



# FLY EMIRATES AIRLINE PERFORMANCE ANALYSIS – 2015

**Prepared by:** Ronak Patel

**Date:** July 2025

---

## 1. Introduction

This project analyzes the performance of Emirates Airlines using the 2015 U.S. domestic flight dataset. The goal is to evaluate operational efficiency, delay patterns, cancellation trends, and provide actionable insights to enhance airline and airport performance. The dashboard enables key stakeholders to make data-driven decisions that improve reliability, customer satisfaction, and logistics planning.

---

## 2. Methodology

### Data Sources:

- flights.csv: Core flight data including schedule, delays, cancellations
- airlines.csv: Airline IATA codes and names
- airports.csv: Airport names, locations, and codes

### Tools Used:

- SQL (PostgreSQL) for data extraction and transformation
- Power BI for dashboard creation and interactive visuals

### Data Preparation Highlights:

- Converted scheduled and actual time fields to proper time formats
  - Mapped cancellation reasons (A/B/C/D) to labels (Carrier, Weather, etc.)
  - Integrated airline and airport metadata into main flight table
  - Removed nulls and blanks from critical fields (airline, airport, delay)
-

### 3. Project Workflow: SQL → Power BI

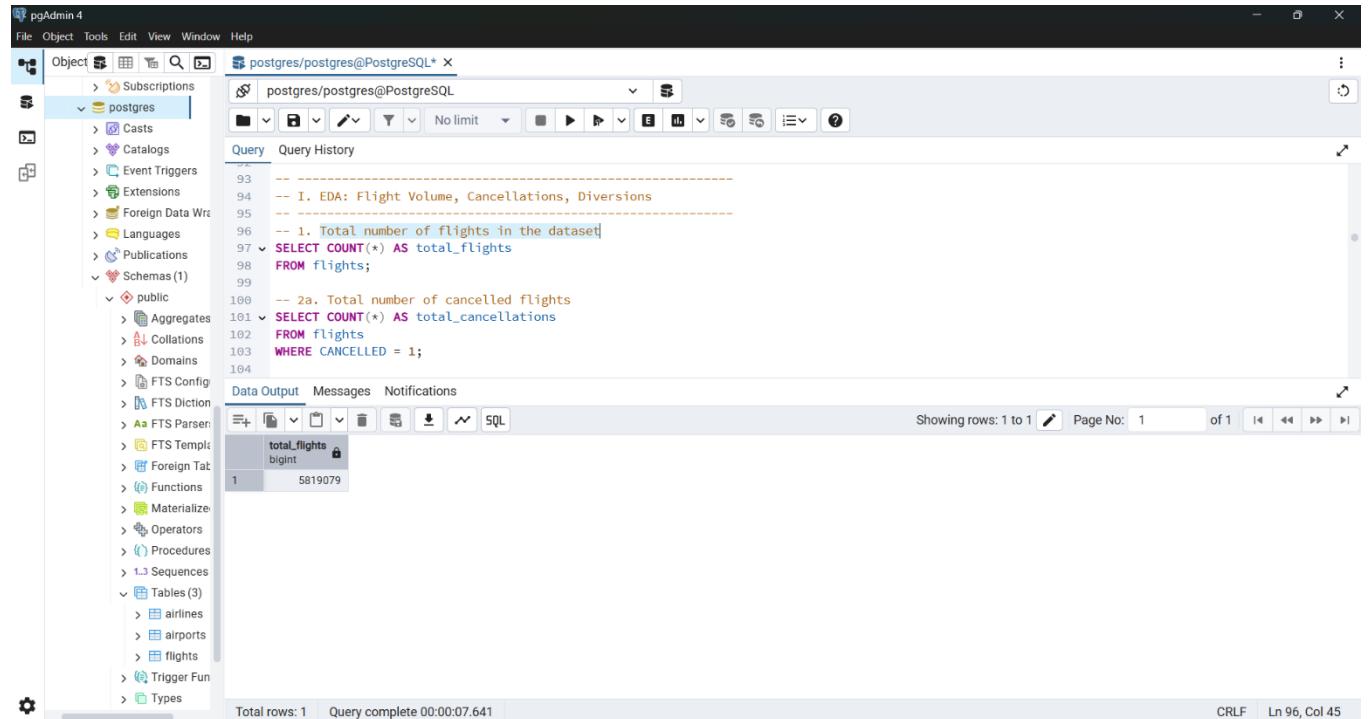
Phase	Task	Tool
1–2	Data Cleaning, Prep, SQL KPIs	PostgreSQL
3–4	Dashboard Development	Power BI
5–6	Insight Generation + Reporting	Word/PDF, PowerPoint

#### Steps:

1. Clean and analyze the data using SQL (join tables, calculate KPIs, handle nulls)
2. Load final output into Power BI for visual analysis
3. Create dashboards with slicers, DAX measures, and visuals
4. Generate insights and present findings in a structured report and slide deck

