Understanding and Visualizing U.S. Flight Delays

Quan Nguyen

Abstract— This research explains why flights get delayed across U.S. airlines, dig ging into detailed records of flight arrivals, cancellations, and delays. The dataset covers all the major airports and carriers, tracking everything from short 15-minute holds to full cancellations and unexpected diversions. By comparing how different airlines perform at various airports, the study re veals which carriers handle disruptions better and where the system tends to break down. These findings matter because they show exactly where airlines and airports you should avoid to prevent travel delays. The data suggests that while the weather will always be unpredictable, many delays could be prevented with better scheduling and contingency planning.

1 Introduction

Air travel has become a vital part of our global infrastructure, playing a central role in how people connect across vast distances and how goods are transported efficiently around the world. Despite its importance and the advances in aviation technology and logistics, flight delays remain a persistent and costly issue. They cause frustration for passengers, who often face missed connections, extended travel times, and added expenses. At the same time, delays disrupt airline operations, complicated scheduling, increase operational costs, and strain airport resources.

This project aims to take a deep and data-driven look at the underlying causes of flight delays in the United States by analyzing detailed flight performance data spanning the past ten years. Rather than focusing on isolated incidents, the goal is to uncover broader patterns—identifying which airports and airlines are most affected, how delays fluctuate throughout the year, and what specific factors are consistently driving these disruptions, whether it's weather, infrastructure, air traffic congestion, or internal airline issues.

By visualizing and interpreting these trends through interactive dashboards, the project not only highlights the scope and complexity of the delay problem but also offers a clearer path toward potential solutions. The hope is that by gaining a better understanding of where and why delays happen, we can inform smarter decisions—from airline scheduling strategies and infrastructure investments to passenger travel planning. Ultimately, the goal is to contribute to a more efficient, predictable, and reliable flying experience for everyone involved.

2 Related Work

Flight delays continue to be one of the most persistent and well-documented challenges within the aviation industry. A substantial body of research has emerged over the years to address this issue, focusing on uncovering root causes, identifying patterns in delay occurrences, and developing effective strategies for prediction and mitigation. This section surveys key contributions in the existing literature, particularly those that have utilized large-scale transportation datasets or offered foundational frameworks that inform current practices in understanding and managing flight delays.

A cornerstone of many studies in this field is the data made available by the Bureau of Transportation Statistics (BTS) [1,6], which classifies delays into five primary categories: air carrier-related delays, weather, National Airspace System (NAS) inefficiencies, security issues, and late-arriving aircraft. This classification system not only standardizes delay reporting across the industry but also provides an essential foundation for operational analysis and regulatory oversight. These categories form the basis for most predictive and diagnostic models used by airlines, airports, and researchers alike.

The Government Accountability Office (GAO) has contributed significantly to the policy side of delay research. In its analysis of delay trends, particularly in the periods before and after the COVID-19 pandemic [2], the GAO identified how systemic disruptions—such as changes in demand, labor shortages, and evolving safety protocols—exacerbated operational challenges for airlines. The GAO also highlighted regulatory efforts aimed at strengthening passenger rights and improving airline accountability, which have influenced how carriers handle scheduling and compensation during delays.

Recent academic studies have turned toward advanced analytical techniques, such as predictive modeling and machine learning, to enhance both the accuracy and efficiency of delay forecasting. One notable approach integrated real-time meteorological data and historical delay durations to improve decision-making processes for flight planning and air traffic control [3]. Another study introduced the Flight Delay Path Previous-based Machine Learning (FDPP-ML) model, which leverages historical route-specific delay data to increase predictive accuracy and help airlines anticipate where and when delays are most likely to occur [5].

The issue of flight delays extends beyond the U.S., and international operations have been examined to draw global comparisons. For instance, a study focusing on European airlines explored additional delay drivers such as airport-specific operational limitations and aircraft maintenance protocols. The research also emphasized the economic implications of delays, particularly in regions with strict compensation laws for disrupted passenger itineraries [4].

Weather remains one of the most unpredictable and impactful contributors to flight delays. Several studies have sought to isolate specific weather phenomena—such as thunderstorms, snowstorms, and low-visibility conditions—and quantify their influence on departure and arrival times. These investigations have informed new strategies to build more resilient flight schedules and develop contingency plans tailored to weather-related risks [7].

