```
-- ========================================
-- Phase 1 - Block 1: Row Count Validation
-- Project: U.S. Airline Performance & Delay Analysis
-- Author: Shruti Sumadhur Ghosh
-- ========================================

-- 🧠 This block confirms the number of rows in each table
-- ☑ Purpose: Ensure all 3 CSVs (airlines, airports, flights)
-- were successfully imported into the SQLite database

SELECT COUNT(*) AS airline_count FROM new_airlines;
SELECT COUNT(*) FROM new_airports;
SELECT COUNT(*) FROM new_flights;
--☑ Phase 1: Data Ingestion - Completed
--📌 Block 1: Row Count Validation
--Table Row Count
--new_airlines  14
--new_airports  322
--new_flights   1,048,575

--💡 All data has been successfully imported from CSVs into the SQLite database.
-- ========================================
-- Phase 1 - Block 2: Sample Data Preview
-- Project: U.S. Airline Performance & Delay Analysis
-- ========================================

-- 🧠 View a few records from each table to verify column structure & values

SELECT * FROM new_airlines LIMIT 5;
SELECT * FROM new_airports LIMIT 5;
SELECT * FROM new_flights LIMIT 5;
--☑ Phase 1 Completed: Summary
--Table Name    Status  Remarks
--new_airlines  ☑ Clean        2 columns, 14 rows
--new_airports  ☑ Clean        7 columns, 322 rows
--new_flights   ☑ Clean        31 columns, 1,048,575 rows

--✔ Key Validations Done:
--Column names are correct ☑

--Data types are appropriate (we'll confirm in PRAGMA check)

--No blank field names (like field1, field2) ☑

--All data appears to have loaded correctly ☑
-- ========================================
-- Phase 1 - Block 3: Table Schema Check
-- Purpose: Confirm correct datatypes are set
-- ========================================
```

```
-- 🧠 This confirms whether columns are using the correct types like INTEGER,
TEXT, REAL

PRAGMA table_info(new_airlines);
PRAGMA table_info(new_airports);
PRAGMA table_info(new_flights);
--☑ Schema Review Summary
--◇ new_airlines
--Column        Type    Notes
--IATA_CODE     TEXT    ☑ Primary Key
--AIRLINE       TEXT    ☑ Clean

--➡ Pass – Clean and minimal.

--◇ new_airports
--Column        Type    Notes
--IATA_CODE     TEXT    ☑ Primary Key
--LATITUDE/LONGITUDE    REAL    ☑ Geolocation Ready

--➡ Pass – All airport details available and well-typed.

--◇ new_flights
--Field Type    Notes
--Date/Time Fields      TEXT/INTEGER    We'll convert properly in Phase 2.
--Delay Fields  INTEGER ☑ Good for math/stats
--Distance      INTEGER ☑ Accurate
--Categorical   TEXT    ☑ Airline, Reason, Tail No., etc.
--Diverted/Cancelled    INTEGER ☑ Boolean-compatible

--➡ Pass – Just needs enrichment & cleanup.
-- ============================================
-- Phase 2 - Block 1: Initial Clean Flight View
-- Dataset: new_flights
-- Author: Shruti Sumadhur Ghosh
-- ============================================

-- 🧠 This view prepares for enrichment by selecting all relevant columns.

