FLIGHT DELAYS ANALYSIS REPORT

Link to visualizations: Flight Delays and Cancellations | Tableau Public

The purpose of this report is to present the findings of an analysis of flight delays, the causes and when those delays get worse. The findings were based on a comprehensive exploration of a flight dataset, and the insights were visually represented through various visualizations using the following questions below:

1. What Airlines Have the Worst Delays?

   To determine the airlines with the worst delays, a tree map visualization was utilized. Each rectangle of the treemap represents an airline, its corresponding delay value, and the percentage of delay it contributed. While colour shades were used to indicate the magnitude of delays across the airlines.

   The analysis revealed that Southwest Airlines Co., American Airlines Inc., and Delta Airlines Inc. had the highest delays, collectively accounting for approximately 46% of the total airline delays. On the other hand, Virgin America, Hawaiian Airlines Inc., and Alaska Airlines Inc. had the least delays among the airlines analyzed.

2. What Causes Delays?

   To illustrate the distribution of delays across these causes throughout the year, a stacked bar chart was employed. We chose a stacked bar chart to visualize months across the year and at the same time also see how each delay type is proportioned in each bar. With this each colour on a bar represents a type of delay and its length represents the level of delay.

   The analysis identified four primary causes of flight delays: Air system, late aircraft, security, and weather. The findings indicated that the impact of delay causes varied over time. Late

aircraft was identified as the major cause of delays, with June 2015 showing a peak in delays attributed to this factor. September, on the other hand, experienced relatively fewer delays. It is worth noting that these delay causes were consistent across all 14 airlines analyzed.

3. Late Aircraft Delays Overtime

   We were interested to dig in more into the Late aircraft as a cause of delay and seeing the behaviour throughout the year.

   To visualize this, we created a dual-axis chart showing a bar chart for the months and a line chart showing the percentage difference trend for the months of the year. We chose a dual-axis chart because we were interested in knowing more about the behaviour of late aircraft as a type of delay and how it differs by month across the year.

   It was interesting to see that the major dropped happened from August to September when we had approximately a -50.4% drop in delays, which later rose in December by approximately 80.1%

4. Relationship Between Airline Flight Distance and Airtime and Their Impact on Delays

   A scatter plot visualization was employed to depict this relationship, allowing for a clear understanding of how these metrics correlated. In this visual, each bubble represents an airline, in which we have 14 of them. The size of each bubble suggests the level of delay while the colour was used to create a gradient where deeper shades indicate a higher level of delay and light shades indicates a lower level of delay.

   The analysis explored the relationship between airline flight distance and airtime and its impact on delays. The findings revealed a positive correlation between airline flight distance

and airtime, indicating that longer distances resulted in higher flight times. Additionally, variations in delays were observed among the different airlines, with airlines operating shorter distances and having shorter airtimes experiencing fewer delays, while those operating longer distances and having longer airtimes tended to have more delays.

5. Why are flights being cancelled?

   Given that we have three reasons which explained why flights are being cancelled in our data, we decided to visualize these reasons using a pie chart, as we can see the proportion of each reason to the other. Here, each chord is a reason, and it's further encoded by colour for emphasis.

   We found that 2,397 flights were cancelled as a result of weather which constitutes approximately 54% of the total reasons, followed by Airline/Carrier at 29% and National Air System at 18%.

6. A view of flight operations over time

   We wanted to see how flight operations occur throughout the year. So to visualize this, we created a line chart. We chose a line chart because we were interested in the trends and patterns over the year.

   We observed that there were an average of 5,499 flights, with the highest on 6th July and the lowest on 26th October with 6,231 and 821 flights respectively.

7. How frequently were flights cancelled?

   To answer this question, we created a heatmap to help us see in the most visually effective way using colour divergence to signify where the frequencies are high as well as low.

This chart type allows us to see the different airline types and how cancellations vary by weekdays. Using this allows us to quickly identify cancellation patterns in the data as regards to where they are high versus where they are low.

Among the airlines, we saw that AA, EV, MQ, and WN have the highest frequencies of flight cancellations, which is common to the four airlines on Mondays. Of all airlines, HA seems to be the least flight-cancelling airline.

Design:

Throughout the report, consistent font types and sizes were used to ensure a cohesive visual experience. Given that flight delays are perceived as negative, the colour red was selected to communicate the varying degrees of delays among different airlines, utilizing a gradient-like approach from the highest to the lowest worst delays.

To enhance the audience's focus on the visualizations, gridlines, line dividers, unnecessary labels, and axis components were removed, allowing users to concentrate on the key insights conveyed through the visualizations.

Conclusion:

Based on the analysis conducted, several key findings have emerged regarding airline delays and cancellations. Southwest Airlines Co., American Airlines Inc., and Delta Airlines Inc. were identified as the airlines with the worst delays, accounting for approximately 46% of total delays. The primary causes of delays were late aircraft, air system issues, security concerns, and weather conditions, with late aircraft being the major contributor. The analysis also revealed a positive correlation between flight distance and airtime, indicating that longer distances resulted in higher flight times and, consequently, more delays. Additionally, flight cancellations were primarily caused by factor 'B,' accounting for approximately 54% of all cancellations.

Examining flight operations over time, it was observed that the average number of flights per day was 5,499, with the highest volume occurring on 6th June and the lowest on 26th October. Furthermore, certain airlines, such as AA, EV, MQ, and WN, experienced higher frequencies of flight cancellations, particularly on Mondays. In contrast, HA demonstrated a lower rate of flight cancellations compared to other airlines.

Having received feedback about the analysis/visualizations, changes were made accordingly to the:

- Cancellation Reasons: Here, we created a new calculated field to reflect the members of 'cancellation reasons' initially; A, B, C, to Airline/Carrier, Weather and National Air System, respectively. Doing this allowed us to relate, gain understanding, and interpret the data more easily.

  Furthermore, in order to create an excellent user experience for the users, we added a filter action to the pie chart where a user is able to click on a 'cancellation reason' and it's able to filter across the other surrounding chart(s).
- We added further notes explaining the reasons for choosing the different chart types we used in our visualizations and as well as what each element in the charts represent and their significance to our analysis.

In conclusion, this analysis highlights the airlines with the worst delays, the primary causes of delays, the relationship between flight distance and airtime, the reasons for flight cancellations, and the frequency of flight cancellations among different airlines. These insights provide valuable information for both airlines and passengers to better understand and address the challenges associated with delays and cancellations in the aviation industry.