## 4. SQL Queries

### 4a. Total number of flights in the dataset



The screenshot shows the pgAdmin 4 interface. On the left is the object browser tree, which includes the 'postgres' database node with its sub-objects like 'public', 'airlines', 'airports', 'flights', and 'Trigger Fun'. The main window contains a SQL query editor with the following code:

```
-- I. EDA: Flight Volume, Cancellations, Diversions
-- 1. Total number of flights in the dataset
SELECT COUNT(*) AS total_flights
FROM flights;

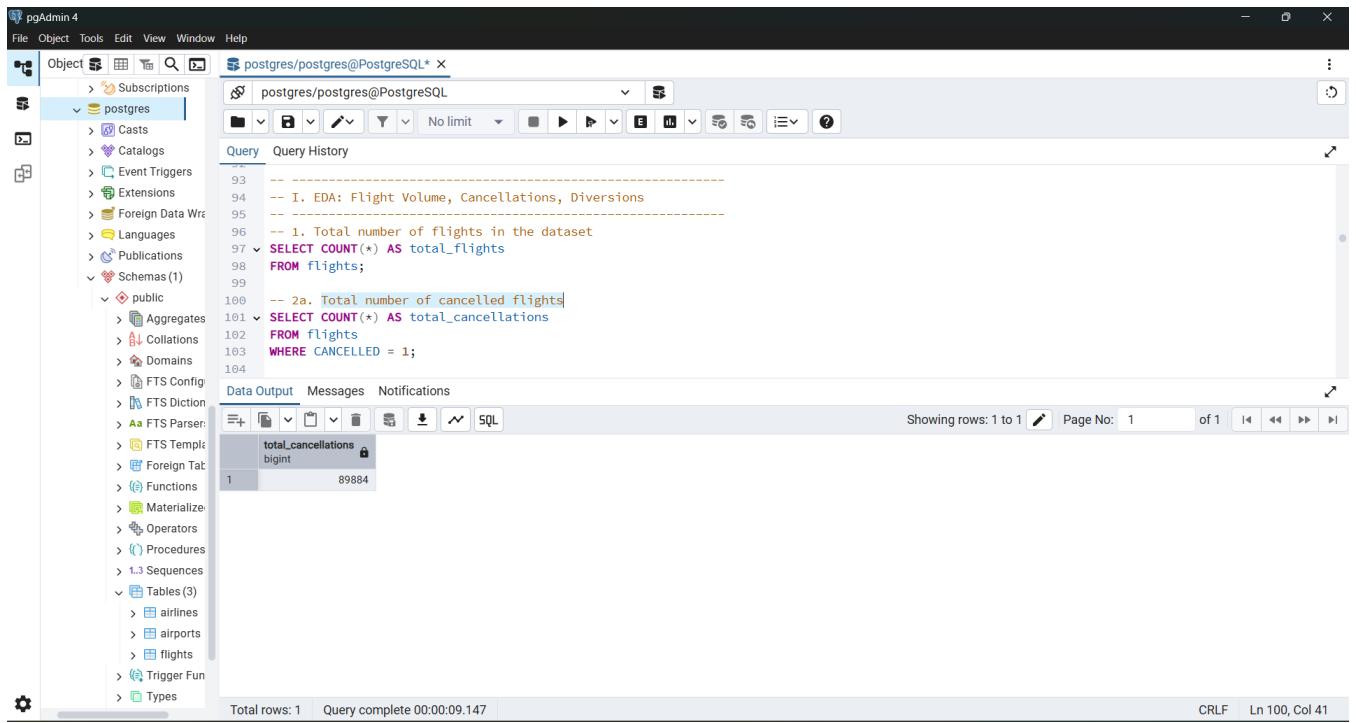
-- 2a. Total number of cancelled flights
SELECT COUNT(*) AS total_cancellations
FROM flights
WHERE CANCELLED = 1;
```

The results pane shows a single row of data:

total_flights	bigrint
1	5819079

At the bottom of the pgAdmin window, status bars show 'Total rows: 1' and 'Query complete 00:00:07.641'.

