

# Flight Performance & Aviation Analytics for AeroStat Airlines

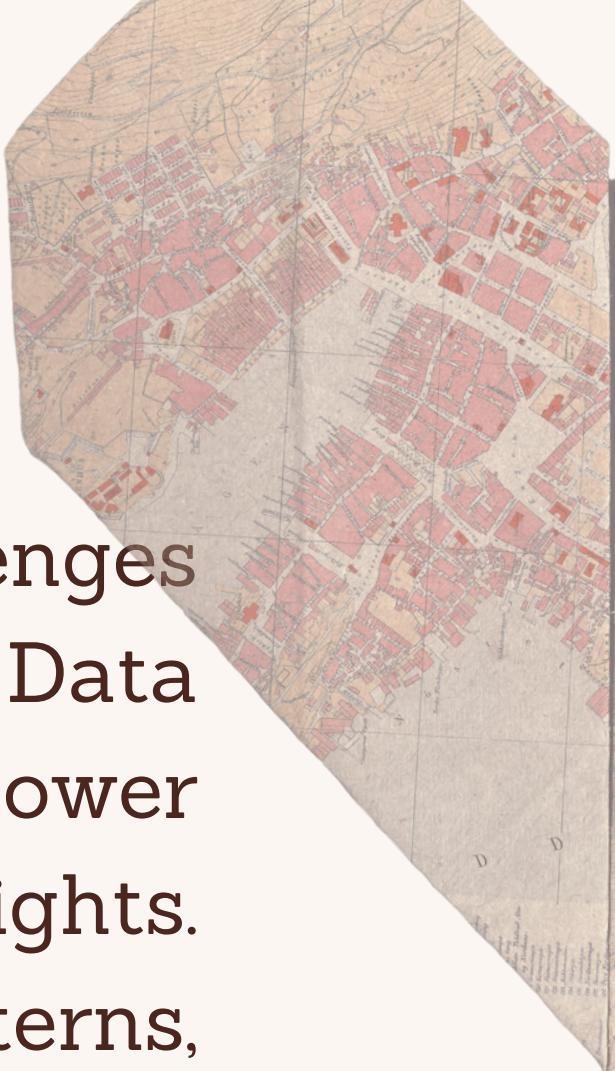
MYSQL | POWER BI | DAX | VISUALIZATION

-by BHUVAN JARI

# Introduction:



AeroStat Airlines operates an extensive flight network, facing challenges like delays, cancellations, and operational inefficiencies. As a Data Analyst, your role is to leverage SQL for data extraction, Power BI/Tableau for visualization, and machine learning for predictive insights. This project focuses on analyzing flight punctuality, cancellation patterns, and route efficiency to identify areas for improvement. By building an interactive dashboard with key performance indicators and trend analysis, you will provide executives with data-driven insights to optimize scheduling, reduce delays, enhance passenger experience, and improve financial efficiency, ensuring AeroStat Airlines remains competitive.



# SQL PROBLEMS:

## 1. Flight Delay Analysis

Calculate the average departure and arrival delay for all flights in the last 6 months.

## 2. Most Frequent Routes

Identify the top 10 most popular flight routes based on the total number of flights.

## 3. Airline Performance Ranking

Rank airlines by their on-time performance (lowest average delay).

## 4. Cancellation Trends

Find the top reasons for flight cancellations and their frequency.

## 5. Airport Congestion Analysis

Identify airports with the highest number of flight departures and arrivals.

## 6. Weather-Related Disruptions

Analyze how weather impacts delays and cancellations.

## 7. Flight Distance & Duration Trends

Calculate the average flight duration for different distance ranges.

## 8. Day-of-Week Flight Performance

Find which day of the week has the highest on-time flight performance.

## 9. Seasonality in Air Traffic

Identify which months have the highest number of flights and delays.

## 10. Passenger Connectivity & Hub Efficiency

Identify which airports act as the largest hubs based on the number of connecting flights.



# Flight Delay Analysis:

Calculate the average departure and arrival delay for all flights in the last 6 months.

```
SELECT
    AVG(DEPARTURE_DELAY) AS avg_departure_delay,
    AVG(ARRIVAL_DELAY) AS avg_arrival_delay
FROM
    flights_cleaned
WHERE
    YEAR = (SELECT
                MAX(YEAR)
            FROM
                flights_cleaned)
    AND MONTH >= (SELECT
                    MAX(MONTH)
            FROM
                flights_cleaned) - 6;
```

avg_departure_delay	avg_arrival_delay
15.6	13.6

# Most Frequent Routes:

Identify the top 10 most popular flight routes based on the total number of flights.

```
SELECT  
    ORIGIN_AIRPORT,  
    DESTINATION_AIRPORT,  
    COUNT(*) AS total_flights  
  
FROM  
    flights_cleaned  
GROUP BY ORIGIN_AIRPORT , DESTINATION_AIRPORT  
ORDER BY total_flights DESC  
LIMIT 10;
```

ORIGIN_AIRPORT	DESTINATION_AIRPORT	total_flights
LAX	JFK	118
JFK	LAX	115
SFO	LAX	102
LAX	SFO	99
LAS	LAX	96
LAX	LAS	96
HNL	OGG	83
OGG	HNL	83
LAX	ORD	76
MCO	ATL	76

# Cancellation Trends:

Find the top reasons for flight cancellations and their frequency.

```
SELECT
    c.CANCELLATION_DESCRIPTION,
    COUNT(f.CANCELLATION_REASON) AS cancellation_count
FROM
    flights_cleaned f
    JOIN
    cancellation_codes_cleaned c ON f.CANCELLATION_REASON = c.CANCELLATION_REASON
WHERE
    f.CANCELLED = 1
GROUP BY c.CANCELLATION_DESCRIPTION
ORDER BY cancellation_count DESC;
```

CANCELLATION_DESCRIPTION	cancellation_count
Weather	551
Airline/Carrier	380
National Air System	207

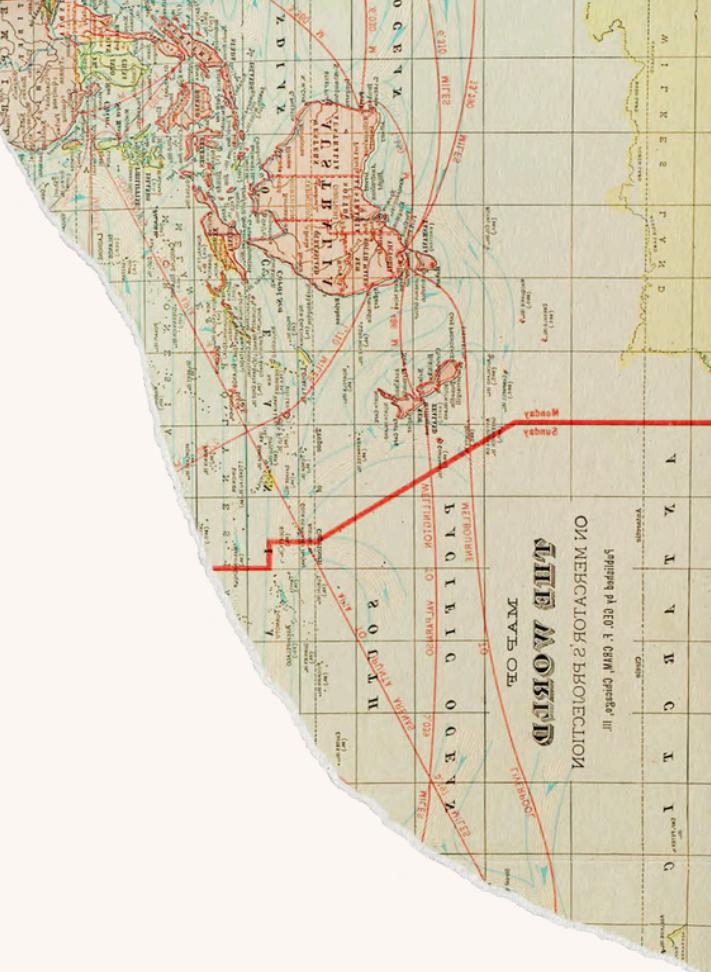
# Weather-Related Disruptions:

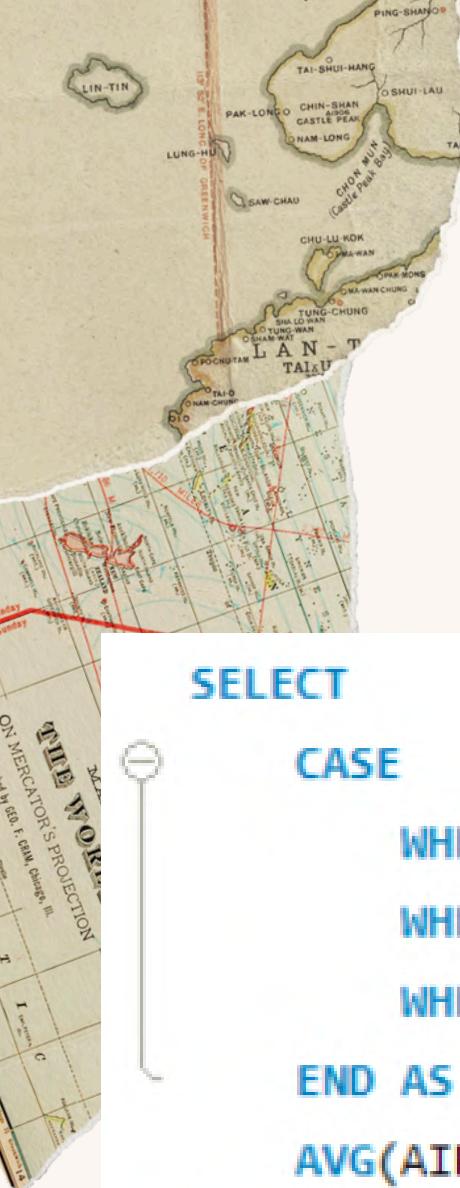
Analyze how weather impacts delays and cancellations.