In terms of operational responses, research has proposed a range of strategies aimed at delay mitigation. Improvements in crew scheduling systems, proactive aircraft maintenance, and advancements in air traffic management technologies have all shown promise in reducing the frequency and severity of delays. A systems-level approach—where airlines, airports, and regulatory bodies collaborate to optimize end-to-end operations—has been particularly effective in minimizing cascading effects and ensuring more efficient flight turnarounds [8].

Taken together, these studies provide a comprehensive and nuanced understanding of the multifactorial nature of flight delays. They underscore the importance of data-driven approaches and cross-sector collaboration in addressing the problem. From predictive algorithms and weather modeling to regulatory reforms and systems engineering, the literature highlights a wide range of tools and frameworks that can be leveraged to make air travel more reliable, efficient, and passenger-friendly.

3 Project Description

As part of a comprehensive analysis of flight delays, I developed three interactive dashboards utilizing D3.js to enable dynamic data exploration. Each dashboard highlights a distinct aspect of the issue, allowing users to examine flight delay patterns from multiple perspectives. The first dashboard visualizes overall delay trends across airlines and airports, helping users identify where delays are most frequent. The second focuses on temporal patterns, such as delays by time of day, day of the week, and seasonal variations. The third dives into contributing factors, such as weather conditions, flight distance, and carrier performance. By combining interactive visual elements with rich datasets, these dashboards empower users to engage with the data intuitively and uncover meaningful insights that might otherwise be hidden in static reports.

3.1 Flight Delays vs. Total Flights

This dashboard presents a side-by-side comparison of the total number of flights and the number of delayed flights for various airlines. The visualization is designed to help users quickly identify which airlines experience the most delays relative to their total operations. The x-axis displays different airline names, while the y-axis shows the number of flights.

For each airline, there are two bars: one blue bar representing the total flights and a yellow bar representing the number of delays. To allow for deeper exploration, users can filter the data by selecting specific airports, which makes it easier to analyze trends at a more localized level. This visualization is important because it helps uncover patterns in airline performance and highlights which airlines or airports are more frequently affected by delays, offering valuable insight into where operational improvements might be needed.

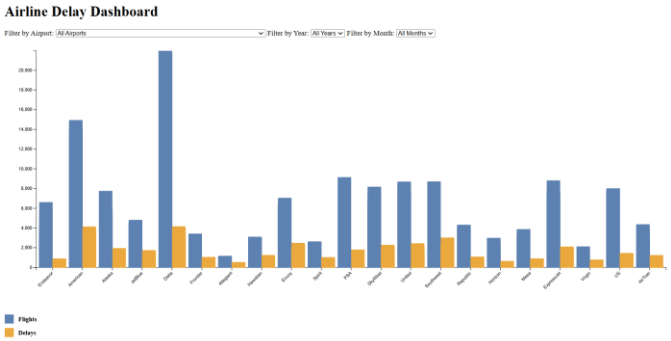

Figure 1. Airline delay Dashboard

3.2 Monthly Flight Delay Percentages

The second dashboard focuses on how the percentage of delayed flights changes throughout the year. It displays a bar chart where each bar corresponds to a single month, with the height of the bar indicating the percentage of flights that were delayed during that time. This format makes it very easy to spot temporal trends in the data, such as increases in delays during the winter or summer months, when weather or high travel volumes are likely to cause disruption. By visualizing monthly delay patterns, this dashboard helps users understand when delays are most common, making it a useful tool for planning around peak seasons or anticipating operational challenges tied to specific times of the year.

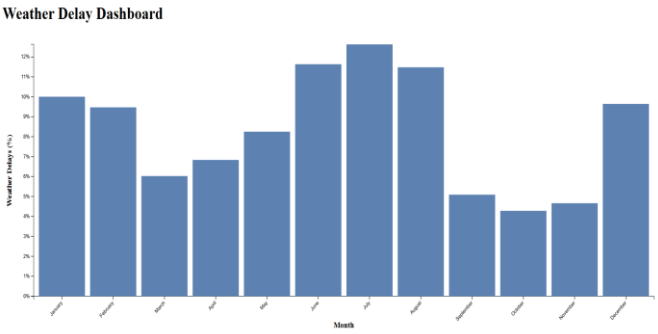

Figure 2. Weather delay dashboard

3.3 Reasons for Delays

The third dashboard provides a breakdown of the reasons behind flight delays. The data is visualized using a pie chart that divides delays into categories such as weather, airline-related issues, National Airspace System (NAS) traffic, security, and late-arriving aircraft. Users can filter this view to focus on specific airlines, allowing them to compare how different carriers are affected by various delay factors. This type of insight is particularly valuable for understanding the root causes of operational inefficiencies. For example, if airline-related issues are responsible for a large share of delays, internal process improvements may be necessary. On the other hand, if delays are mostly due to NAS congestion or weather, broader infrastructure or systemic solutions might be more appropriate.