## 4b. Total number of cancelled flights

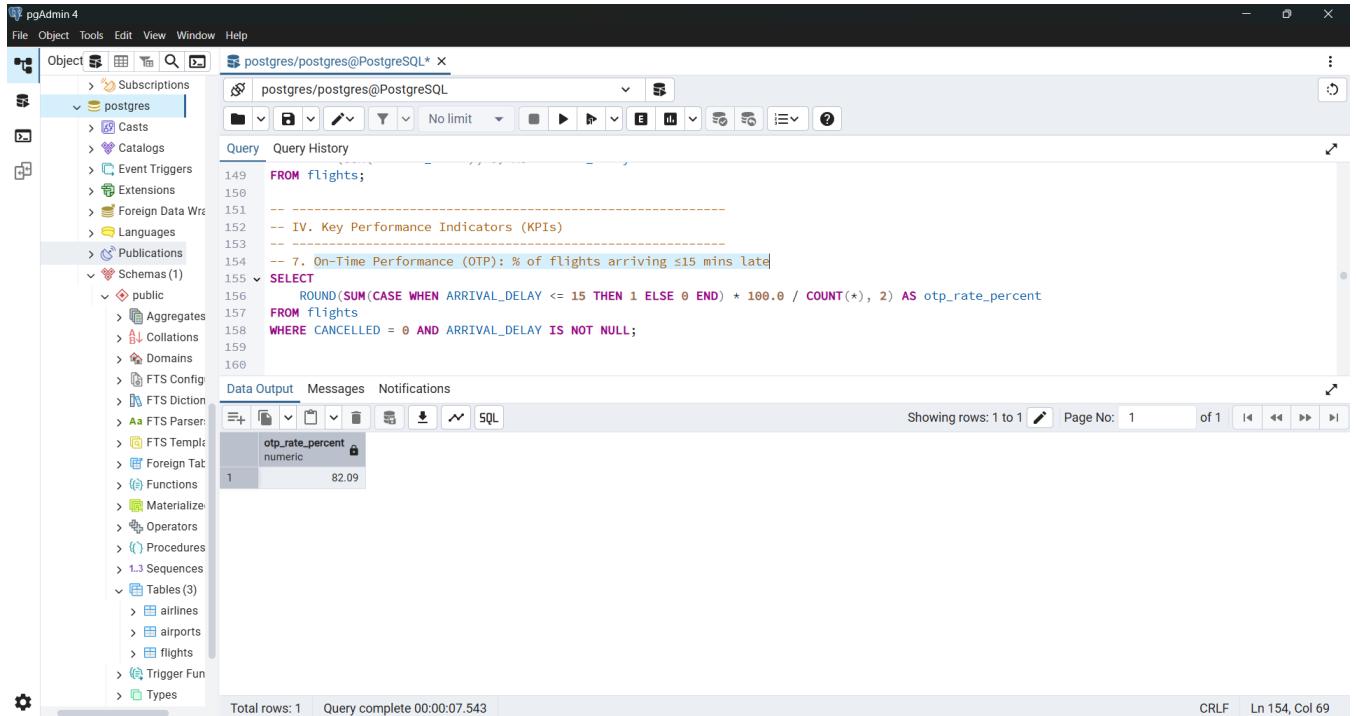


The screenshot shows the pgAdmin 4 interface with the 'Query' tab selected. The left sidebar shows the database schema with the 'flights' table highlighted. The main window displays the following SQL code:

```
93 -- -----
94 -- I. EDA: Flight Volume, Cancellations, Diversions
95 --
96 -- 1. Total number of flights in the dataset
97 SELECT COUNT(*) AS total_flights
98 FROM flights;
99
100 -- 2a. Total number of cancelled flights
101 SELECT COUNT(*) AS total_cancellations
102 FROM flights
103 WHERE CANCELLED = 1;
104
```

The results table shows one row with the value 89884 for 'total\_cancellations'. The status bar at the bottom indicates 'Query complete 00:00:09.147'.

## 4c. On-Time Performance (OTP): % of flights arriving ≤15 mins late

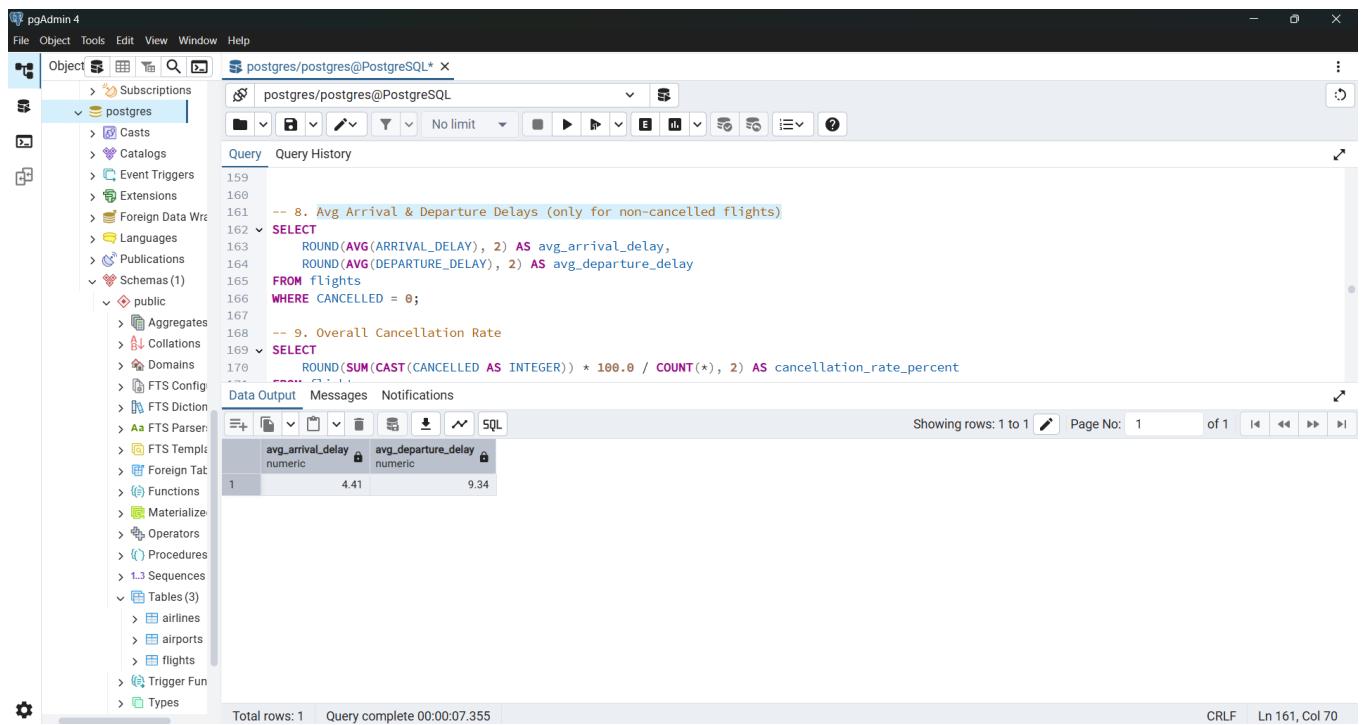


The screenshot shows the pgAdmin 4 interface with the 'Query' tab selected. The left sidebar shows the database schema with the 'flights' table highlighted. The main window displays the following SQL code:

```
149 FROM flights;
150
151 -- -----
152 -- IV. Key Performance Indicators (KPIs)
153 --
154 -- 7. On-Time Performance (OTP): % of flights arriving ≤15 mins late
155 SELECT
156     ROUND(SUM(CASE WHEN ARRIVAL_DELAY <= 15 THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS otp_rate_percent
157 FROM flights
158 WHERE CANCELLED = 0 AND ARRIVAL_DELAY IS NOT NULL;
159
160
```

The results table shows one row with the value 82.09 for 'otp\_rate\_percent'. The status bar at the bottom indicates 'Query complete 00:00:07.543'.

