

Domestic Arrivals Delay Analysis

**A review of United States domestic flights delays and cancellations in
2003-23**

Executive Summary

This report outlines a consultancy project for the U.S. Department of Transportation (DOT) focused on analysing arrival flight delays at public airports across the United States. The analysis encompasses data from American domestic flights spanning June 2003 to December 2023. The DOT seeks insights on the following issues:

- **Avoidable Delays:** Caused by on-ground and aviation operations, or air traffic control.
- **Unavoidable Delays:** Caused by harsh weather conditions or security reasons.
- **Flight Cancellations:** Distribution, rate, and patterns based on season, regions, and states.

In addition to passenger security, the DOT is concerned with the financial and temporal impacts of delays and cancellations in the aviation industry. In 2023, one in five flights were delayed due to avoidable or unavoidable reasons. These delays can cause passengers to miss connecting flights or accommodation reservations, placing additional strain on the National Aviation System to maintain the schedule. A reduction of even 10% in flight delays could save over \$2 billion and 40,000 hours, significantly enhancing passenger satisfaction. This analysis aims to identify opportunities to minimise these effects.

The approach involves analysing general categories of delay, distinguishing between avoidable and unavoidable delays, and considering the roles of airlines and airports. This includes a historical review and examination of delay ratios for arrival flights, delay duration per delayed arrival, and the causes of delays. Pareto charts are used to identify the most influential elements in the avoidable delay category, such as airlines responsible for 80% of domestic flights and states hosting the highest number of airports, to address NAS-induced delays. Seasonal trends for weather and security-related delays are also examined to anticipate future occurrences.

Key airport characteristics evaluated include elevation, area, and number of runways. Conclusive results were derived from the type of airports (international, regional, and domestic), age of facilities, and total number of incoming flights, using techniques such as A/B testing, scatterplots, and cohorts.

Findings

- The average number of yearly flights decreased slightly from 2003 to 2023, but the median delay duration per delayed flight and per total arrivals increased by 31% and 37%, respectively.
- 19.2% of total arrivals were delayed, with avoidable delays accounting for 96% of these. Late aircraft had the highest share at 35.8%.

- June, July, and December had the highest delay rates due to increased holiday travel and seasonal weather conditions.
- Southwest Airlines had the highest number of total delays by airline (17.8%) and by late aircraft (25.8%), but their delay rate was just below the average.
- California experienced the highest number of NAS delays (10.9%), attributable to its high number of airports (34).
- Airports established after 2000 showed better results in avoidable delay rates (15.59% compared to 18.46%) but higher average delay duration per delayed flight.

A digital dashboard with filter options has been created to facilitate further exploration of the data.

Introduction

Each month, between 450,000 and 650,000 domestic flights take off and land at U.S. airports, transporting 751.4 million passengers in 2023. Despite this, airports and airlines face numerous challenges that result in delays, cancellations, and diverted flights. These issues can stem from technical problems, weather conditions, busy airspace, luggage handling delays, or security notices. In 2023, 20.25% of domestic flights were delayed, and 1.3% were cancelled, affecting nearly 170 million passengers annually.

Airlines offer amenities and services to mitigate the inconvenience of delays and cancellations, such as rebooking, meal vouchers, or overnight accommodations. According to the Federal Aviation Administration, the operational costs of delayed and cancelled flights exceed \$25 billion annually. Understanding the root causes of these issues is essential for identifying improvement opportunities.

Since June 2003, airlines have reported on-time performance and causes of delays and cancellations to the Bureau of Transportation Statistics of the DOT. Delays are categorized into five broad groups: Air Carrier, National Aviation System (NAS), Weather, Late-Arriving Aircraft, and Security. A flight is considered delayed if it arrives 15 or more minutes later than scheduled.

Delay Categories:

- **Airline:** Delays due to airline-controlled circumstances such as maintenance or crew issues (*Avoidable*).
- **Late-Arriving Aircraft:** Delays due to a previous flight with the same aircraft arriving late (*Avoidable*).
- **National Aviation System (NAS):** Delays due to non-extreme weather, airport operations, heavy traffic, and air traffic control (*Avoidable*).

- **Weather:** Delays due to significant meteorological conditions such as tornadoes, blizzards, or hurricanes (*Unavoidable*).
- **Security:** Delays due to security breaches or other security reasons (*Unavoidable*).

Approach

The analysis is based on arrival flights of various carriers at U.S. public airports from June 2003 to December 2023. Data includes total incoming flights, delays, cancellations, diverted flights, reasons for each occurrence, and total delay duration. This data is combined with information on aviation facilities, such as location, state, and date of activation, for a comprehensive analysis. The data, sourced from the DOT's official website, has been cleaned and organised using BigQuery.

The analysis is divided into two major parts:

- **Avoidability:** Inspection based on avoidable and unavoidable roots of delays
- **Technicality:** Evaluation of performance of airlines, airports and air control (NAS) as functioning units of the system

Metrics are converted into rates, shares, and percentages to account for the varying sizes of airlines and airports. Pareto charts are used to prioritise areas for improvement by weighing the effect of major elements in avoidable delays.

Analysis

A historical overview shows a direct correlation between the number of flights and delays. Significant events such as the Great Recession and the COVID-19 pandemic have markedly affected flight numbers and delays. The median delay duration per total arrivals increased from 7 to 11 minutes, while the median delay duration per delayed arrival increased by 35%.