- **SELECT**  
    **SUM(WEATHER\_DELAY) AS total\_weather\_delay,**  
    **COUNT(CASE**  
        **WHEN CANCELLATION\_REASON = 'B' THEN 1**  
    **END) AS total\_weather\_cancellations**  
**FROM**  
    **flights\_cleaned;**



total_weather_delay	total_weather_cancellations
29846	551





# Flight Distance & Duration Trends:

Calculate the average flight duration for different distance ranges.

```
SELECT
CASE
    WHEN DISTANCE < 500 THEN 'Short (<500 miles)'
    WHEN DISTANCE BETWEEN 500 AND 1500 THEN 'Medium (500-1500 miles)'
    WHEN DISTANCE > 1500 THEN 'Long (>1500 miles)'
END AS distance_category,
AVG(AIR_TIME) AS avg_flight_duration
FROM
flights_cleaned
GROUP BY distance_category;
```

distance_category	avg_flight_duration
Medium (500-1500 miles)	123.77
Long (>1500 miles)	261.27
Short (<500 miles)	52.53

# Day-of-Week Flight Performance

Find which day of the week has the highest on-time flight performance.

SELECT

DAY\_OF\_WEEK, AVG(ARRIVAL\_DELAY) AS avg\_arrival\_delay

FROM

flights\_cleaned

GROUP BY DAY\_OF\_WEEK

ORDER BY avg\_arrival\_delay ASC;

DAY_OF_WEEK	avg_arrival_delay
4	5.35
5	9.84
7	12.08
6	25.46

# Seasonality in Air Traffic:

Identify which months have the highest number of flights and delays.

```
SELECT  
    MONTH,  
    COUNT(*) AS total_flights,  
    AVG(ARRIVAL_DELAY) AS avg_arrival_delay  
FROM  
    flights_cleaned  
GROUP BY MONTH  
ORDER BY total_flights DESC;
```

MONTH	total_flights	avg_arrival_delay
1	49998	13.6



# Passenger Connectivity & Hub Efficiency:

Identify which airports act as the largest hubs based on the number of connecting flights.

SELECT

ORIGIN\_AIRPORT, COUNT(\*) AS total\_flights

FROM

flights\_cleaned

GROUP BY ORIGIN\_AIRPORT

ORDER BY total\_flights DESC

LIMIT 10;

ORIGIN_AIRPORT	total_flights
ATL	2864
DFW	2498
ORD	2395
DEN	1970
LAX	1913
IAH	1479
SFO	1461
PHX	1457
LAS	1186
MCO	1125

# SQL QUERIES EXECUTIVE SUMMARY :

This set of SQL queries provides a comprehensive analysis of airline performance and operational efficiency. It examines flight delays, highlighting average departure and arrival delays over the last six months, and ranks airlines based on on-time performance. The queries also explore route popularity, identifying the busiest flight routes, and analyze cancellation trends, pinpointing the most common reasons for flight cancellations. Additionally, the airport congestion analysis identifies the busiest airports by departures and arrivals, while weather-related disruptions are assessed in terms of delays and cancellations. The study further investigates flight duration trends based on distance categories, evaluates the best days for on-time flights, and examines seasonal patterns in air traffic. Lastly, hub efficiency is analyzed by identifying the airports with the highest number of connecting flights. This data-driven approach helps in understanding key operational bottlenecks, improving scheduling, and enhancing passenger experience.



# Power BI - DASHBOARD VISUALIZATION

## 1. Flight Operations Overview

**Problem:** Airlines need to monitor operational efficiency and punctuality to minimize delays and improve service reliability. Understanding flight activity trends, average delays, and on-time performance helps optimize scheduling and resource allocation.

## 2. Airline Performance Comparison

**Problem:** Airlines have varying levels of punctuality, delay management, and cancellation rates. Comparing their performance can help identify operational weaknesses and best practices, leading to improved industry standards and customer satisfaction.

## 3. Route & Airport Analytics

**Problem:** Certain airports and routes experience higher congestion and delays, affecting efficiency. Identifying the busiest airports and frequent routes helps in capacity planning, airport infrastructure development, and air traffic management.

## 4. Flight Cancellation & Delay Patterns

**Problem:** Unpredictable disruptions like weather, technical issues, and late aircraft affect airline operations. Analyzing the causes of cancellations and delays enables proactive strategies to minimize disruptions and enhance reliability.

## 5. Passenger Experience & Service Quality

**Problem:** Flight delays and missed connections directly impact customer satisfaction. Measuring factors like delay recovery and service efficiency helps airlines enhance passenger experience and loyalty.

## 6. Financial & Operational Efficiency

**Problem:** Flight delays, cancellations, and fuel inefficiencies result in financial losses for airlines. Analyzing cost-per-mile usage, and revenue loss helps optimize profitability and operational effectiveness.