#### 4d. Avg Arrival & Departure Delays (only for non-cancelled flights)



```

-- 8. Avg Arrival & Departure Delays (only for non-cancelled flights)
SELECT
    ROUND(AVG(ARRIVAL_DELAY), 2) AS avg_arrival_delay,
    ROUND(AVG(DEPARTURE_DELAY), 2) AS avg_departure_delay
FROM flights
WHERE CANCELLED = 0;

-- 9. Overall Cancellation Rate
SELECT
    ROUND(SUM(CAST(CANCELLED AS INTEGER)) * 100.0 / COUNT(*), 2) AS cancellation_rate_percent

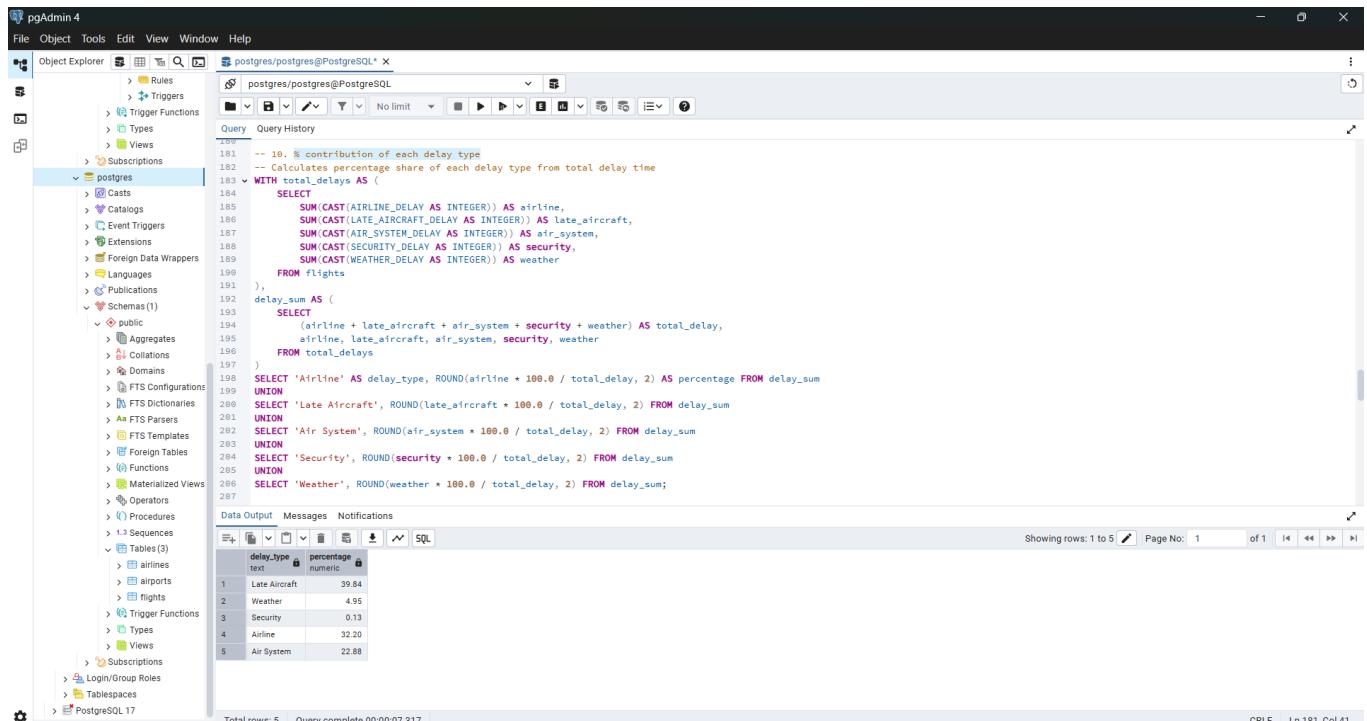
```

The screenshot shows the pgAdmin 4 interface with the query results displayed in the Data Output tab. The results are as follows:

	avg_arrival_delay	avg_departure_delay
1	4.41	9.34

Total rows: 1 Query complete 00:00:07.355 CRLF Ln 161, Col 70

#### 4e. % contribution of each delay type



```

-- 10. % contribution of each delay type
-- Calculates percentage share of each delay type from total delay time
WITH total_delays AS (
    SELECT
        SUM(CAST(AIRLINE_DELAY AS INTEGER)) AS airline,
        SUM(CAST(LATE_AIRCRAFT_DELAY AS INTEGER)) AS late_aircraft,
        SUM(CAST(ATR_SYSTEM_DELAY AS INTEGER)) AS air_system,
        SUM(CAST(SECURITY_DELAY AS INTEGER)) AS security,
        SUM(CAST(WEATHER_DELAY AS INTEGER)) AS weather
    FROM flights
),
delay_sum AS (
    SELECT
        (airline + late_aircraft + air_system + security + weather) AS total_delay,
        airline, late_aircraft, air_system, security, weather
    FROM total_delays
)
SELECT
    'Airline' AS delay_type, ROUND(airline * 100.0 / total_delay, 2) AS percentage
FROM delay_sum
UNION
SELECT
    'Late Aircraft', ROUND(late_aircraft * 100.0 / total_delay, 2) FROM delay_sum
UNION
SELECT
    'Air System', ROUND(air_system * 100.0 / total_delay, 2) FROM delay_sum
UNION
SELECT
    'Security', ROUND(security * 100.0 / total_delay, 2) FROM delay_sum
UNION
SELECT
    'Weather', ROUND(weather * 100.0 / total_delay, 2) FROM delay_sum;

```

The screenshot shows the pgAdmin 4 interface with the query results displayed in the Data Output tab. The results are as follows:

delay_type	percentage
Late Aircraft	39.84
Weather	4.95
Security	0.13
Airline	32.20
Air System	22.88