CREATE VIEW IF NOT EXISTS v_clean_flight_data AS
SELECT *
FROM new_flights
WHERE airline IS NOT NULL
  AND origin_airport IS NOT NULL
  AND destination_airport IS NOT NULL;
  -- ============================================
-- Phase 2 - Block 2: Add FLIGHT_DATE and FLIGHT_DATETIME
-- Project: U.S. Airline Performance & Delay Analysis
-- Dataset: v_clean_flight_data
-- Author: Shruti Sumadhur Ghosh
-- ============================================
```

```sql
-- 🧠 Purpose:
-- This block creates a view that adds two new columns:
-- 1. FLIGHT_DATE: Combines year, month, and day into 'YYYY-MM-DD'
-- 2. FLIGHT_DATETIME: Combines flight date and scheduled departure (HHMM) into
full datetime
--    Handles cases where scheduled_departure is < 1000 by padding with leading 0

-- ☑ Output View: v_flight_with_datetime

CREATE VIEW IF NOT EXISTS v_flight_with_datetime AS
SELECT
    *,

    -- Format YYYY-MM-DD
    printf('%04d-%02d-%02d', year, month, day) AS flight_date,

    -- Format full datetime: YYYY-MM-DD HH:MM
    printf(
        '%04d-%02d-%02d %02d:%02d',
        year, month, day,
        CAST(SUBSTR('0000' || scheduled_departure, -4, 2) AS INTEGER),
        CAST(SUBSTR('0000' || scheduled_departure, -2, 2) AS INTEGER)
    ) AS flight_datetime

FROM v_clean_flight_data;
-- ============================================
-- Phase 2 - Block 3: Add CANCELLATION_REASON_DESC
-- Project: U.S. Airline Performance & Delay Analysis
-- Dataset: v_flight_with_datetime
-- Author: Shruti Sumadhur Ghosh
-- ============================================

-- 🧠 Purpose:
-- Adds descriptive labels for cancellation reasons using CASE statement

-- ☑ Output View: v_flight_with_cancel_desc

CREATE VIEW IF NOT EXISTS v_flight_with_cancel_desc AS
SELECT
    *,

    -- Human-readable cancellation reason
    CASE cancellation_reason
        WHEN 'A' THEN 'Airline'
        WHEN 'B' THEN 'Weather'
        WHEN 'C' THEN 'National Air System'
        WHEN 'D' THEN 'Security'
        ELSE 'Not Cancelled'
    END AS cancellation_reason_desc
```

```
FROM v_flight_with_datetime;
-- ============================================
-- Phase 2 - Block 4: Final Enriched Analytical View
-- Project: U.S. Airline Performance & Delay Analysis
-- Author: Shruti Sumadhur Ghosh
-- ============================================

-- 🧠 Purpose:
-- Join flight data with airline names and origin/destination airport details

-- ☑ Output View: v_flight_data_enriched

CREATE VIEW IF NOT EXISTS v_flight_data_enriched AS
SELECT
    f.*,

    -- Airline Full Name
    al.airline AS airline_name,

    -- Origin Airport Details
    ao.airport AS origin_airport_name,
    ao.city AS origin_city,
    ao.state AS origin_state,
    ao.country AS origin_country,

    -- Destination Airport Details
    ad.airport AS dest_airport_name,
    ad.city AS dest_city,
    ad.state AS dest_state,
    ad.country AS dest_country

FROM v_flight_with_cancel_desc f
LEFT JOIN new_airlines al ON f.airline = al.iata_code
LEFT JOIN new_airports ao ON f.origin_airport = ao.iata_code
LEFT JOIN new_airports ad ON f.destination_airport = ad.iata_code;
-- ============================================
-- Phase 3 - Block 1: Overall Flight Statistics
-- Project: U.S. Airline Performance & Delay Analysis
-- Dataset: v_flight_data_enriched
-- Author: Shruti Sumadhur Ghosh
-- ============================================

-- 🧠 Explanation:
-- This block gives a high-level overview of:
-- 1. Total number of flights
-- 2. Number of cancelled flights
-- 3. Cancellation rate (%)
-- 4. (Optional) Diverted flight count and rate
```

```
SELECT
    COUNT(*) AS total_flights,
    SUM(CASE WHEN cancelled = 1 THEN 1 ELSE 0 END) AS cancelled_flights,
    ROUND(
        100.0 * SUM(CASE WHEN cancelled = 1 THEN 1 ELSE 0 END) / COUNT(*),
        2
    ) AS cancellation_rate_pct

-- Optional: Uncomment to include diverted stats
-- , SUM(CASE WHEN diverted = 1 THEN 1 ELSE 0 END) AS diverted_flights
-- , ROUND(100.0 * SUM(CASE WHEN diverted = 1 THEN 1 ELSE 0 END) / COUNT(*), 2) AS
diversion_rate_pct

FROM v_flight_data_enriched;
--📊 Query Output:
--total_flights cancelled_flights       cancellation_rate_pct
--1,048,575     40,527  3.86%

--☑ Interpretation:
--Out of ~1.05 million domestic flights in 2015, 3.86% were cancelled, representing
a relatively low cancellation rate across the U.S. airline industry for that year.