# Flight Operations Overview

**50K**

Total Flights Operated

**47.9%**

On-Time Performance Rate

**15.25**

Avg\_Departure\_Delay

**116.78**

Average\_Flight\_Duration

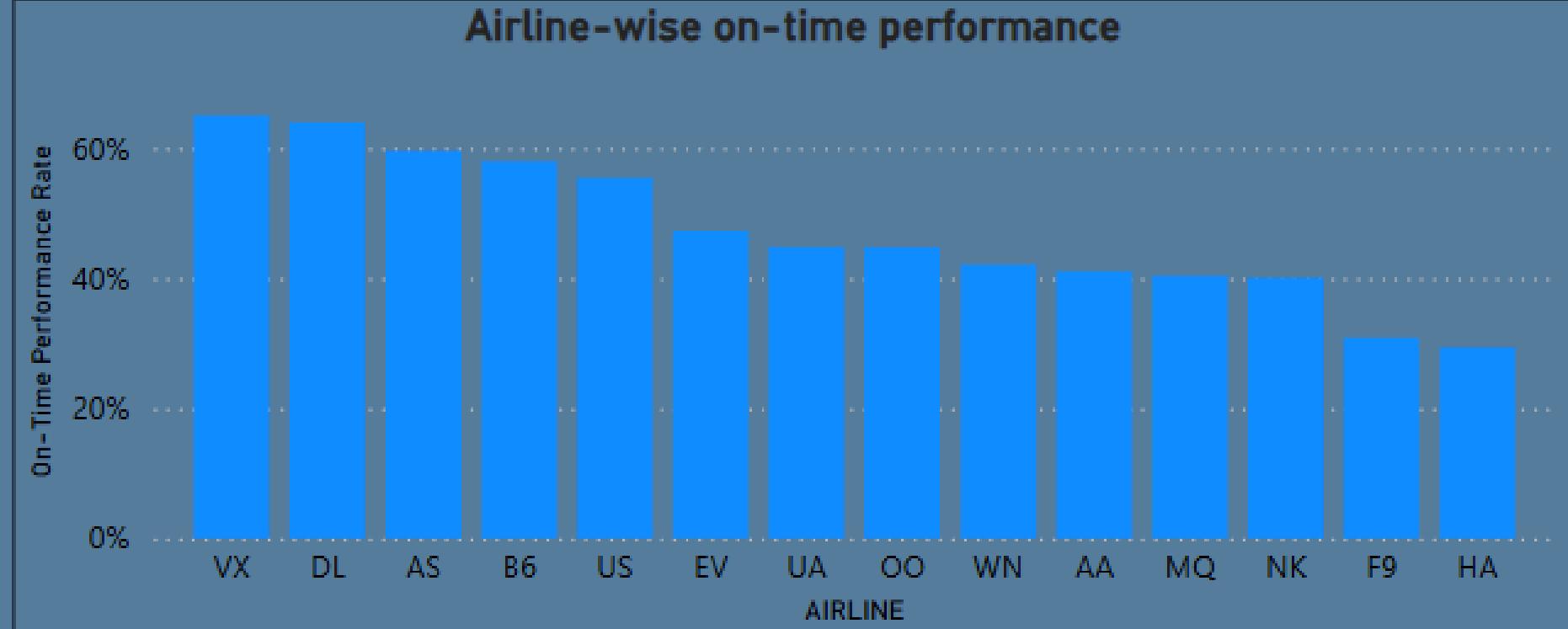
**13.26**

Avg Arrival Delay

Flight Activity Trends Over Time



Airline-wise on-time performance



AIRLINE

All

DATE

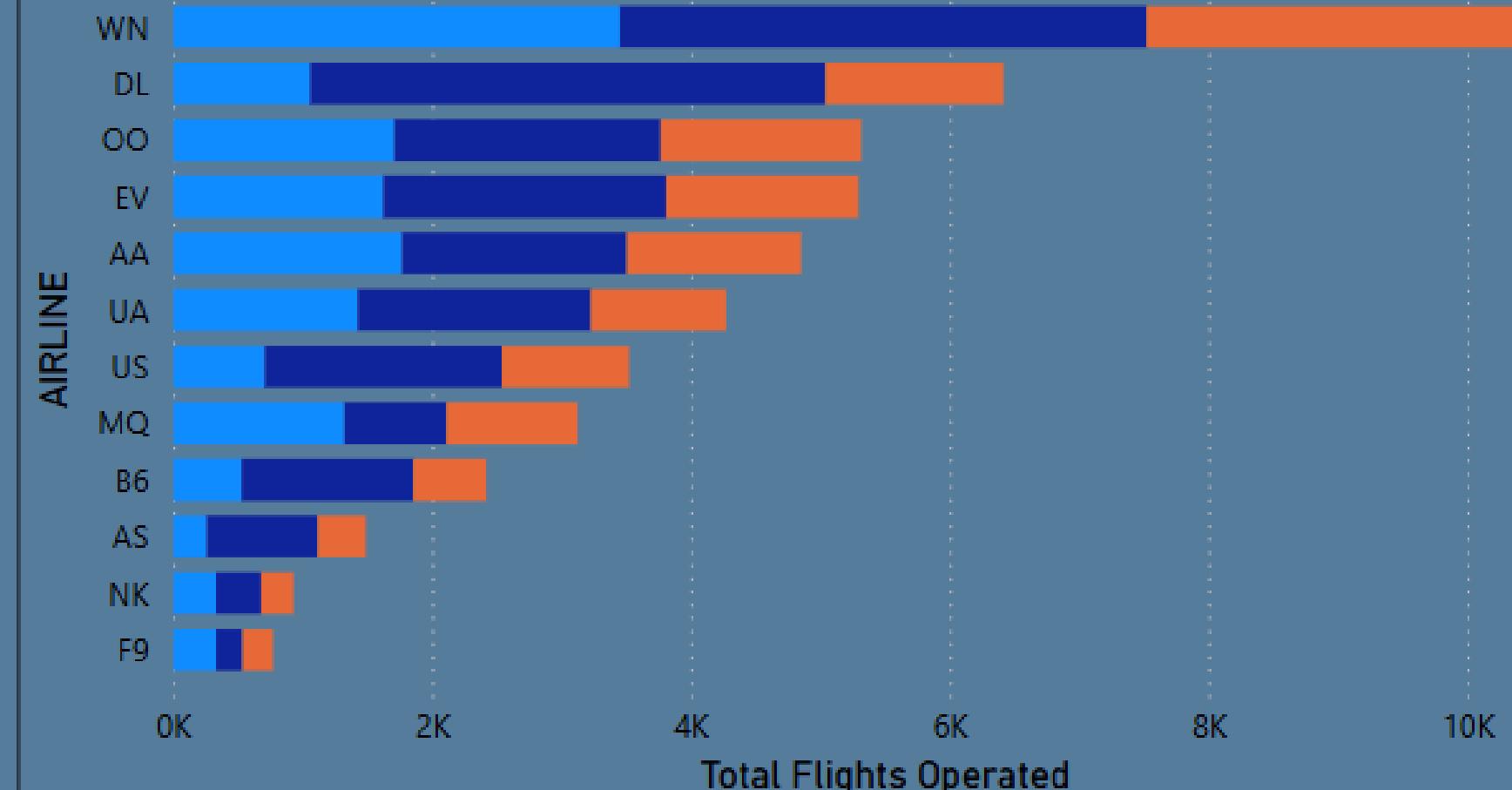
DATE

01-01-2015

04-01-2015

Flight punctuality distribution

FlightPunctualityCategory: Delayed (Blue), Early Arrival (Dark Blue), On-Time (Orange)