Total rows: 5 Query complete 00:00:07.317 CRLF Ln 181, Col 41

## 4f. KPI Breakdown by Airline:

pgAdmin 4

File Object Tools Edit View Window Help

Object Explorer

postgres/postgres@PostgreSQL\*

```

209 -- V. Aggregation by Airline, Airport, Time
210
211 -- 11. KPI Breakdown by Airlines
212 -- Shows: OTP %, Avg Arrival Delay, Avg Departure Delay, Cancellation Rate
213 <-- SEL
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232

```

AIRLINE	otp_percent	avg_arrival_delay	avg_departure_delay	cancellation_percent
HA	89.20	2.02	0.49	0.22
AS	87.11	-0.98	1.79	0.39
DL	86.45	0.19	7.37	0.44
AA	86.93	3.43	8.90	1.00
VX	86.61	4.74	9.02	0.88
WN	86.44	4.37	10.58	1.27
OO	86.39	5.85	7.80	1.49
US	86.16	3.71	6.14	2.05
UA	78.81	5.43	14.44	1.27
EV	78.37	6.59	8.72	2.66
BB	76.68	6.68	11.51	1.00
MQ	74.73	6.46	10.13	5.10
F9	74.02	12.50	13.35	0.65
NK	69.88	14.47	15.94	1.71

Total rows: 14 Query complete 00:00:08.024

CRLF Ln 211, Col 33

## 4g. KPIs by Month

pgAdmin 4

File Object Tools Edit View Window Help

Object Explorer

postgres/postgres@PostgreSQL\*

```

248 -- 13. KPIs by Month
249 <-- SEL
250
251
252
253
254
255
256
257

```

month	avg_arrival_delay	avg_departure_delay	cancellation_percent
1	5.81	9.76	2.55
2	8.32	11.89	4.78
3	4.92	9.66	2.18
4	3.16	7.72	0.93
5	4.49	9.45	1.15
6	9.60	13.99	1.81
7	6.43	11.39	0.92
8	4.61	9.93	0.99
9	-0.77	4.82	0.45
10	-0.78	4.98	0.50
11	1.10	6.94	0.98
12	6.09	11.78	1.68

Total rows: 12 Query complete 00:00:06.421

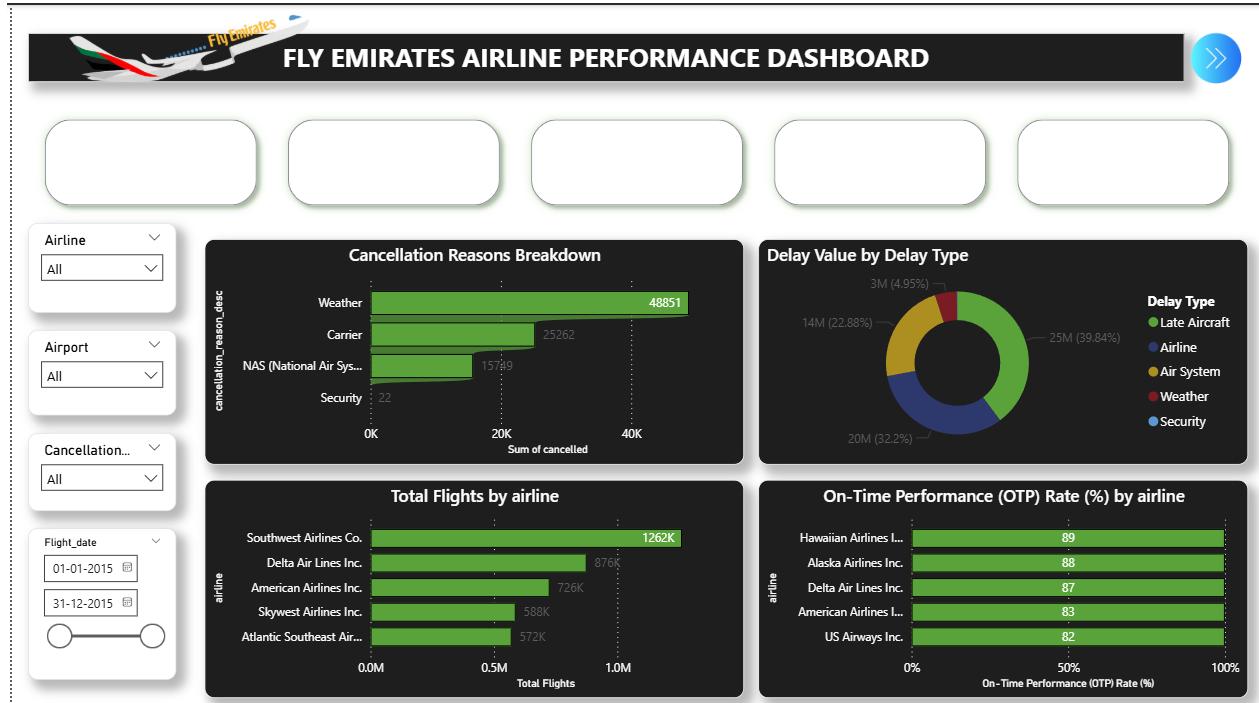
CRLF Ln 248, Col 8

## 5. Dashboard Overview

### Page 1: Overview Dashboard

Key KPIs: Total Flights, On-Time Rate, Avg Delay, Cancellation Rate

Delay by cause, cancellation by reason, top airlines by volume



### Key Insights – Page 1: Overview Dashboard (Summary)

#### 1. On-Time Rate at 82.4%

Industry-wide performance is moderate; Hawaiian & Alaska Airlines set the benchmark (>88%).

#### 2. Weather & Carrier = Top Cancellation Reasons

Weather caused ~49K cancellations, followed by ~25K due to carrier-related issues.

#### 3. Late Aircraft Drives Most Delays

Accounts for ~40% of total delay minutes — highlights turnaround inefficiencies.

