

On-Time Performance Analysis of Airlines.

Flight delays are influenced by seasonal travel spikes, time-of-day trends, and weather-related issues. July and August are the worst months for delays due to the heavy summer travel load. On the flip side, December sees better punctuality. Morning flights are the most reliable, with an average delay of just 7 minutes, while night flights face the longest delays—up to 22 minutes—because of disruptions piling up throughout the day.

Regional airports like Rhineland, WI, and Aspen, CO, struggle with weather-related delays the most, underscoring the need for better infrastructure. Meanwhile, older planes are more prone to significant delays, making fleet upgrades a top priority.

To tackle these challenges, airlines should focus on fine-tuning schedules during peak months, prioritizing morning flights for key routes, investing in weather resilience for regional airports, and improving maintenance programs for aging planes. These steps are critical for improving operations and ensuring passengers have a better overall experience.

FLIGHT DATA FOR YEARS AND AIRLINES.

Fig. 1(2023)

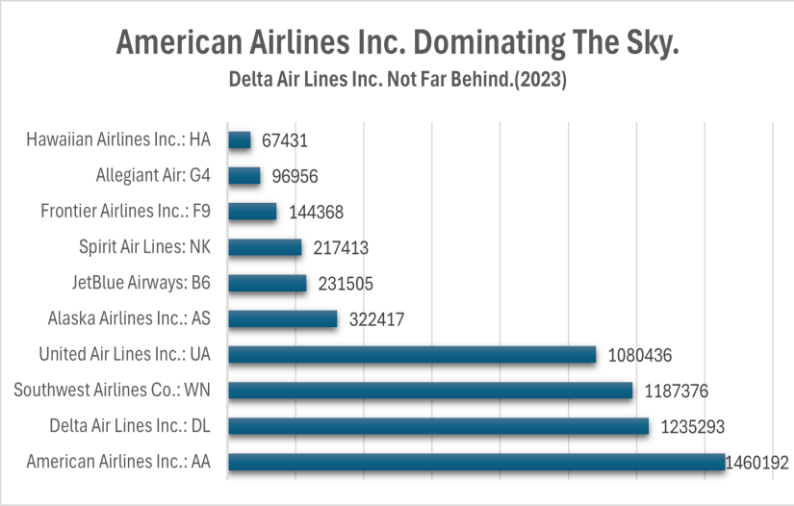


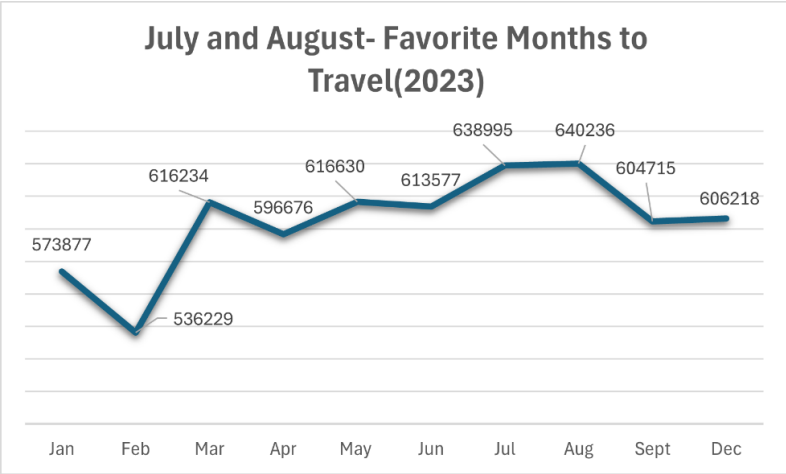
Fig. 2 (2024)