# Airline Performance Comparison

ORIGIN\_AIRPORT

AIRLINE

**47.89%**  
Punctuality Score

**2.28%**  
Cancellation Rate

**13.26**  
Avg Arrival Delay

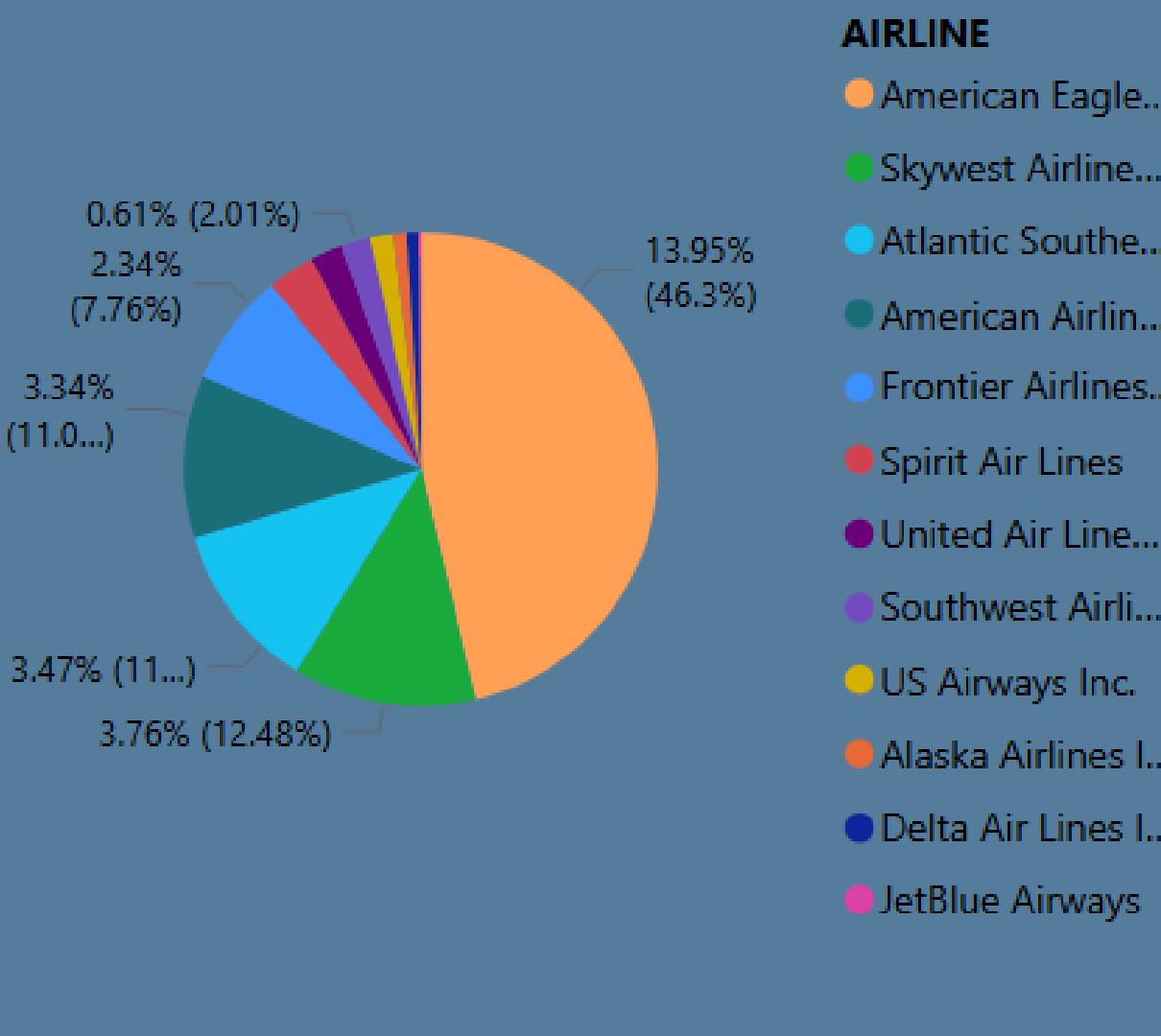
DATE



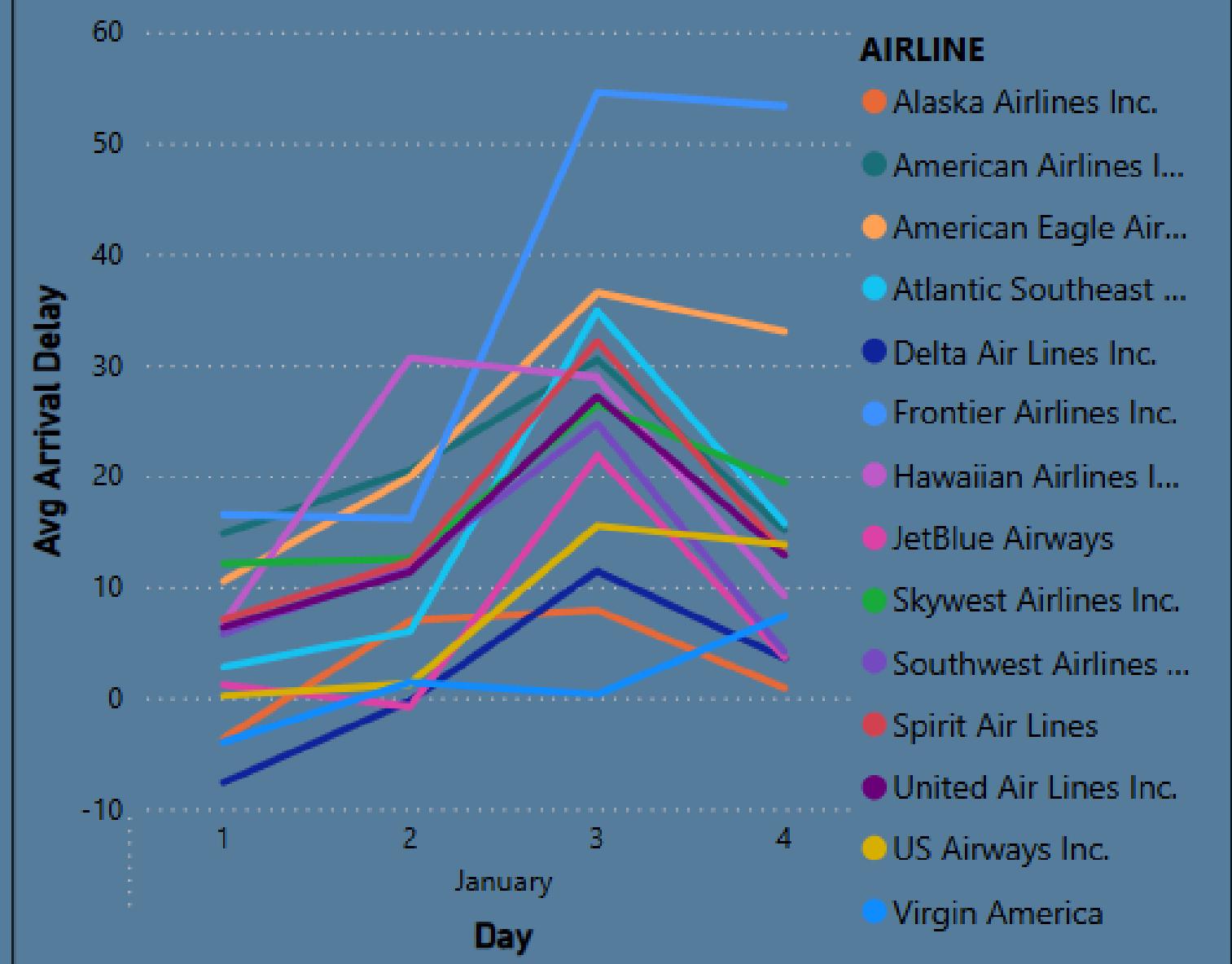
AIR\_SYSTEM\_DELAY



Airline-specific cancellation patterns



Delay trends across airlines



Ranking of airlines based on punctuality

AIRLINE	Punctuality Rank	Punctuality Score
AA	1	41.07%
AS	1	59.60%
B6	1	58.07%
DL	1	64.00%
EV	1	47.30%
F9	1	30.91%
HA	1	29.39%
MQ	1	40.38%
NK	1	40.11%
OO	1	44.79%
UA	1	44.83%
US	1	55.48%
VX	1	65.15%
WN	1	42.18%
<b>Total</b>	<b>6</b>	<b>47.89%</b>

# Route & Airport Analytics

Heatmap of flight movements across airports



Top 10 Busiest Airports

ORIGIN_AIRPORT	Sum of Total Flights
ATL	2864
DEN	1970
DFW	2498
IAH	1479
LAS	1186
LAX	1913
MCO	1125
ORD	2395
PHX	1457
SFO	1461
<b>Total</b>	<b>18348</b>

2.28%

Cancellation Rate

50K

Total Flights Operated

DATE

01-01-2015

04-01-2015

ORIGIN\_AIRPORT

All

AIRPORT

All

Delay trends by airport

AIRPORT

- Aberdeen Regional Airport
- Abilene Regional Airport
- Abraham Lincoln Capital A...
- Adak Airport
- Akron-Canton Regional Ai...
- Albany International Airport
- Albert J. Ellis Airport
- Albuquerque Internation...
- Alexandria International Ai...
- Alpena County Regional A...