-- ============================================
-- Phase 3 - Block 2: Delay Statistics Summary
-- Project: U.S. Airline Performance & Delay Analysis
-- Dataset: v_flight_data_enriched
-- Author: Shruti Sumadhur Ghosh
-- ============================================

-- 🧠 Explanation:
-- This block provides:
-- 1. Average, minimum, and maximum departure delay
-- 2. Average, minimum, and maximum arrival delay
-- ☑ Excludes cancelled flights to avoid skewed stats

SELECT
    ROUND(AVG(departure_delay), 2) AS avg_departure_delay,
    MIN(departure_delay) AS min_departure_delay,
    MAX(departure_delay) AS max_departure_delay,

    ROUND(AVG(arrival_delay), 2) AS avg_arrival_delay,
    MIN(arrival_delay) AS min_arrival_delay,
    MAX(arrival_delay) AS max_arrival_delay

FROM v_flight_data_enriched
WHERE cancelled = 0;
-- ============================================
-- ☑ Phase 3 - Block 2: Delay Statistics Summary
-- Project: U.S. Airline Performance & Delay Analysis
-- Dataset: v_flight_data_enriched
-- Author: Shruti Sumadhur Ghosh
```

```
-- ========================================

-- 🧠 Objective:
-- Analyze the central tendency and range of delays.
-- Specifically:
--  ◇ Average, Minimum, and Maximum Departure Delay
--  ◇ Average, Minimum, and Maximum Arrival Delay
--  ◇ Note: Cancelled flights excluded to avoid skewed delay stats.

-- 📊 Query Output:
-- avg_departure_delay | min_departure_delay | max_departure_delay |
avg_arrival_delay | min_arrival_delay | max_arrival_delay
--
--------------------|---------------------|---------------------|----------------
---|--------------------|-------------------
--        11.28        |         -61          |        1988          |       7.61
     |        -82         |        1971

-- ☑ Interpretation:
-- On average, flights experienced an **11.28-minute departure delay** and a
**7.61-minute arrival delay**.
-- However, the data also includes **extreme delays**, with some flights delayed by
over 30 hours.
-- Negative delay values suggest flights that departed or arrived ahead of
schedule.

-- ▧ Next: Block 3 – Delay Type Contribution Breakdown
-- ========================================
-- ☑ Phase 3 - Block 3: Delay Type Contribution Breakdown
-- Project: U.S. Airline Performance & Delay Analysis
-- Dataset: v_flight_data_enriched
-- Author: Shruti Sumadhur Ghosh
-- ========================================

-- 🧠 Objective:
-- Analyze what portion of total delay is caused by each category:
-- Airline, Weather, Air System (NAS), Security, Late Aircraft

SELECT
    -- Total delay in minutes by category
    SUM(airline_delay) AS airline_delay_total,
    ROUND(100.0 * SUM(airline_delay) / total_delay, 2) AS airline_pct,

    SUM(weather_delay) AS weather_delay_total,
    ROUND(100.0 * SUM(weather_delay) / total_delay, 2) AS weather_pct,

    SUM(air_system_delay) AS nas_delay_total,
    ROUND(100.0 * SUM(air_system_delay) / total_delay, 2) AS nas_pct,

    SUM(security_delay) AS security_delay_total,
```

```
    ROUND(100.0 * SUM(security_delay) / total_delay, 2) AS security_pct,

    SUM(late_aircraft_delay) AS late_aircraft_delay_total,
    ROUND(100.0 * SUM(late_aircraft_delay) / total_delay, 2) AS late_aircraft_pct

FROM v_flight_data_enriched,

-- Subquery to get total delay from all types (cancelled excluded)
(
    SELECT
        SUM(
            airline_delay + weather_delay + air_system_delay +
            security_delay + late_aircraft_delay
        ) AS total_delay
    FROM v_flight_data_enriched
    WHERE cancelled = 0
) AS delay_totals

WHERE cancelled = 0;
-- 📊 Delay Type Contribution Breakdown:
--
+----------------------+-------------+----------------------+--------------+------
------------+-----------+--------------------------+------------------+------------