Ref: Fig 1 and 2

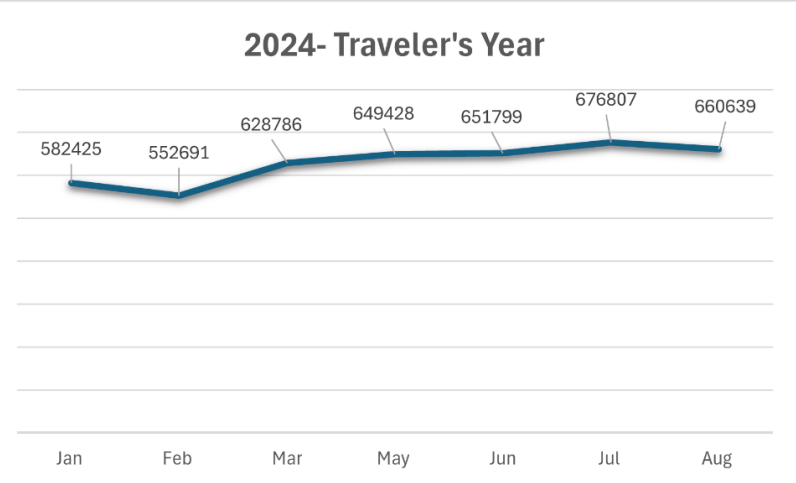
The 2023 and 2024 airline performance analysis highlight the dominance of major carriers such as American Airlines, Delta Air Lines, Southwest Airlines, and United Airlines, which collectively operated over 5 million flights. American Airlines having the highest number of flights in 2023 and 2024. Regional airlines like Alaska Airlines and Hawaiian Airlines contributed significantly within their niches, while low-cost carriers such as JetBlue and Spirit Airlines demonstrated strong market competition in the budget segment. The data underscores a concentrated market among the top four airlines, while smaller and regional competitors cater to specific customer demands. These insights provide a comprehensive overview of flight volumes and market dynamics, offering a foundation for strategic operational improvements.

Fig.3



Ref: Fig 3. In 2023, flight volumes peaked during summer, with July and August recording the highest numbers at 638,995 and 640,236, reflecting strong vacation demand. February saw the lowest activity at 536,229 due to post-holiday slowdowns, while volumes steadily grew from March to June. December showed a recovery with 606,218 flights, driven by holiday travel.

Fig.4



Ref: Fig.4 In 2024, flight volumes showed steady growth with smaller monthly variations. July peaked at **676,807** flights, followed by August at **660,639**, reflecting sustained summer demand. Early months like February and March also saw higher activity, indicating improved operational consistency and demand planning.

July is the peak of the summer travel season, driven by vacation plans, school breaks, and good weather. Passenger numbers are at their highest this month, with a noticeable spike in flight volumes compared to other months. This surge increases operational demands, requiring airlines to handle tighter schedules, allocate resources efficiently, and maintain on-time performance. For industry, July is a key period for revenue and customer satisfaction, making operational efficiency crucial to meeting the high demand.

FLIGHT DATA FOR DELAYS:

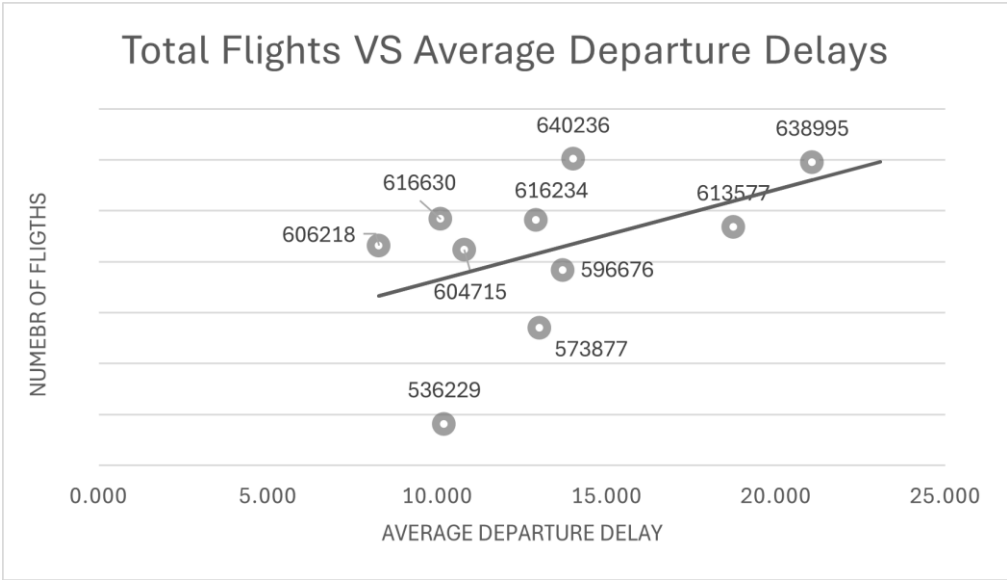


Fig.5

Ref Fig.5 The moderate positive correlation of **0.46** between flight numbers and average departure delays shows that more flights tend to lead to longer delays, but the connection isn't very strong. This means that while higher traffic during busy months like July and August adds to delays, other factors such as operational issues, weather, or airport capacity also play a major role. To reduce delays, airlines should focus on improving operations and allocating resources effectively during peak travel months. Examining other causes of delays could also reveal specific areas for improvement to boost on-time performance.

Analysis shows that flight delays tend to increase as the day goes on. Morning flights have the shortest average delays at **7.01** minutes, while delays peak at night with an average of **22.08** minutes. Afternoon flights face delays of **13.93** minutes on average, and evening flights see delays rise to **19.70** minutes. This pattern reflects how congestion and operational strain build up throughout the day. Flying in the morning is the best option for minimizing delays, while flights later in the day are more likely to face disruptions. Ref: Fig 6

Fig.6 – Best Time of the Year



Fig.7- Best Time of the Week

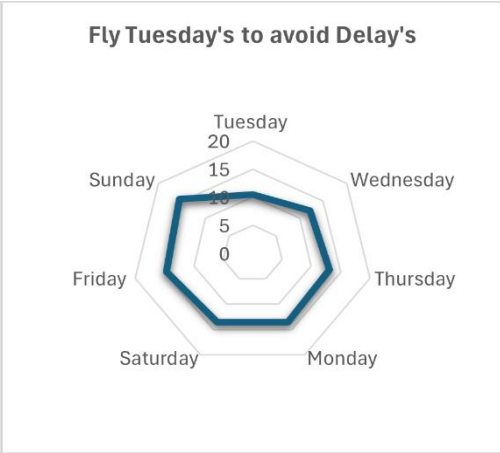
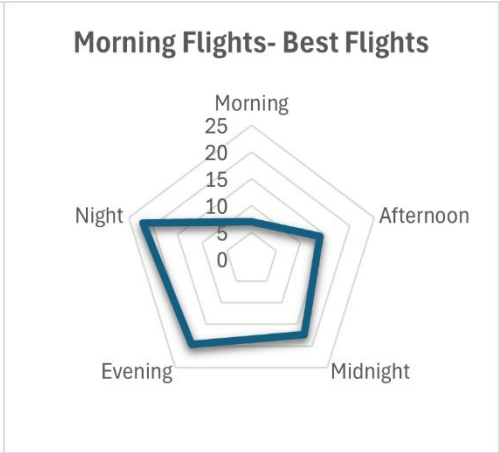


Fig.8- Best Time of the Day



The analysis shows clear patterns in average departure delays based on the month, day of the week, and time of day. Ref: Fig.6 December has the lowest average delay at 8.28 minutes, likely due to lower travel volumes and more efficient winter scheduling. On the other hand, delays peak in July (21.07 minutes) and June (18.74 minutes), aligning with the busy summer travel season. This highlights how seasonal demand impacts delays.

Looking at days of the week Ref: Fig.7, Tuesday and Wednesday are the most reliable for on-time departures, with average delays of 10.49 and 12.25 minutes. Sundays, however, see the highest delays at 15.50 minutes, likely due to heavier weekend travel. Time of day matters Ref: Fig 8, mornings have the shortest delays at 7.01 minutes, while delays steadily rise throughout the day, peaking at 22.08 minutes at night. These trends show how both timing and operational pressures affect flight delays.