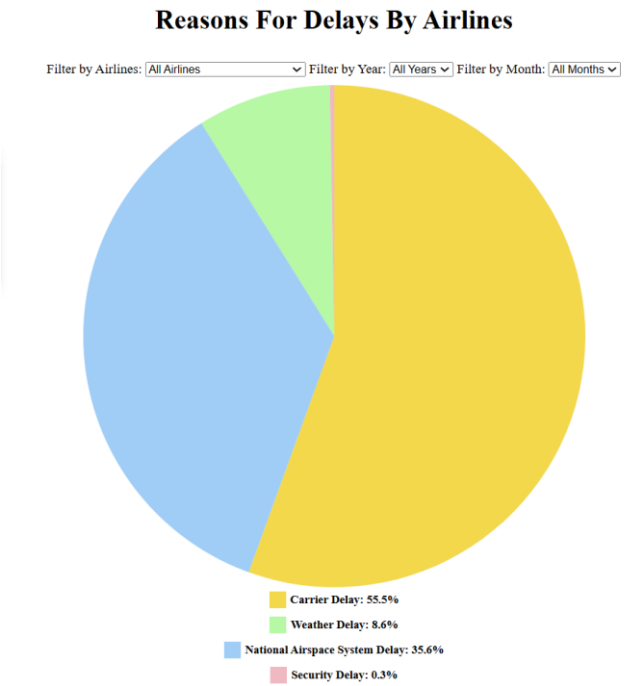

Figure 3. Reason for delay dashboard

4 Design Justifications

The overall layout of the dashboard system was created with simplicity and usability in mind. Each dashboard addresses a specific aspect of the flight delay data and is arranged in a way that encourages intuitive navigation and quick understanding. The design avoids visual clutter and focuses on clarity, ensuring that even users without a technical background can engage with the data meaningfully.

For the Flight Delays vs. Total Flights dashboard, the dual-bar chart format was chosen because it allows for immediate visual comparison.

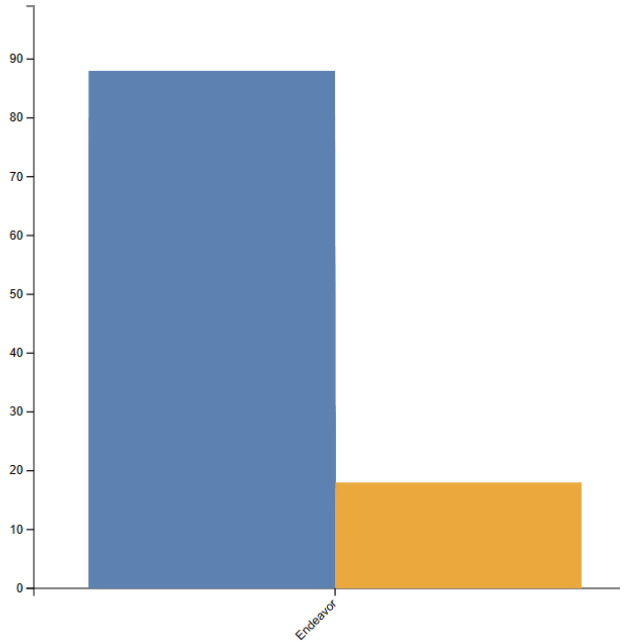

Figure 4. dual-bar chart

This makes it easy to see how many of an airline's flights are delayed relative to its total volume. Including the option to filter by airport adds another layer of utility, enabling users to examine trends on a more granular, regional level.

The Monthly Delay Percentages dashboard uses a simple bar chart to represent month-over-month changes in delay frequency. This straightforward visualization was chosen for its effectiveness in highlighting patterns and seasonal trends, which are crucial for planning staffing levels, allocating resources, and forecasting demand.