Total Avg Delay

100

50

0

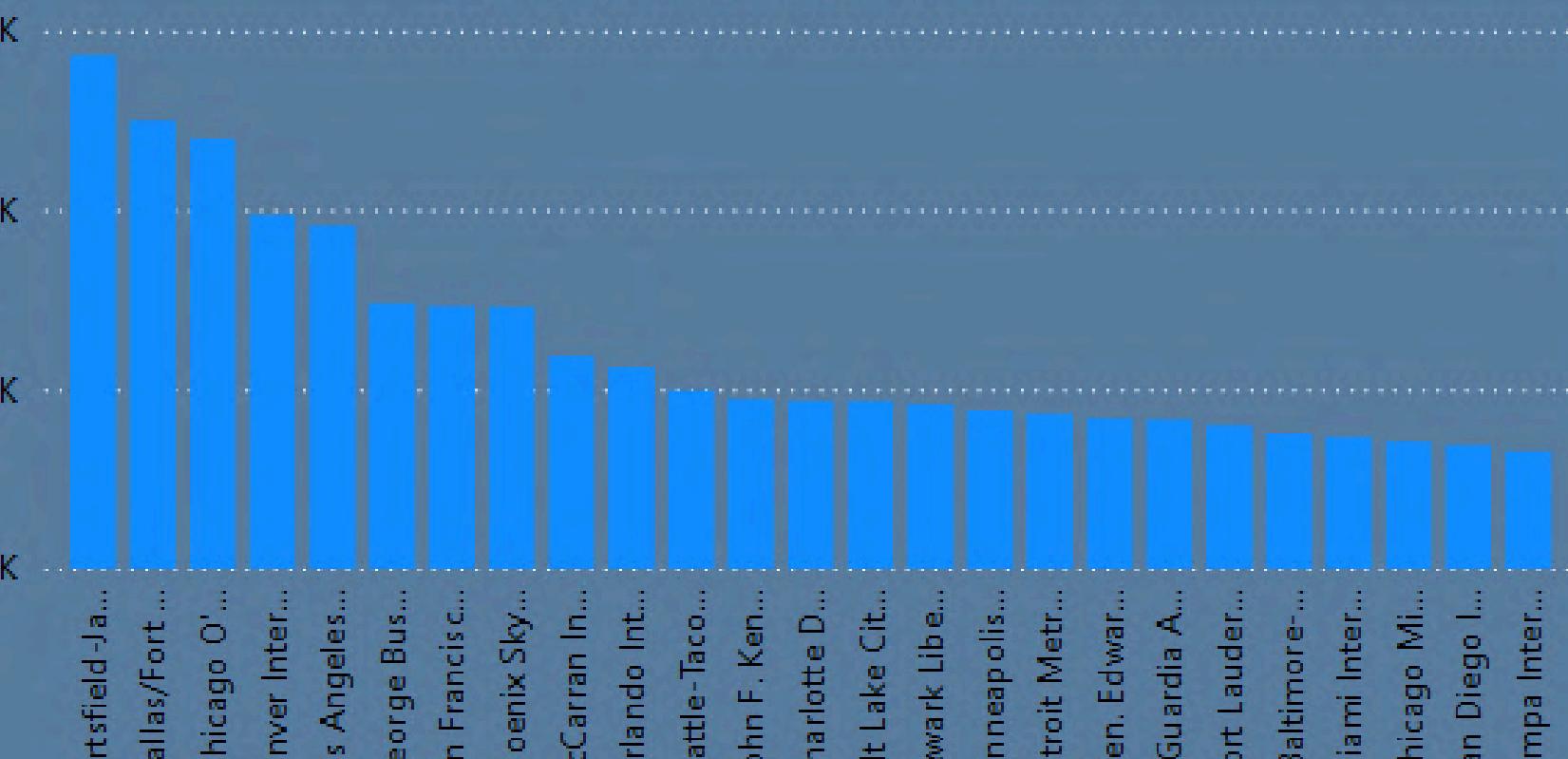
January

Day

Airport Congestion Index

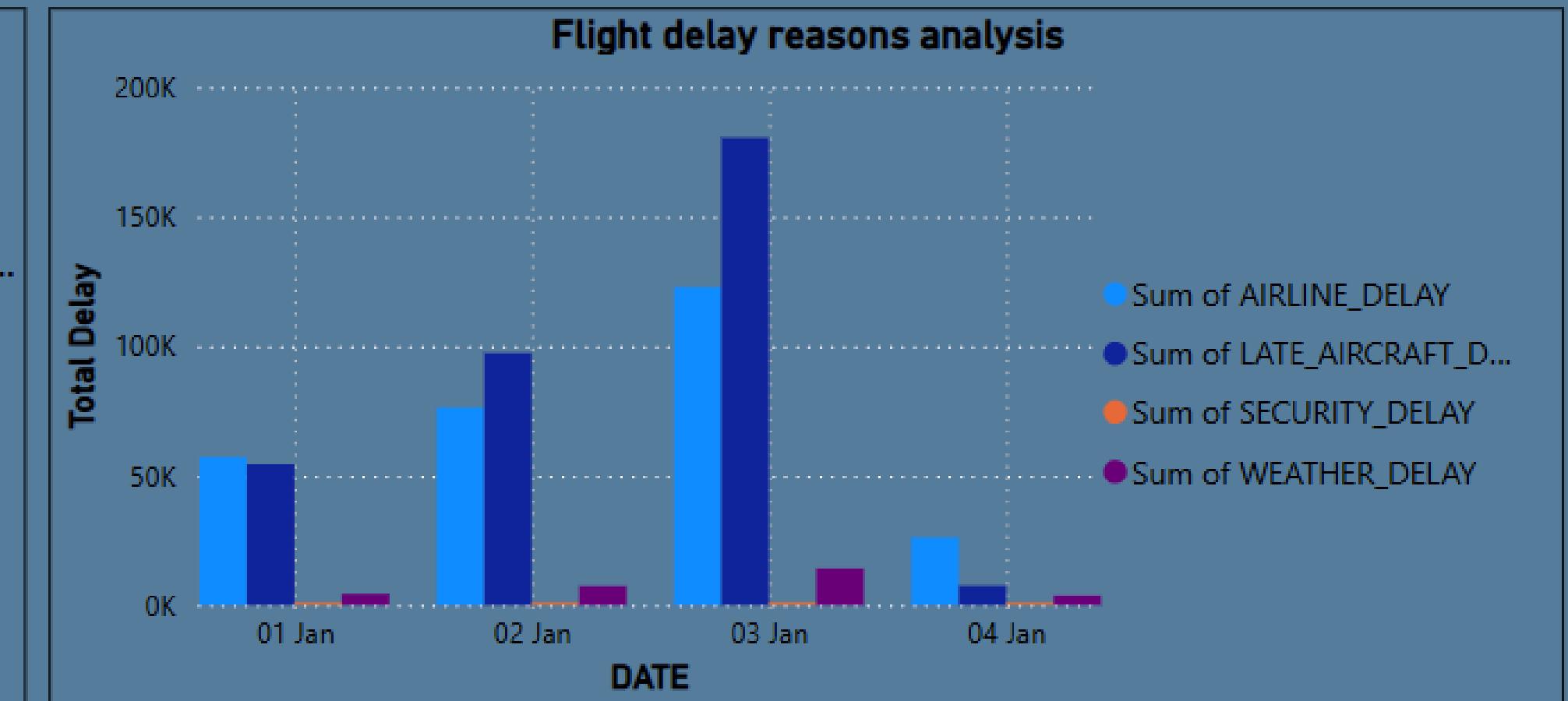