----------------+-------------------+
-- | airline_delay_total | airline_pct | weather_delay_total  | weather_pct |
nas_delay_total | nas_pct  | security_delay_total    | security_pct     |
late_aircraft_delay_total | late_aircraft_pct |
--
+----------------------+-------------+----------------------+--------------+------
------------+-----------+--------------------------+------------------+------------
----------------+-------------------+
-- |      4,160,027       |    31.16    |       810,195         |    6.07     |
3,129,132     |  23.44   |         13,101           |     0.10         |
5,238,195          |      39.24        |
--
+----------------------+-------------+----------------------+--------------+------
------------+-----------+--------------------------+------------------+------------
----------------+-------------------+
-- ☑ Interpretation:
-- In 2015, the most significant contributor to total delay time was:
--  ◇ Late-arriving aircraft (39.24%) – cascading delays from previous flights.
--  ◇ Airline-related delays (31.16%) – internal operational issues.
--  ◇ Airspace/NAS-related delays (23.44%) – air traffic congestion and system
constraints.
--  ◇ Weather and Security delays were minimal contributors at 6.07% and 0.10%,
respectively.
-- 🛩 Together, these figures help identify key pressure points in flight
punctuality.
-- ============================================
-- ☑ Phase 3 - Block 4: KPI Definitions Summary
```

```
-- Project: U.S. Airline Performance & Delay Analysis
-- Dataset: v_flight_data_enriched
-- Author: Shruti Sumadhur Ghosh
-- ========================================

-- 🧠 Purpose:
-- Compute key performance indicators (KPIs) for airline operations:
--  ◇ On-Time Performance Rate (OTP)
--  ◇ Average Arrival & Departure Delay (excluding cancelled)
--  ◇ Cancellation Rate (overall)

SELECT
    -- Total number of flights
    COUNT(*) AS total_flights,

    -- On-time flights: arrival delay ≤ 15 mins
    SUM(CASE WHEN cancelled = 0 AND arrival_delay <= 15 THEN 1 ELSE 0 END) AS
on_time_flights,

    -- OTP Rate = On-time flights / Total non-cancelled
    ROUND(
        100.0 * SUM(CASE WHEN cancelled = 0 AND arrival_delay <= 15 THEN 1 ELSE 0
END) /
        SUM(CASE WHEN cancelled = 0 THEN 1 ELSE 0 END), 2
    ) AS otp_rate_pct,

    -- Avg departure and arrival delays
    ROUND(AVG(CASE WHEN cancelled = 0 THEN departure_delay END), 2) AS
avg_departure_delay,
    ROUND(AVG(CASE WHEN cancelled = 0 THEN arrival_delay END), 2) AS
avg_arrival_delay,

    -- Cancellation rate
    SUM(CASE WHEN cancelled = 1 THEN 1 ELSE 0 END) AS cancelled_flights,
    ROUND(
        100.0 * SUM(CASE WHEN cancelled = 1 THEN 1 ELSE 0 END) / COUNT(*),
        2
    ) AS cancellation_rate_pct

FROM v_flight_data_enriched;
-- 📊 KPI Definitions Summary:
--
+---------------+------------------+----------------+---------------------+--------
-----------+--------------------+-------------------------+
-- | total_flights | on_time_flights  | otp_rate_pct   | avg_departure_delay |
avg_arrival_delay | cancelled_flights  | cancellation_rate_pct  |
--
+---------------+------------------+----------------+---------------------+--------
-----------+--------------------+-------------------------+
-- |   1,048,575   |     785,087      |     77.88%     |        11.28        |
```

```
 7.61        |       40,527       |          3.86%          |
--
+--------------+-----------------+---------------+--------------------+--------
-----------+--------------------+------------------------+

-- ☑ Interpretation:
--  ◇ **77.88%** of all non-cancelled flights arrived on time (within 15 minutes
of schedule), indicating a moderate industry-level On-Time Performance (OTP).