#### 4. Southwest Airlines = Highest Volume

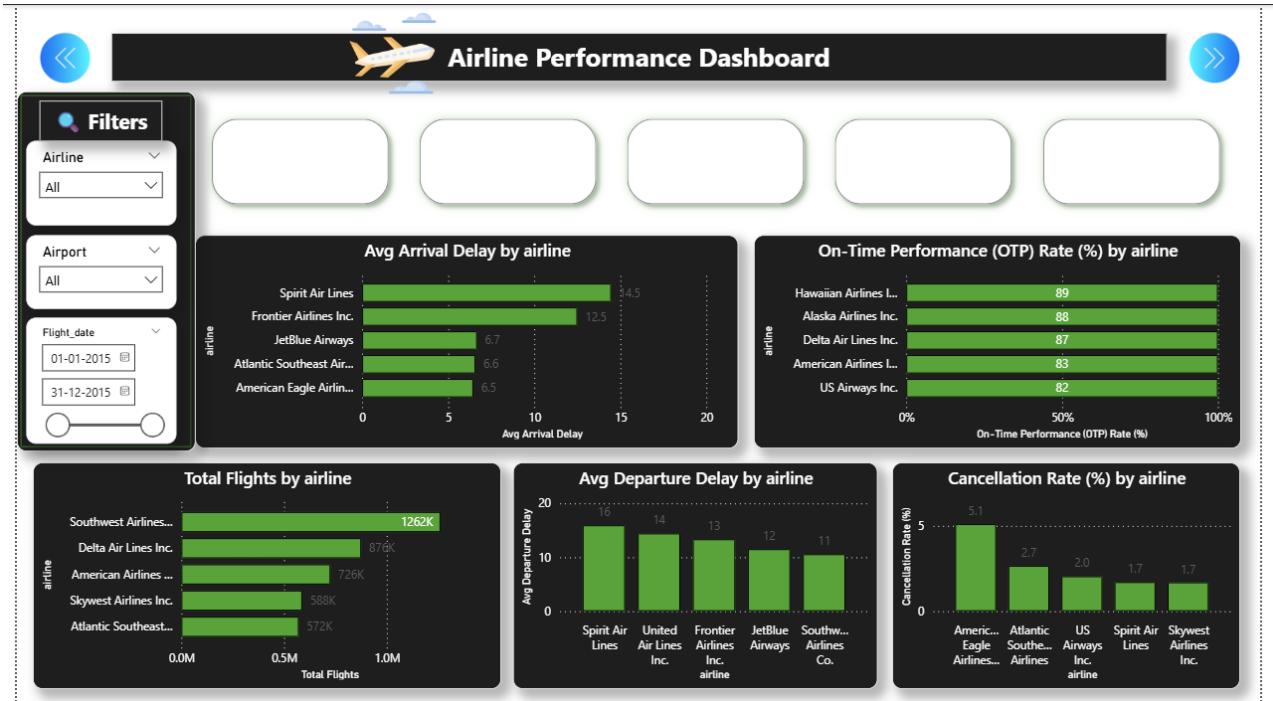
Operated 1.26M flights, showing significant exposure to delays and cancellations.

#### 5. Departure Delays Worse than Arrival

Avg departure delay: 9.37 mins vs. arrival delay: 4.41 mins — suggests ground-level issues.

## Page 2: Airline Performance

Comparison of airlines by average arrival/departure delay, total flights, and cancellation rates