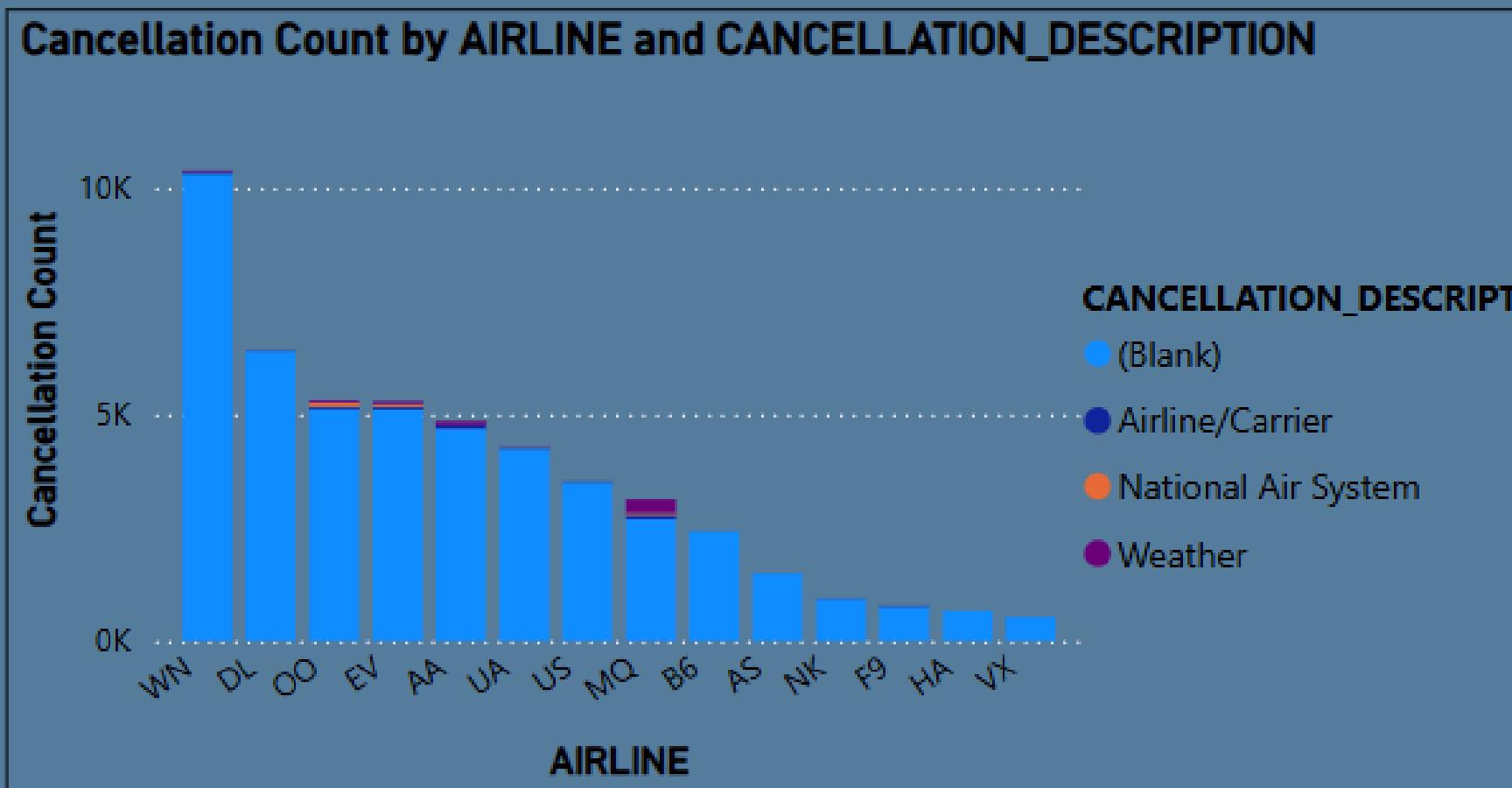
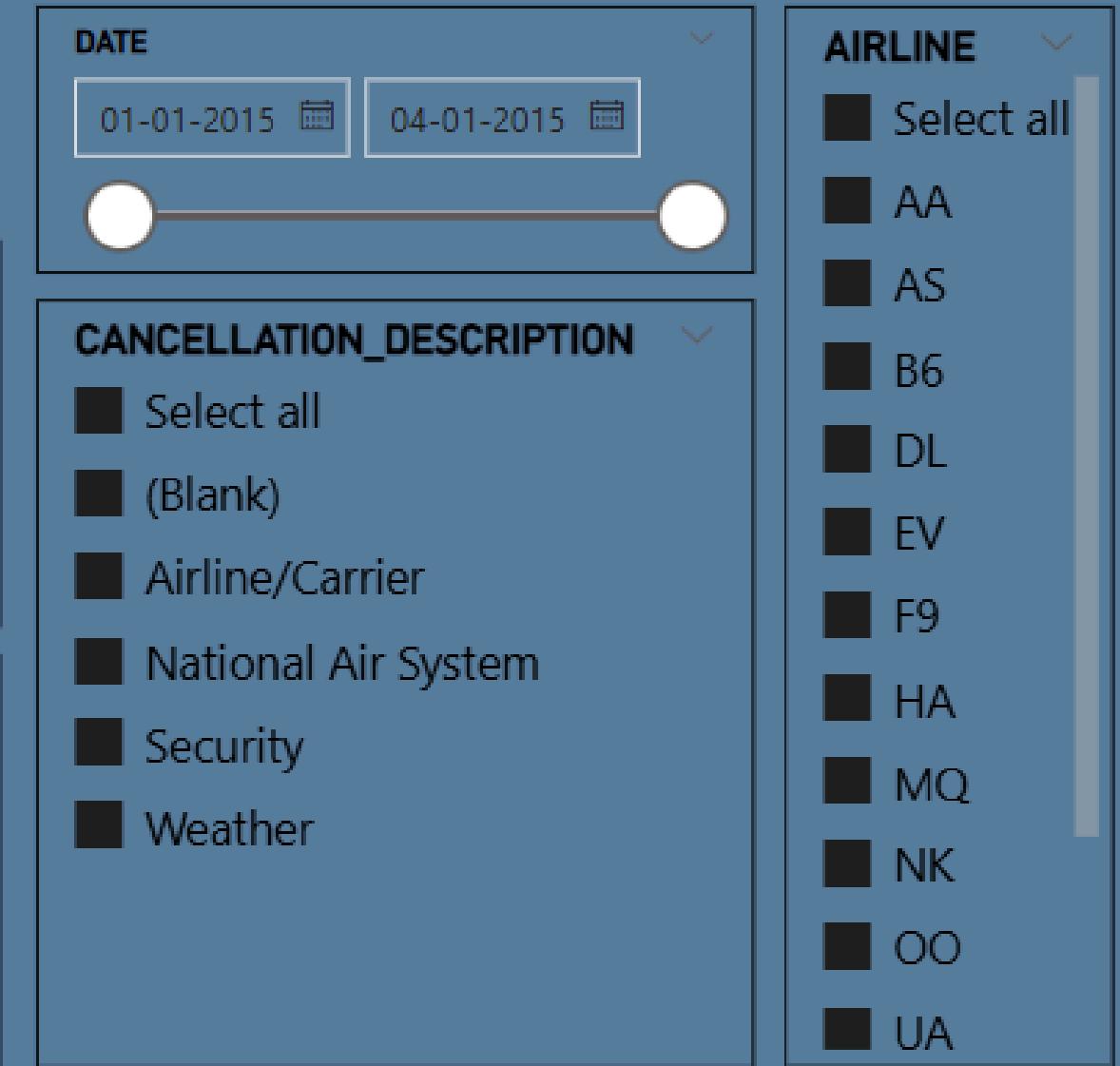
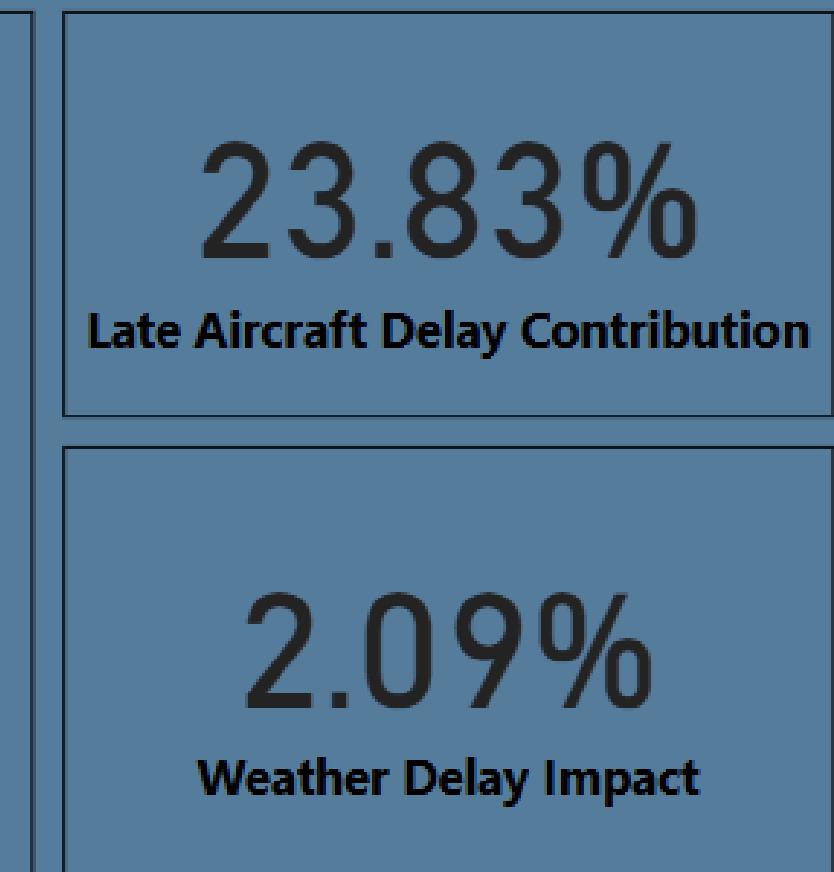
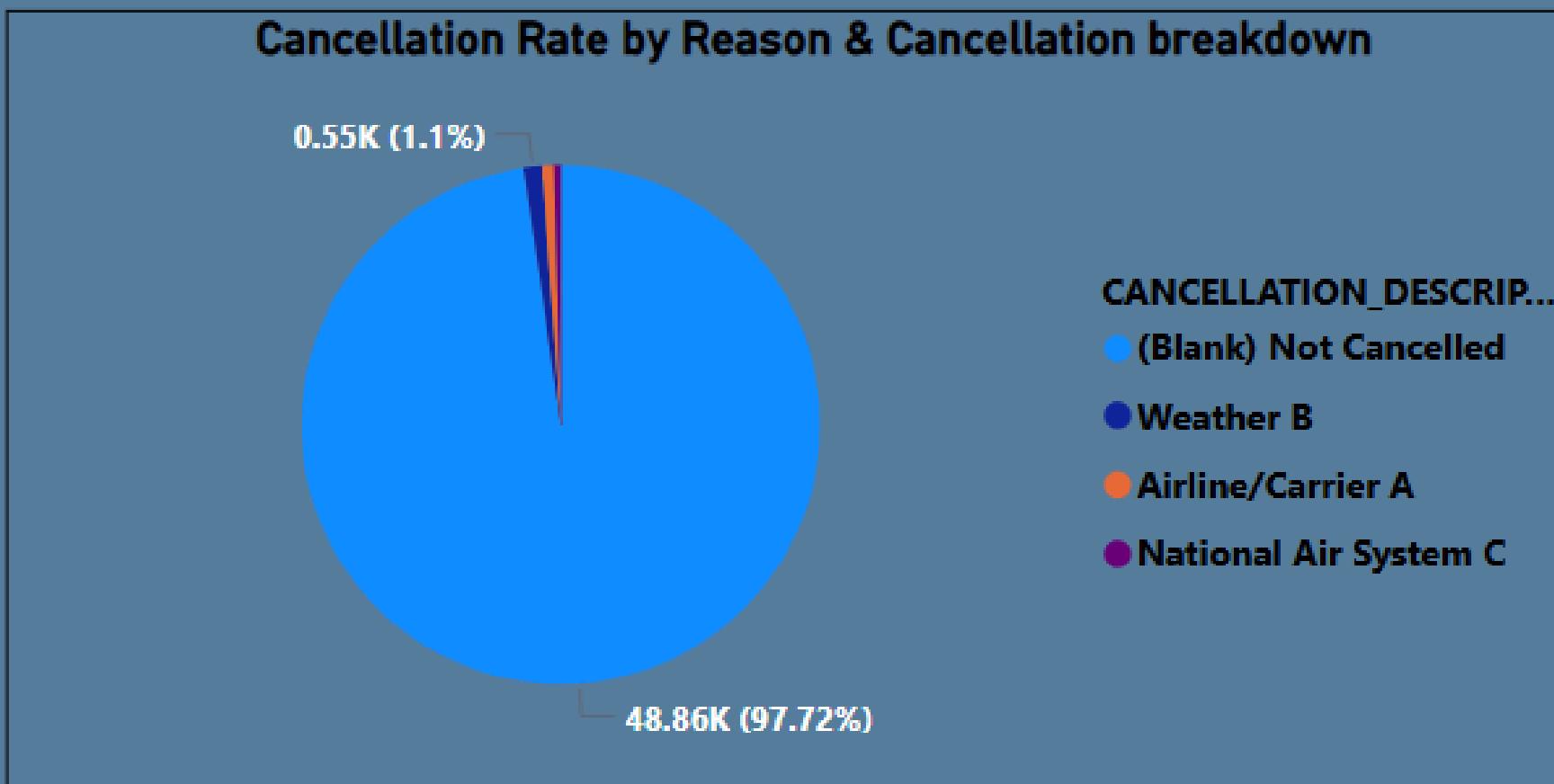
Airport Congestion Index