--  ◇ The **average departure delay** was **11.28 minutes**, while the **average
arrival delay** was **7.61 minutes**, suggesting partial recovery of delays during
flight.
--  ◇ **3.86%** of flights were cancelled — a relatively low cancellation rate.
-- 📌 These KPIs form a strong baseline to compare airline and airport performance
in further analysis.
PRAGMA table_info(v_flight_data_enriched);
-- ============================================
-- ☑ Phase 3 - Block 5: Airline-Level Performance Summary
-- Project: U.S. Airline Performance & Delay Analysis
-- Dataset: v_flight_data_enriched
-- Author: Shruti Sumadhur Ghosh
-- ============================================

-- 🧠 Objective:
-- Analyze airline-wise performance across key metrics:
--  ◇ Total Flights
--  ◇ On-Time Performance (%)
--  ◇ Average Delays (Departure & Arrival)
--  ◇ Cancellation Rate (%)

SELECT
    airline_name,

    COUNT(*) AS total_flights,

    -- On-time flights: arrival delay <= 15 mins
    SUM(CASE WHEN cancelled = 0 AND arrival_delay <= 15 THEN 1 ELSE 0 END) AS
on_time_flights,
    ROUND(
        100.0 * SUM(CASE WHEN cancelled = 0 AND arrival_delay <= 15 THEN 1 ELSE 0
END) /
        NULLIF(SUM(CASE WHEN cancelled = 0 THEN 1 ELSE 0 END), 0),
        2
    ) AS otp_rate_pct,

    -- Average delays (exclude cancelled)
    ROUND(AVG(CASE WHEN cancelled = 0 THEN departure_delay END), 2) AS
avg_departure_delay,
    ROUND(AVG(CASE WHEN cancelled = 0 THEN arrival_delay END), 2) AS
avg_arrival_delay,
```

```
    -- Cancellation metrics
    SUM(CASE WHEN cancelled = 1 THEN 1 ELSE 0 END) AS cancelled_flights,
    ROUND(
        100.0 * SUM(CASE WHEN cancelled = 1 THEN 1 ELSE 0 END) / COUNT(*),
        2
    ) AS cancellation_rate_pct

FROM v_flight_data_enriched
GROUP BY airline_name
ORDER BY total_flights DESC;
-- ==========================================
-- ☑ Phase 3 - Block 5: Airline-Level Performance Summary
-- Project: U.S. Airline Performance & Delay Analysis
-- Dataset: v_flight_data_enriched
-- Author: Shruti Sumadhur Ghosh
-- ==========================================

-- 🤔 Objective:
-- Analyze airline-wise performance across key metrics:
--  ◇ Total Flights
--  ◇ On-Time Performance (%)
--  ◇ Average Delays (Departure & Arrival)
--  ◇ Cancellation Rate (%)

-- 📊 Airline Performance Summary Table:

--
+------------------------------+--------------+-----------------+--------------
-+--------------------+------------------+--------------------+-----------------
-------+
-- |         airline_name         | total_flights | on_time_flights  | otp_rate_pct
  | avg_departure_delay | avg_arrival_delay | cancelled_flights  |
cancellation_rate_pct  |
--
+------------------------------+--------------+-----------------+--------------
-+--------------------+------------------+--------------------+-----------------
-------+
-- | Southwest Airlines Co.       |     221,586    |     173,933      |     80.91%
   |        10.05        |       3.84        |       6,606        |          2.98%
       |
-- | Delta Air Lines Inc.         |     147,486    |     120,721      |     83.47%
   |        9.62         |       2.76        |       2,861        |          1.94%
       |
-- | Atlantic Southeast Airlines  |     111,206    |      79,331      |     75.60%
   |        11.27        |       10.34       |       6,274        |          5.64%
       |
-- | Skywest Airlines Inc.        |     107,099    |      80,166      |     77.00%
   |        11.22        |       9.85        |       2,983        |          2.79%
       |
-- | American Airlines Inc.       |      97,549    |      72,193      |     77.74%
```

```