FLIGHT DELAY RELATED ISSUES AND PLANE AGE:

The data indicates a connection between airplane age and average delays, with older aircraft like N669UA (**25 years**) and N665UA (26 years) experiencing significantly higher delays of **319** and 168.5 minutes, respectively. In contrast, newer planes such as N512SY (3 years) and N413SY (2 years) have much lower delays, averaging around 40 minutes, reflecting the superior reliability of modern aircraft. However, exceptions exist; for instance, N664UA (**26 years**) records a delay of just 1 minute, highlighting the role of

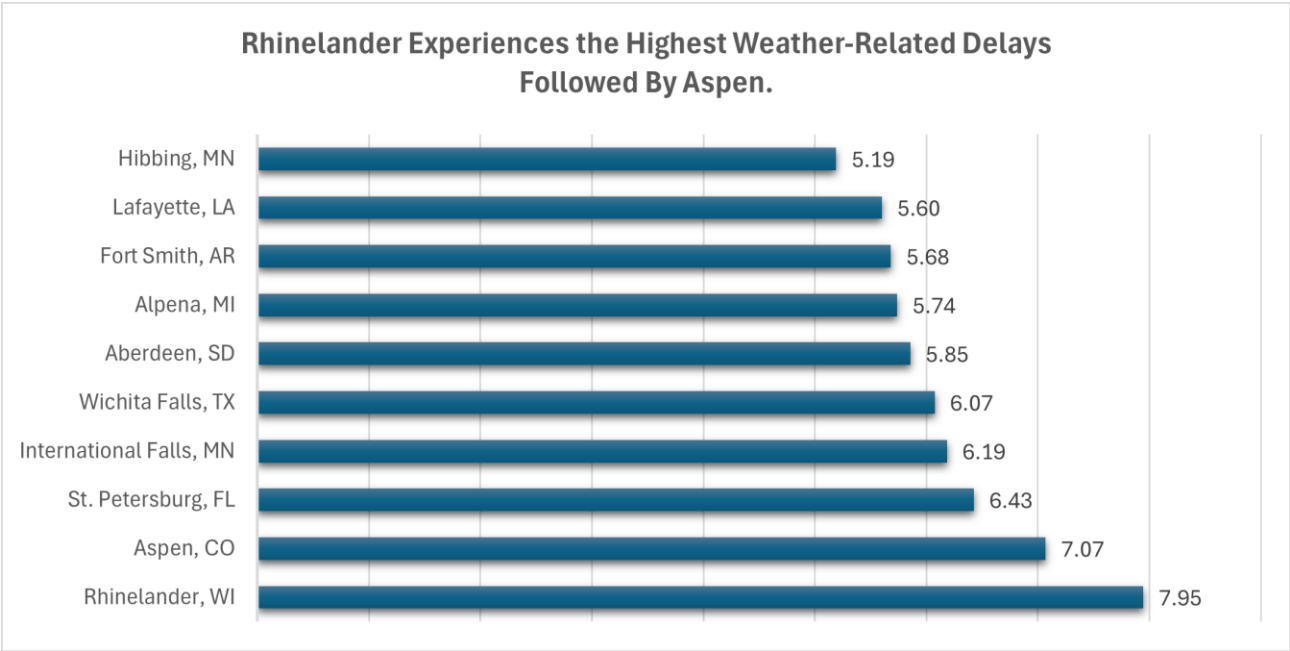
strong maintenance. Conversely, N886BR (1 year) shows delays of 178.75 minutes, suggesting operational factors like routes and scheduling also influence performance. While older planes generally show more delays, maintenance and operations are equally critical in determining reliability.

ROUTES WITH MOST FLIGHTS:

Passenger traffic patterns shift significantly across the four quarters, reflecting varying travel demands. In the 1st Quarter, business routes dominate, with Washington, DC to New York City, NY (6,198 flights) and Chicago, IL to New York City, NY (5,920 flights) among the busiest, alongside the leisure-heavy Las Vegas to Los Angeles route (6,017 flights). The 2nd Quarter sees continued demand for regional travel, particularly within California, as routes like Los Angeles to San Francisco (5,589 flights) gain prominence, driven by spring and summer leisure activity.

In the 3rd Quarter, California routes grow further, with Los Angeles to San Francisco (5,436 flights) and the reverse route (5,309 flights) ranking highly, while Las Vegas to Los Angeles remains the busiest (6,098 flights). By the 4th Quarter, flight volumes decline significantly, with new routes like Orlando, FL to New York City, NY (1,853 flights) and New York City to Miami, FL (1,759 flights) reflecting a seasonal shift toward holiday and leisure travel, favoring warmer destinations. These trends highlight the dynamic nature of travel demand across seasons.