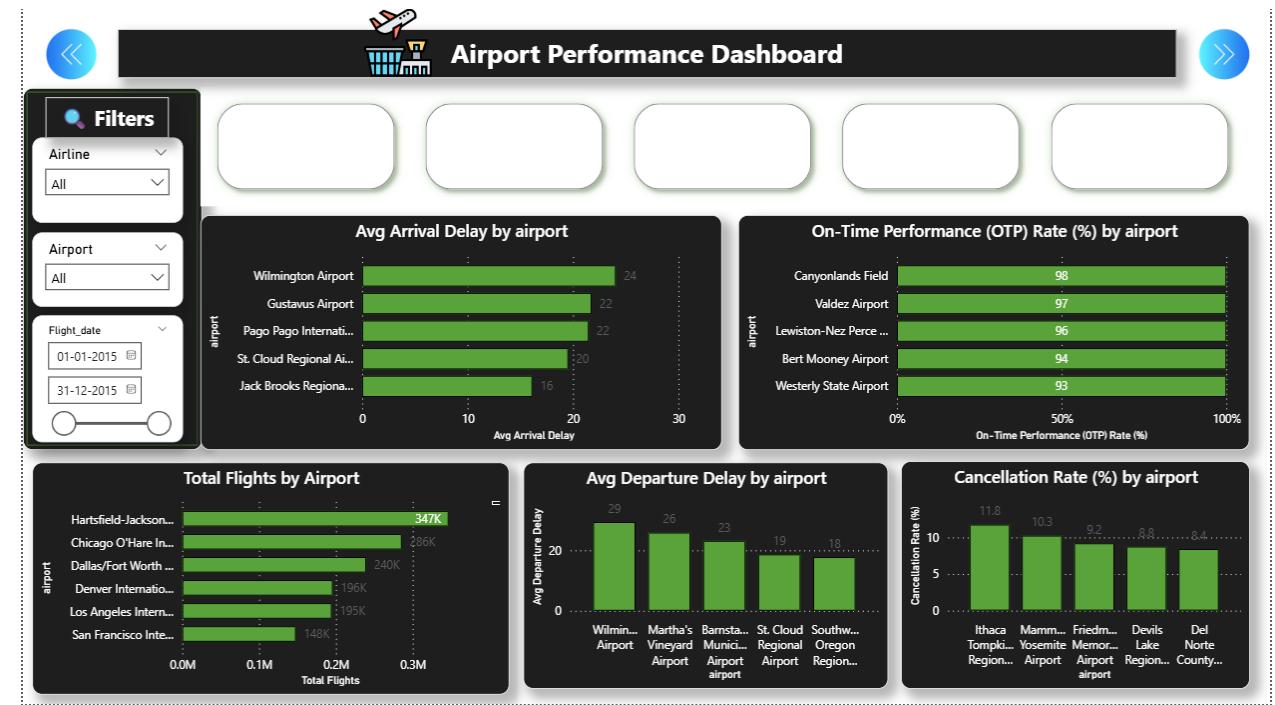


### Key Insights – Page 2: Airline Performance (Summary)

- Spirit & Frontier = Worst Delay Offenders**  
Spirit: Avg Arrival 14.5 mins, Departure 16 mins — Frontier close behind.  
Indicates poor turnaround efficiency.
- Southwest Leads in Flight Volume**  
Operated 1.26M flights — ~45% more than Delta — yet maintains low delays, showcasing strong operations.
- American Eagle = High Cancellation Rate**  
5.1% cancellations — nearly 2x higher than peers — signals major reliability issues.
- JetBlue & Delta Perform Consistently Well**  
Balanced metrics across delays and OTP — Delta especially shows scalable best practices.

## Page 3: Airport Performance

Average delays and cancellation rates across major and regional U.S. airports

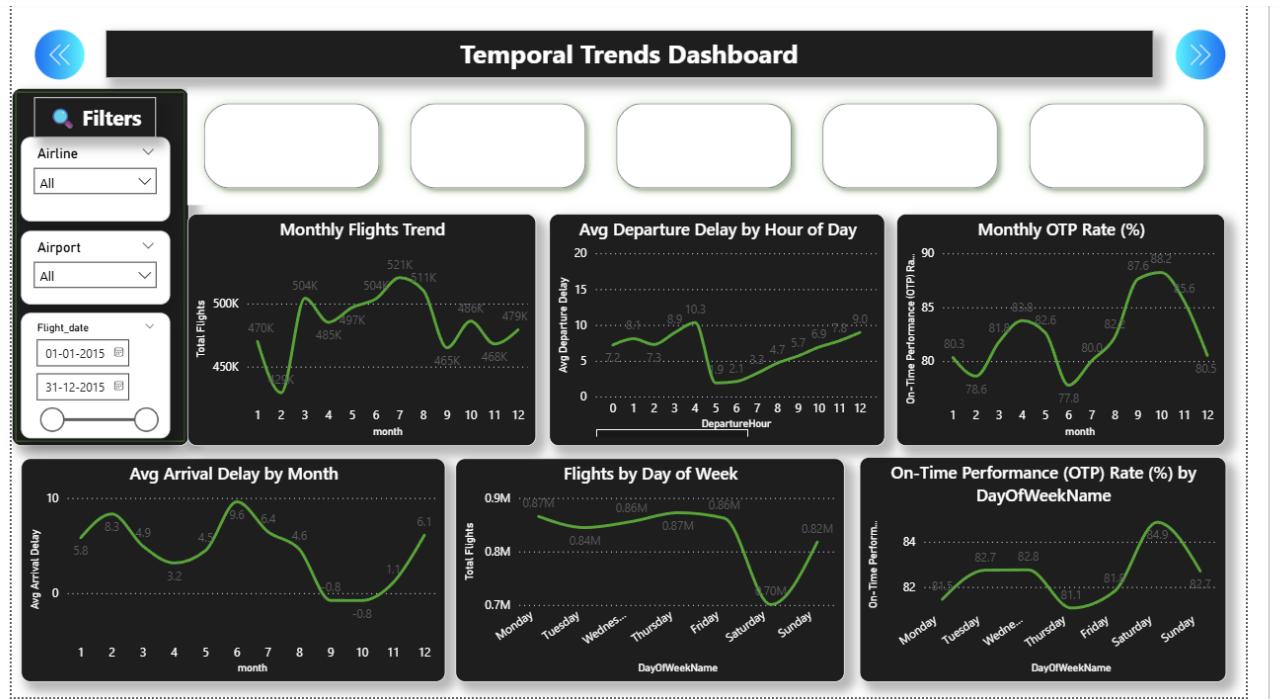


### Key Insights – Page 3: Airport Performance (Summary)

- Wilmington Airport = Highest Delay**  
Avg Departure: 29 mins, Avg Arrival: 24 mins — indicates persistent inefficiencies.
- Top 3 Busiest Airports Operate Smoothly**  
ATL (347K), ORD (286K), DFW (240K) — manage high volume with stable delays, showing strong capacity planning.
- Ithaca Airport = Highest Cancellation Rate**  
11.8% cancellations — ~8× above average; other weak performers: Mammoth Yosemite, Friedman Memorial.
- Regional Airports Shine in OTP**  
Canyonlands (98%), Valdez (97%) excel in punctuality — aided by low congestion and simpler logistics.