3K  
2K  
1K  
0K



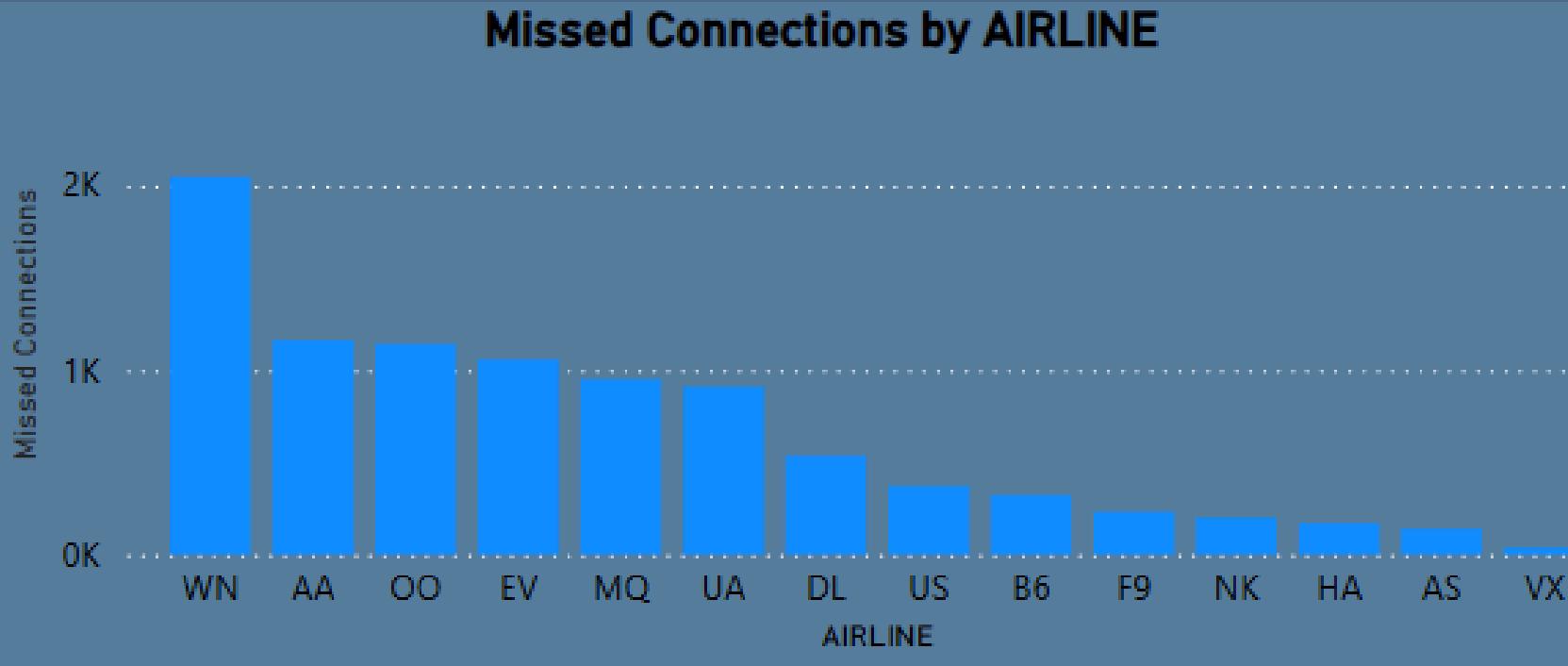
AIRPORT

# Flight Cancellation & Delay Patterns



# Passenger Experience & Service Quality

Missed Connections by AIRLINE



18.58

Missed Connections Rate

13.1%

Delay Recovery Effectiveness

96.91

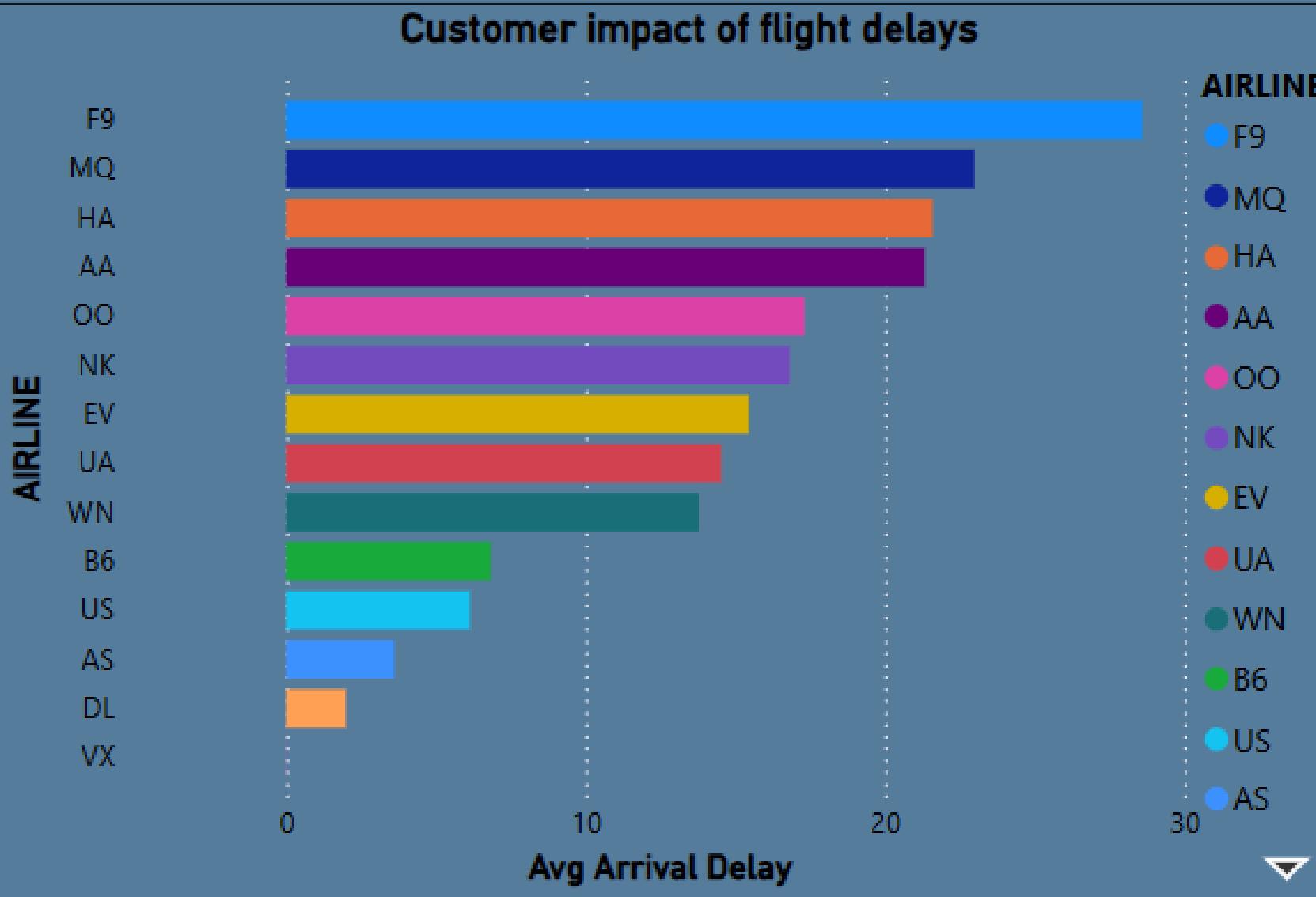
Customer Satisfaction Score

22.11%

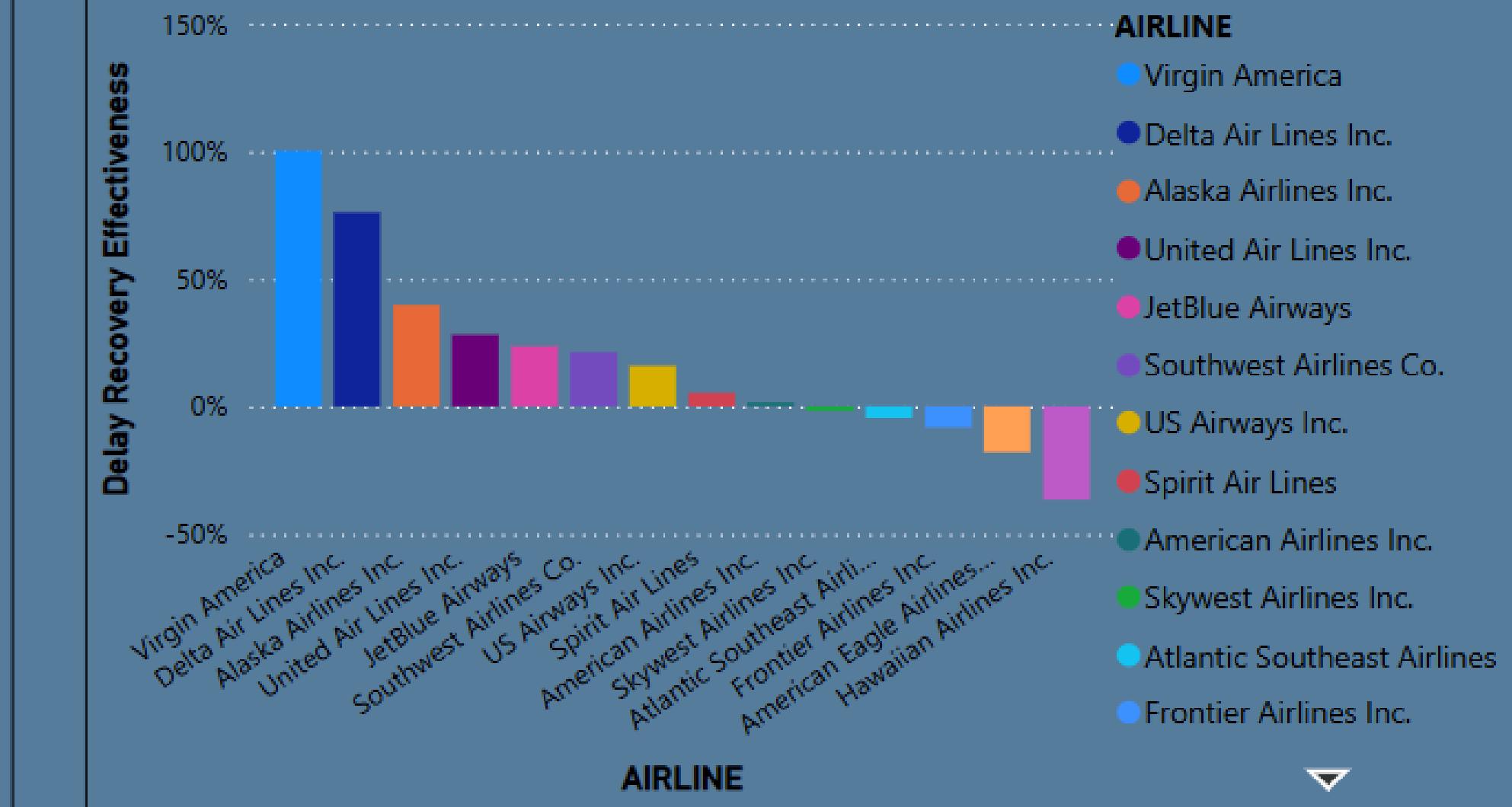
Delay Contribution %

- AIRLINE
- AA
  - AS
  - B6
  - DL
  - EV
  - F9
  - HA

Customer impact of flight delays



Delay Recovery Effectiveness



# Financial & Operational Efficiency

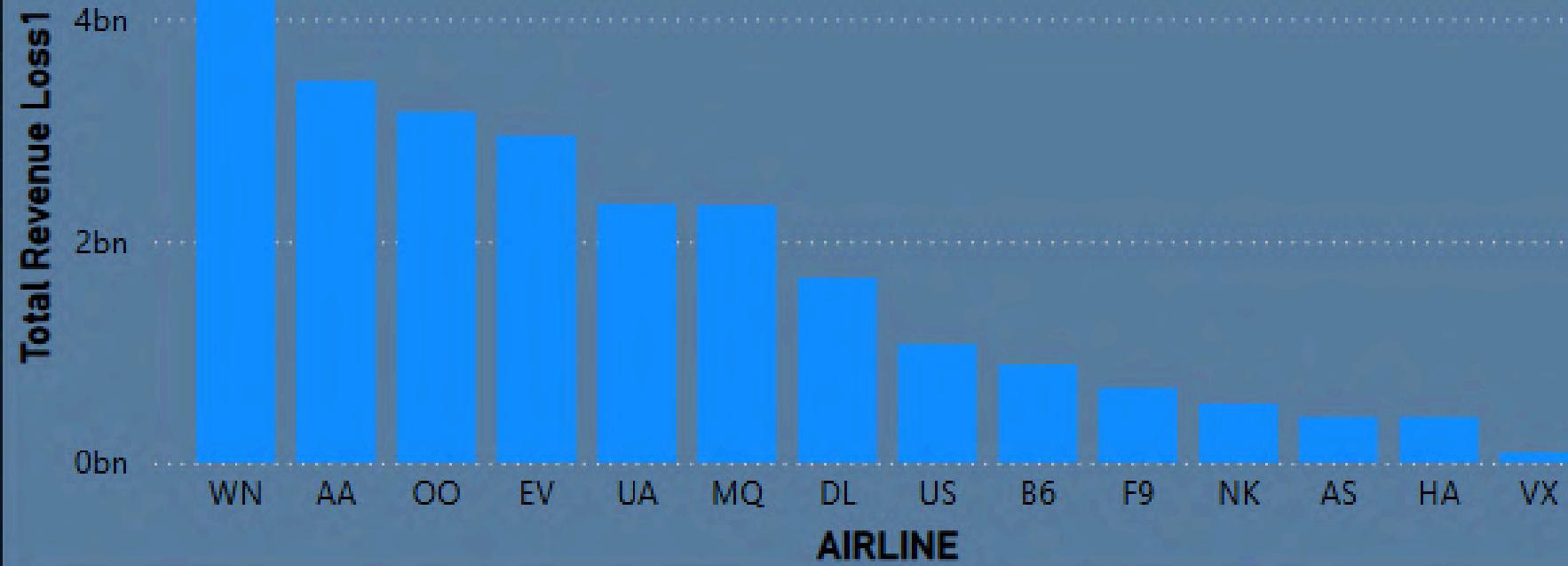
CANCELLATION\_DESCRIPTION

All

AIRLINE

- AA
- AS
- B6
- DL
- EV
- F9
- HA
- MQ
- NK
- OO
- UA
- US

Financial impact of cancellations and delays



\$24.92bn

Total Revenue Loss

\$12.40

Flight Cost per Mile

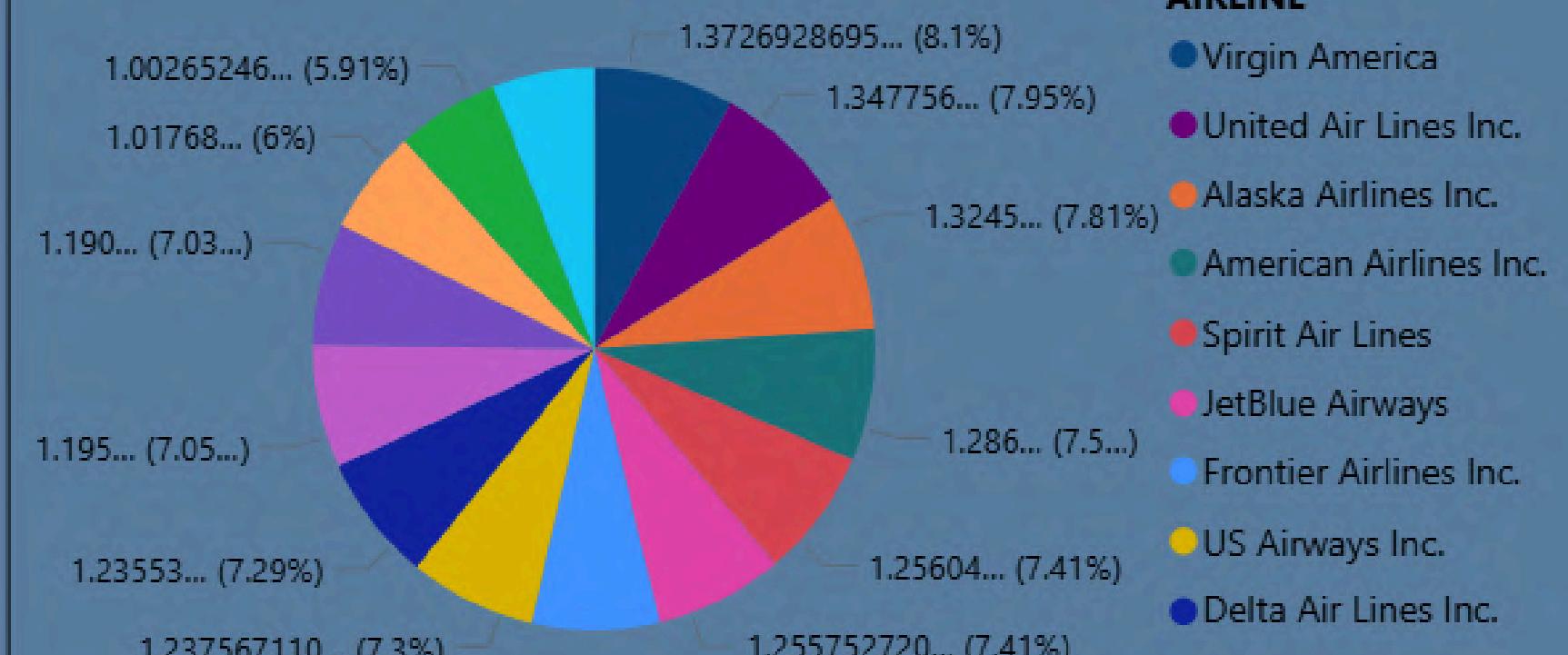
1.21

Fuel Efficiency

Cost efficiency comparison across different routes



Airline fuel and operational efficiency insights



# POWER BI - EXECUTIVE SUMMARY:

The airline industry faces significant challenges in operational efficiency, passenger satisfaction, and financial management. Our analysis across six key areas—Flight Operations, Airline Performance, Route & Airport Analytics, Flight Cancellations & Delays, Passenger Experience, and Financial & Operational Efficiency—highlights critical trends and opportunities for improvement. Flight operations reveal that some airlines excel in on-time performance, while others struggle with delays and congestion at certain airports, impacting overall efficiency. A comparative performance analysis shows disparities in flight