Flight Delay Analysis

Summary

As a frequent traveller and former airline employee, I'm always interested in air travel. Delays are especially annoying and costly, so I've tried to learn ways to minimize or avoid them. Using data downloaded from RITA/BTS, I've attempted to isolate a useful metric for comparing departure delays across several lenses, use it to interrogate the cause of delays, and recommend some basic ways to avoid encountering them if possible.

> Flight Analysis v1

> Flight Analysis v2

Design

The intended audience for this set of visualizations is any traveller who wants to know more about flight delays. Based on that, I've tried to keep things simple, informative, and high level, so that a user of any sophistication level. As such, minimalism and appropriate use of bold color, along with clear explanations of each visual, were important to me.

I initially planned for 5 slides, but added a 6th after getting initial feedback.

- An introductory slide to set the stage for the overall analysis, and to explain the data used and how it was transformed. Since this was a text heavy slide without any data or presentation component, I opted for a moderately large font size and a large image related to the analysis topic. I also highlighted keywords within the text.
- A slide to help define some of the ways flight delays can be explored, and to show the metrics we'll be considering during the analysis. I used an interactive map of the top 100 US airports to show how different approaches to measurement can impact the data story, and to allow the user to see a geographic overview of the data. Color choices are designed for simplicity and immediate cultural understanding - red = poor performance and green = good performance. *Post Feedback: Color choices were reevaluated in order to include color blind users. Blue for good performance and yellow for poor performance are still easily understandable, but work for all three major types of color blindness. I've also adjusted my measurement standards to focus on Delays vs Expected Delays instead of delay rate. Delay*

rate was confusing to some users, and didn't allow for the apples to apples comparisons I wanted to achieve.

- A breakdown of the FAA standard flight delay categories, how they are defined and how much they influence the overall departure delay landscape. I used color here throughout to tie the visualization back to the explanation, and to draw attention to the important topics. *POst Feedback: I've gone from a pie chart to a standard bar chart. My peers suggested a pie chart is a bit déclassé for a data professional in 2022, and there are other ways to get the point across.*
- An examination on how airport size/flight volume impacts delays. Prior to my exploratory analysis of this data, I would have assumed that busier airports would be more impacted by delays. This wasn't the case, so I wanted to showcase that fact clearly in case other travelers face the same misapprehension. *Post Feedback: I've replaced the bar charts with a simpler to understand dumb-bell chart of the top 20 airports by volume. Some feedback found the bar charts too busy, and they were less representative of the new metrics.*
- A slide showing the two biggest factors I found in causing delay - time and carrier. The goal of this slide is to show how flight volumes and delay/cancelation rates change based on month and day of the week, along with showcasing the wide variability of delay among carriers. Bar charts in a dark color are used to show the background data of flight volume, while area charts in front of that showcase delay and cancellation risk. *Post Feedback: This page saw the biggest change between v1 and v2. Most people found the visualizations busy and confusing, which lead to going back to the drawing board. V2 eliminates much of the visual complexity, showcasing expected delays vs actual delays with a standard overlay bar chart. I also dropped the cancellation metric, as it wasn't particularly useful. "If the colored bar is higher than the grey bar, delays are worse" is a much easier visual connection to make.*
- *Post Feedback: I've added a final slide that allows a user to check any arbitrary US airport and create similar visuals to what was used in the previous slides. Some of the feedback I got wanted more of a conclusion and a concrete application of the findings from the analysis.*

Feedback

I sought feedback from two peers, a frequent travel partner (my wife), and my mother-in-law who is a visual artist. All feedback was verbal after a brief run through of the analysis and some hands on time. I've transcribed the most important criticism from my notes here.

Peer 1 (Brian, data analytics director)

- I don't think I'd use a pie chart if I could avoid it.
- The fifth slide is a little messy. Can you make it clearer?

- Some of the tooltips aren't doing much. I'd either turn them off or expand them to be more useful.

Peer 2 (Bisan, data scientist)

- Delay rate is probably not the best measurement to use
- Red/Green is bad for color blind users. Maybe look into another color scheme?

Travel Partner (Deb, wife)

- I don't get the when to travel slide. And the colors are not my fave.
- There isn't much of a conclusion. It is good to know, but how is it applicable?

Resources

RITA/BTS data

https://www.tableau.com/about/blog/5-unusual-alternatives-pie-charts

https://financebuzz.com/departure-delays-cost-dollars-and-years

https://www.faa.gov/air_traffic/by_the_numbers