Fig,9



ADDRESSING WEATHER DELAYS:

Ref: Fig.9 Weather-related flight delays show clear patterns, with Rhinelanders, WI, topping the list at 7.95%, far exceeding other airports. Aspen, CO, follows at 7.07%, likely due to its challenging mountainous terrain and unpredictable weather. St. Petersburg, FL (6.43%) also ranks high, reflecting its exposure to tropical weather, while International Falls, MN (6.19%) deals with harsh winter conditions. Wichita Falls, TX (6.07%) rounds out the top five, often impacted by severe storms and strong winds.

Smaller regional airports like Aberdeen, SD (5.85%), Alpena, MI (5.74%), and Fort Smith, AR (5.68%) also report higher-than-average weather-related delays, likely influenced by extreme local weather. Hibbing, MN (5.19%) has the lowest delay percentage among the top 10 but still shows significant weather-related disruptions.

This analysis highlights a trend: airports in regions prone to extreme weather—whether it's snow, storms, or mountainous terrain—face higher delays. Smaller regional airports are especially affected, likely due to fewer resources and less robust infrastructure to handle adverse conditions. Investing in better weather management systems, particularly at regional hubs and high-risk locations like Rhinelander and Aspen, could help improve reliability and reduce disruptions for travelers.

RECOMMENDATION:

Based on flight data analysis and operational trends, airlines can improve performance and passenger satisfaction by addressing key areas. Morning flights, with average delays of just 7 minutes, should be prioritized for critical routes to enhance on-time performance. Meanwhile, summer months like July and August require better planning, including increased staffing, improved ground services, and contingency measures to handle peak demand and minimize delays.

Weather-related disruptions at regional airports, such as Rhinelander, WI (7.95%) and Aspen, CO (7.07%), highlight the need for better forecasting tools, real-time communication, and infrastructure investments like de-icing facilities. Collaborating with airport authorities to develop contingency plans for extreme weather will further enhance reliability.

Upgrading fleets is another critical focus. Older planes, especially those over 20 years old, contribute significantly to delays. Replacing aging aircraft and leveraging predictive maintenance tools for all planes can reduce disruptions and lower long-term costs. Finally, improving passenger communication during delays through real-time updates, notifications, and compensation options will help maintain customer trust and satisfaction. By addressing these areas, airlines can enhance operational efficiency and deliver a better passenger experience.

References: OpenAI. (n.d.) ChatGPT. Accessed on November 21, 2024. Retrieved from <https://chatgpt.com/share/67400651-79b4-800d-ad37-1391273b9ae7>

```
(AVG(delay) WHEN dep_delay > 0 THEN 1 ELSE 0) AS avg_delay,
COUNT(*) AS total_flights
FROM
  "AwsDataCatalog"."business analysis"."202301" -- Replace with
  your table name
WHERE
  tail_num IS NOT NULL -- Exclude rows with missing tail
  numbers
GROUP BY
  tail_num
ORDER BY
  avg_delay ASC
LIMIT 15;
```

Access denied to glue:GetDatabase
You don't have permission to glue:GetDatabase. To request access, copy the following text and send it to your AWS administrator. Learn more about troubleshooting access denied errors.

The error indicates that your AWS user does not have the required permissions to access the AWS Glue database for executing the query. Specifically, the permission `glue:GetDatabase` is missing. This permission is essential to retrieve the database schema information when using Athena.

Steps to Re...

How do weather delays compare yearly? What trends emerge from this data?

```
JOIN
  "AwsDataCatalog"."business analysis"."airline_id" as airlines
ON
  twentythree.mkt_carrier_airline_id = airlines.code
WHERE fl_date like '%2023%'
GROUP BY
  airline_code
ORDER BY
  total_flights DESC;
```

COLUMN_NOT_FOUND: line 15:6: Column 'airline_code' cannot be resolved or requester is not authorized to access requested resources

The error occurs because the column `airline_code` is not explicitly included in the `GROUP BY` clause as it is an alias. Athena requires grouping by the actual column or expression, not its alias. Let's fix the query by referencing the actual column name used in the `SELECT` clause.