The Reasons for Delays dashboard uses a pie chart because it clearly communicates how delay causes are distributed. This type of visualization is familiar to most users and allows for quick assessments of which issues are most prevalent. Allowing users to filter by airline makes the insights even more specific and actionable.

5 Interactive Features

Each dashboard includes a number of interactive features that enhance the user experience and support data exploration. Dropdown filters allow users to customize the data they see based on airports or airlines, which helps them tailor the insights to their particular interests or concerns.

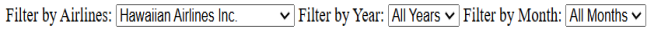

Figure 5. Drop down filters

Tooltips provide additional context when users hover over a data point, such as the exact number or percentage represented, without cluttering the main visual. Legends are included to clearly explain what each color or chart segment means, ensuring that the dashboards remain accessible and easy to understand.

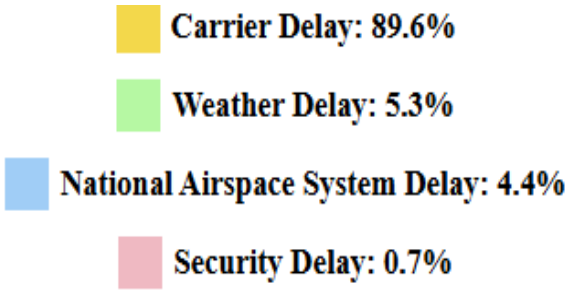

Figure 6. Legends

6 Visual Design

The dashboards were designed with a clean, centered alignment to keep them visually organized and easy to scan. The color palette—primarily blue and yellow—was chosen to provide high contrast without overwhelming the viewer. Blue is used for total flights or neutral categories, while yellow is used to highlight delays, drawing attention to problem areas and making comparisons simple and intuitive.

7 Key Takeaways

The dashboards revealed several important insights about flight delays. Some airlines and airports consistently experience more delays than others, which could be attributed to factors such as congestion,

staffing shortages, or infrastructure limitations. Delay rates tend to increase during certain months, most likely due to weather-related challenges or increased travel demand during holidays and peak vacation seasons. When examining the causes of delays, the data shows that NAS-related issues and airline-side problems are the most common, suggesting that improvements in air traffic control efficiency and internal airline operations could have the greatest impact on reducing delays.

Conclusions

These D3.js dashboards transform complex and often overwhelming flight data into clear, accessible, and interactive visual tools that users can actually learn from. Rather than presenting raw numbers or static charts, each dashboard offers a dynamic way to explore flight delays from multiple perspectives—volume, timing, and cause. This multifaceted approach allows users to dive deep into the data and uncover patterns that might otherwise go unnoticed. Whether it's identifying which airlines have the highest delay rates, spotting seasonal trends in on-time performance, or understanding the most common causes of delays, the dashboards bring clarity to a traditionally messy dataset. For travelers, this means being empowered to make more informed decisions—choosing flights or carriers based on performance, adjusting travel plans around likely delay periods, or simply gaining a better understanding of what's happening behind the scenes. Ultimately, the goal is to turn data into actionable insights that can contribute to smoother, more efficient travel experiences.

References

[1] Bureau of Transportation Statistics," Understanding the Reporting of Causes of Flight Delays and Cancellations," [Online]. Available: https://www.bts.gov.

[2] U.S. Government Accountability Office," Airline Passenger Protections: Observations on Flight Delays and Cancellations," GAO-23-105524, 2023. [Online]. Available: https://www.gao.gov.

[3] Predictive Analytics," Flight Delay Prediction Based on Delay Time Using Predictive Analytics," Springer Link, 2024.

[4] Sustainability," Factors Affecting International Flight Delays and Their Impact on Airline Operations," MDPI, 2022.

[5] Machine Learning," A Novel Intelligent Approach for Flight Delay Prediction," Springer Link, 2023

[6] Bureau of Transportation Statistics," Airline On-Time Performance and Causes of Flight Delays," [Online]. Available: https://www.bts.gov.

[7] Journal of Transportation," Impact of Weather on Flight Delays," 2023.

[8] Aviation Research," Operational Challenges and Solutions for Mitigating Flight Delays," 2024.