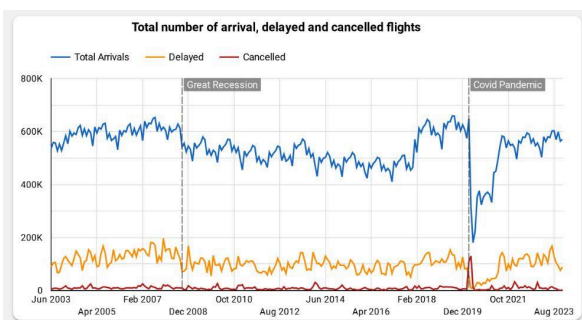


Fig. 1: Total number of arrivals, delays and cancellations from 2003-2023

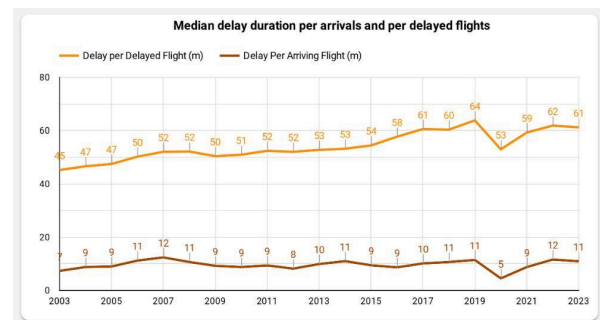


Fig. 2: Median delay duration by arrival and delayed arrival

The breakdown of delayed arrivals indicates that the major causes of delay are late aircraft (6.62%), NAS (6.2%), and airlines (5.64%). Seasonal trends reveal a 5-6% rise in delays during

summer and December, with a sharper increase in airline and late aircraft delays during these periods.

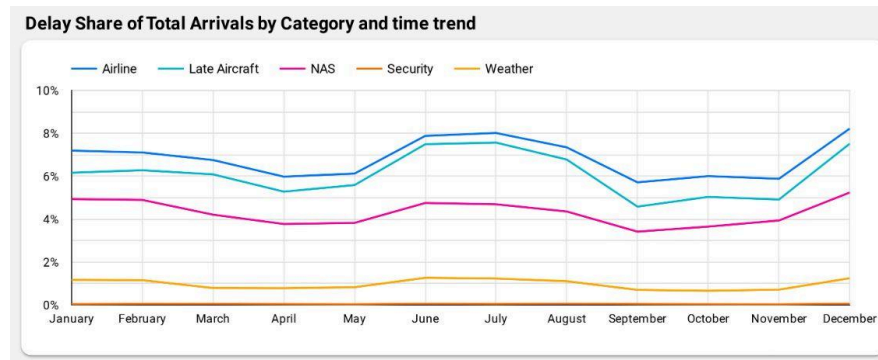


Fig3: Seasonal trend of delays by reason

Results

Avoidable delays, which can be addressed by authorities or airlines, constitute the primary focus. Pareto charts show that the top 10 airlines are responsible for airline-originated delays, with Southwest Airlines accounting for 17.9% of total delays and for 80% of late aircraft delays, with Southwest Airlines taking 25.8%.

NAS delays are analysed based on states, with California, Texas, Illinois, Florida, and New York having the highest numbers. Sixteen states account for 80% of NAS delays. Scatter plots show that variance in cancellation and delay rates reduces significantly for airlines with over 350,000 annual arrivals.

Evaluating airports, it was found that elevation, area, and proximity to cities have no significant impact on delay or cancellation rates. By doing cohort analysis and hypothesis testing, it was concluded that newer airports show better avoidable delay rates but not necessarily in the unavoidable ones. Although there are no patterns visible in delays of airport cohorted by 10 years age groups.

	Activation Period	Total Arrival	Avoidable delay Count	RoD%	Standard Error (std)	Z score	P score	Probability of variant being better than control
Control	After 2000	140434	21897	15.59%	0.10%	29.74	0.00	0%
Variant	Before 2000	132824438	24519008	18.46%	0.00%			
	Activation Period	Total Arrival	Unavoidable delay Count	RoD%	Standard Error (std)	Z score	P score	Probability of variant being better than control
Control	After 2000	140434	1068	0.76%	0.02%	-0.92	0.82	84.57%
Variant	Before 2000	132824438	981918	0.74%	0.00%			

Figure 3 & 4: A/B test results for comparison of avoidable and unavoidable delays in airports built after 2000 and before 2000

Recommendations

Regarding Airlines:

- Reduce loading time by digitising boarding and luggage claim process
- Considering reserve aircrafts in hubs for replacing the late ones
- Increase or relocate ground personnel to expedite pre-flight operations.
- Introduce incentives and seasonal KPIs to reduce delays gradually.

Regarding DOT Operations:

- Equip airports and fleets with better tools for handling harsh weather.
- Analyse seasonal climate patterns to define temporary reroutes.
- Implement flight caps for major airlines and incentives for smaller ones to balance flight capacities.
- Increase flight tax in states with higher delay rates to compensate for losses.
- Consider establishing new regional hubs to distribute incoming flight loads.

Conclusion

The aviation industry is a direct indicator of economic growth and has become more popular due to the increasing value of time. Maintaining it requires extensive infrastructure, a significant workforce, and long-term planning. While it offers substantial profits, repeated delays and cancellations can lead to significant losses. Addressing these issues is crucial for sustaining industry growth and passenger satisfaction.