Generated Query

```
month, origin_city_name, aci.description
GROUP BY
  month, origin_city_name, aci.description
ORDER BY
  month ASC, flight_count DESC;
```

can you modify this code to quarterly rather than montly

To modify the query for quarterly data instead of monthly, you can adjust the `SELECT`, `WHERE`, and `GROUP BY` clauses to use `quarter` instead of `month`. Here's the modified query:

```
SQL
Copy code
SELECT
  quarter,
  origin_city_name AS depart_city,
  aci.description AS destination_city,
  COUNT(*) AS flight_count
FROM
  "AwsDataCatalog"."business analysis"."202301" df
JOIN
  city_ma
ON
  acode = df.destination_market_id;
```

How do weather delays compare yearly? What trends emerge from this data?

```
FROM
  "AwsDataCatalog"."business analysis"."202301"
JOIN
  "AwsDataCatalog"."business analysis"."airport_id" AS airports
ON
  flights.origin_airport_id = airports.code
WHERE
  fl_date LIKE '%2023%'
GROUP BY
  origin
ORDER BY
  delay_probability_percentage DESC;
```

Your query has an issue related to the JOIN clause and GROUP BY. Let's analyze the query step-by-step and address the problems:

Issues in the Query

- Incorrect Column in JOIN Clause:
 - The `JOIN` clause uses `airports.code` as the join key, but the `GROUP BY` clause uses `origin` as the grouping key. This inconsistency can lead to incorrect results. Instead, you should use `airline_id` or whatever column is used in the `JOIN` clause for grouping.

Message ChatGPT

Appendix:

Data about flights between location for all quarters.

Quarter	From	To	Number of Flights
1st Quarter	Washington, DC	New York City, NY (Metropo	6198
1st Quarter	Las Vegas, NV	Los Angeles, CA (Metropolii	6017
1st Quarter	Chicago, IL	New York City, NY (Metropo	5920
2nd Quarter	Las Vegas, NV	Los Angeles, CA (Metropolii	6139
2nd Quarter	Chicago, IL	New York City, NY (Metropo	5756
2nd Quarter	Los Angeles, CA	San Francisco, CA (Metropo	5589
3rd Quarter	Las Vegas, NV	Los Angeles, CA (Metropolii	6098
3rd Quarter	Los Angeles, CA	San Francisco, CA (Metropo	5436
3rd Quarter	San Francisco, CA	Los Angeles, CA (Metropolii	5309
4th Quarter	Las Vegas, NV	Los Angeles, CA (Metropolii	2006
4th Quarter	Orlando, FL	New York City, NY (Metropo	1853
4th Quarter	New York, NY	Miami, FL (Metropolitan Arc	1759

Data About planes for Top and bottom 5 and random in Between Delays

airplar	avg_delay	total_flights	MFR-Date	Todays Date	Age	Manufacturer
N923SW	347.125	13	2002	2024	22	BOMBARDIER INC
N669UA	319	1	1999	2024	25	BOEING
N886BR	178.75	41	2023	2024	1	BOEING
N672UA	178	3	1999	2024	25	BOEING
N665UA	168.5	4	1998	2024	26	BOEING
N512SY	40.025	1083	2021	2024	3	YABORA INDUSTRIA AERONAUTICA S
N135SY	40.02028986	1148	2015	2024	9	EMBRAER S A
N910DU	40.00680272	958	2018	2024	6	BOEING
N413SY	40.00291545	1447	2022	2024	2	EMBRAER S A
N579NN	40	1629	2015	2024	9	BOMBARDIER INC
N430QX	4.666666667	38	2007	2024	17	BOMBARDIER INC
N725AN	2.666666667	7	2013	2024	11	BOEING
N720AN	2	5	2013	2024	11	BOEING
N664UA	1	7	1998	2024	26	BOEING
N661UA	1	4	1993	2024	31	BOEING

Rudrani Sakhare

MBAN-2