## Page 4: Temporal Trends

Trends in delays, flight volume, and cancellations by month, day, and hour

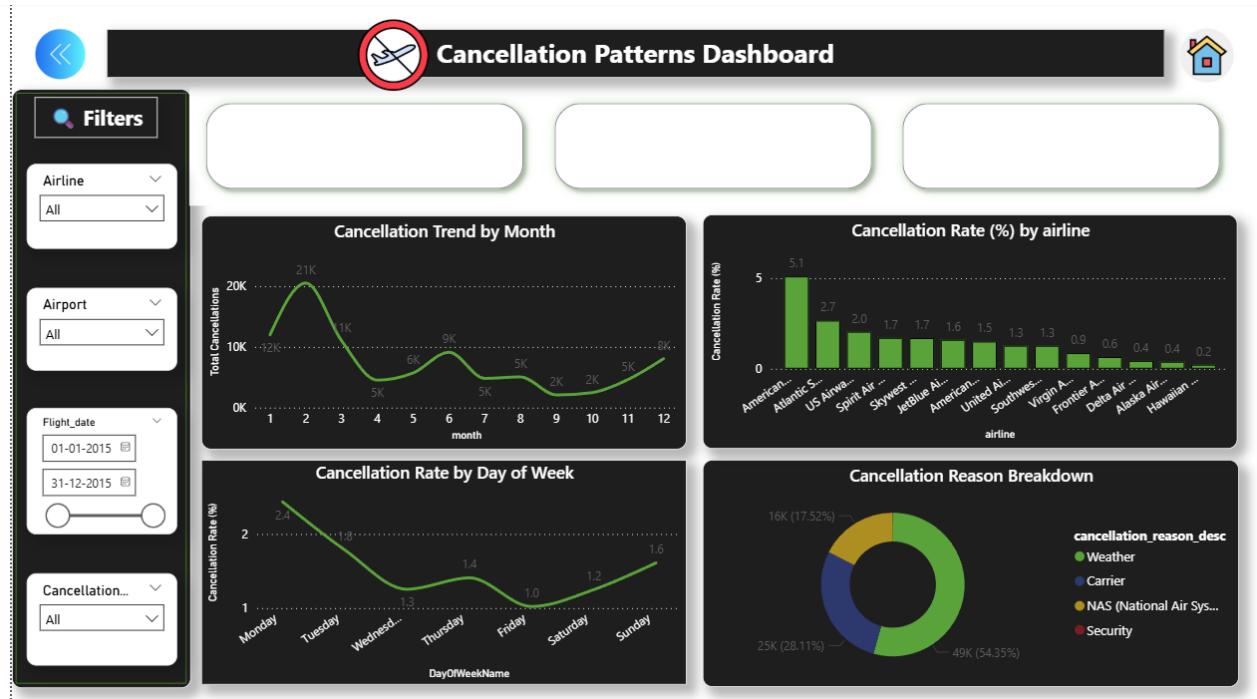


### Key Insights – Page 4: Temporal Trends (Summary)

- Best OTP in November (88.2%)**  
Worst in March (77.8%) — shows seasonal efficiency shifts tied to weather/travel surges.
- 5 AM Has Highest Departure Delays**  
Avg Delay: 10.3 mins — suggests early-morning prep and maintenance issues.
- April = Worst Arrival Delays**  
Avg: 9.6 mins — winter/spring weather may drive disruptions.
- Sunday = Least Busy but Best OTP**  
Lowest volume (0.70M), highest OTP (84.9%) — low congestion = better performance.
- Summer = High Volume, Poor Performance**  
June–August: peak flights but lower OTP and higher delays — likely system overload.

## Page 5: Cancellation Patterns

Cancellations by airline, reason, and time periods (month/day)



### ✖ Key Insights – Page 5: Cancellation Patterns (Summary)

- Overall Cancellation Rate = 1.54%**  
~90K out of 5.8M flights cancelled — moderately stable, but optimization is needed.
- March Has the Most Cancellations (~21K)**  
Peaks in Jan and Dec too — likely due to winter weather and holiday congestion.
- American Airlines = Highest Cancellation Rate (5.1%)**  
Nearly double that of second-highest (Atlantic Southeast) — suggests internal inefficiencies.
- Cancellations Spike on Mondays & Sundays**  
Mondays: 2.4%, Sundays: 1.6% — linked to travel peaks, crew shifts, or congestion.
- Weather = #1 Cause (54.35%)**  
Followed by Carrier (28.11%) and NAS (17.52%) — calls for better weather planning and carrier-side process improvements.

## 6. Key Insights

### From Page 1: Overview Dashboard

- Overall On-Time Performance: 82.41%
- Weather-related cancellations are highest (48,851 flights)
- Late Aircraft is the top delay contributor (~40% of total delay minutes)
- Southwest Airlines handled 1.26M flights

### From Page 2: Airline Performance

- Spirit Airlines has highest delays: 14.5 min (arrival), 16 min (departure)
- American Eagle has the highest cancellation rate at 5.1%
- Delta and JetBlue balance high volume with lower delays

### From Page 3: Airport Performance

- Wilmington Airport has the worst delays (29 min avg departure)
- Ithaca shows highest cancellation rate (11.8%)
- Canyonlands Field and Valdez Airport show 95%+ OTP

### From Page 4: Temporal Trends

- March has the highest cancellations (~21K flights)
- 5 AM has the worst average departure delays
- Sundays have best OTP and lowest volume

### From Page 5: Cancellation Patterns

- 54% of cancellations due to weather
- American Airlines has 5.1% cancellation rate
- Mondays show the most cancellations by day

---

## 7. Recommendations

1. **Reduce Late Aircraft Delays:** Improve gate turnaround and ground team efficiency
2. **AI-Driven Weather Planning:** Use predictive models to manage high-risk months (March, Dec, Jan)
3. **Redistribute Early Morning Flights:** Avoid congestion and delays at 5 AM
4. **Optimize American Airlines Operations:** Focus on hub-level audits and real-time staffing

5. **Support Regional Airports:** Improve weather-readiness and alternate routing at airports like Ithaca
- 

## 8. Conclusion

The dashboard reveals strong operational patterns and clear opportunities for improvement. By optimizing schedules, enhancing coordination, and investing in predictive planning, stakeholders can reduce delays and cancellations while improving the passenger experience. Future analysis may include 3-year trends and external factors (e.g., aircraft type, crew rotation).

---