completion rates and punctuality, suggesting the need for underperforming airlines to benchmark against industry leaders. Route and airport analytics further emphasize the importance of demand-driven scheduling and strategic airport selection to optimize efficiency.

Flight cancellations and delays remain a major concern, with weather, technical issues, and congestion contributing to disruptions. These delays not only inconvenience passengers but also impact airline reputation and revenue. The passenger experience analysis indicates that while the industry maintains a strong overall customer satisfaction score of 96.91, missed connections and inconsistent delay recovery effectiveness continue to affect service quality. Financially, delays and cancellations have led to revenue losses of \$24.92 billion, with fuel efficiency and operational costs varying significantly across airlines. To address these issues, airlines must focus on optimizing scheduling, improving communication with passengers, reducing operational costs, and leveraging AI-driven analytics to enhance decision-making. By implementing these strategies, airlines can improve performance, boost customer trust, and ensure long-term profitability in an increasingly competitive market.

# Let's Connect !!

I hope you found these insights valuable.

Feel free to connect with me on [Linkedin](#) to explore more about my projects, share feedback, or discuss new opportunities:

LinkedIn Profile: 

Github Profile: 



Or drop me a message directly at:

Email: [bhuvanjarioo1@gmail.com](mailto:bhuvanjarioo1@gmail.com)

Together, let's turn data into meaningful stories!