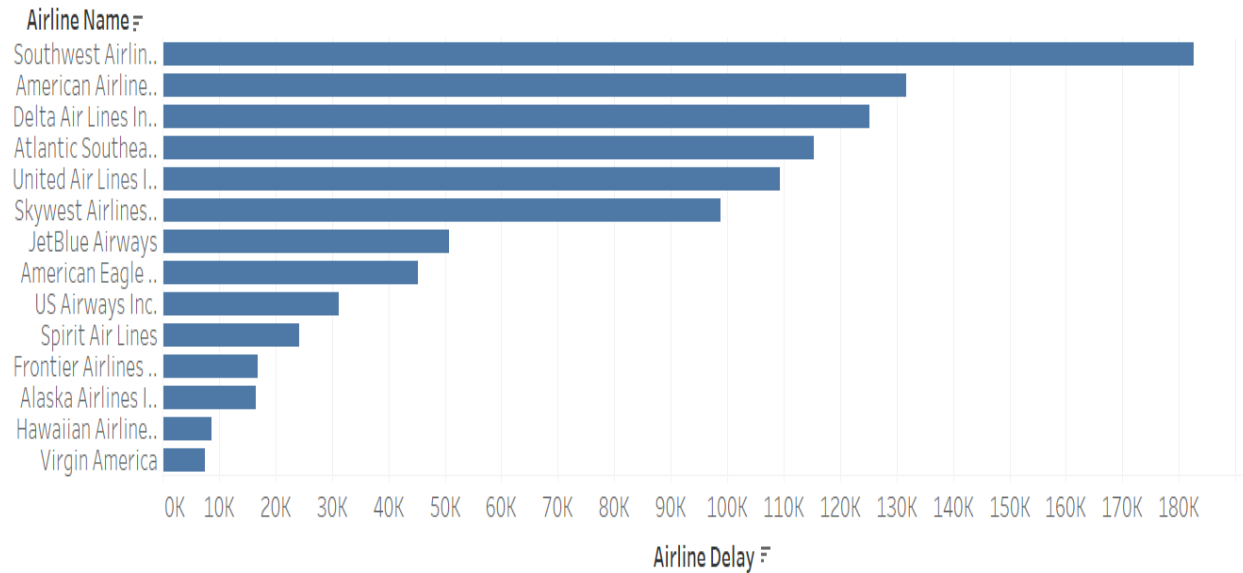
In the month of June and July, weather is the least common reason for cancellations as 'National Air System' takes it over. We also see a high spike in 'Weather' as a cancellation reason in the month of February.

Since we want to see the number of cancellation based on different reasons on different month a continuous line chart is ideal for this scenario as it allows us to compare the cancellation based on the reasons for every month in a single graph. The steepness in the curve shows the spike or fall in the number of cancellation between two months.

URL: [https://public.tableau.com/profile/harsh.jain2653#!/vizhome/FlightDelaysandCancellationDashboard\\_15930180145860/CancellationReasonsbyMonth](https://public.tableau.com/profile/harsh.jain2653#!/vizhome/FlightDelaysandCancellationDashboard_15930180145860/CancellationReasonsbyMonth)

### Insight 5: Which airline has highest number of delayed flights?

Which airline has more number of delayed flights?



Description: From the visual we can conclude that the airline that has the highest number of cases for flight delays is 'Southwest Airlines' followed by 'American Airlines'. I chose bar chart to show this insight as we have one categorical and one quantitative variable i.e. Airline Name and number of delays by airline. The plot is basic and has no color encoding for the bins as the airline name is already present in the Y-axis label.

Moreover, adding color encoding would be unnecessary effect in this scenario as it does not facilitate readability for the visual rather it would be distracting as it would decrease the data-ink ratio for the visual.

This visual is also made interactive in the dashboard to show the insights based on the selected state.

URL: [https://public.tableau.com/profile/harsh.jain2653#!/vizhome/FlightDelaysandCancellationsDashboard\\_15930180145860/FlightsDelaybyAirlines](https://public.tableau.com/profile/harsh.jain2653#!/vizhome/FlightDelaysandCancellationsDashboard_15930180145860/FlightsDelaybyAirlines)