   |       11.27        |       8.23        |       4,685        |          4.80%
       |
-- | United Air Lines Inc.        |      87,606    |      65,458      |     76.83%
   |       15.03        |       7.44        |       2,403        |          2.74%
       |
-- | US Airways Inc.              |      73,942    |      56,254      |     79.46%
   |       7.53         |       5.52        |       3,143        |          4.25%
       |
-- | American Eagle Airlines Inc. |      65,513    |      36,898      |     63.85%
   |       17.96        |       20.41       |       7,727        |         11.79%
       |
-- | JetBlue Airways              |      48,157    |      31,940      |     70.18%
   |       15.87        |       13.94       |       2,645        |          5.49%
       |
-- | Alaska Airlines Inc.         |      29,614    |      25,356      |     86.20%
   |       2.90         |       -0.61       |        197         |          0.67%
       |
-- | Spirit Air Lines             |      19,612    |      13,383      |     69.88%
   |       16.03        |       14.96       |        461         |          2.35%
       |
-- | Frontier Airlines Inc.       |      14,669    |       9,123      |     63.20%
   |       23.09        |       24.33       |        233         |          1.59%
       |
-- | Hawaiian Airlines Inc.       |      14,133    |      12,119      |     85.95%
   |       1.55         |       4.39        |         33         |          0.23%
       |
-- | Virgin America               |      10,403    |       8,212      |     81.09%
   |       10.24        |       5.24        |        276         |          2.65%
       |
--
+------------------------------+---------------+------------------+---------------
-+---------------------+-------------------+--------------------+-----------------
-------+

-- ☑ Interpretation:
--  ◇ **Alaska Airlines** (86.2%) and **Hawaiian Airlines** (85.95%) lead in
On-Time Performance, with minimal delays and cancellations.
--  ◇ **Frontier** and **American Eagle** report the **worst delays and lowest
OTP**, with departure delays averaging over 17–23 minutes.
--  ◇ **Southwest** and **Delta** operate the largest number of flights and
maintain strong OTPs (~81–83%) with low cancellation rates (<3%).
--  ◇ **American Eagle** has the **highest cancellation rate** at **11.79%**,
highlighting operational inefficiencies.
-- 🛩 These airline-level insights will help benchmark future KPIs and inform
route/partner optimization.

-- ▨ Next Step: Airport-level performance analysis (Block 6)
-- ========================================
-- ☑ Phase 3 - Block 6: Airport-Level Performance Summary
-- Project: U.S. Airline Performance & Delay Analysis
```

```
-- Dataset: v_flight_data_enriched
-- Author: Shruti Sumadhur Ghosh
-- ==========================================

-- 🧐 Objective:
-- Analyze origin airport performance on:
--  ◇ Total Departures
--  ◇ On-Time Performance (%)
--  ◇ Average Departure Delay
--  ◇ Cancellation Rate

SELECT
    origin_airport,
    origin_airport_name,
    origin_city,
    origin_state,

    COUNT(*) AS total_departures,

    -- On-time departures: departure delay ≤ 15 mins
    SUM(CASE WHEN cancelled = 0 AND departure_delay <= 15 THEN 1 ELSE 0 END) AS
on_time_departures,
    ROUND(
        100.0 * SUM(CASE WHEN cancelled = 0 AND departure_delay <= 15 THEN 1 ELSE 0
END) /
        NULLIF(SUM(CASE WHEN cancelled = 0 THEN 1 ELSE 0 END), 0),
        2
    ) AS otp_departure_pct,

    -- Average departure delay
    ROUND(AVG(CASE WHEN cancelled = 0 THEN departure_delay END), 2) AS
avg_departure_delay,

    -- Cancellation metrics
    SUM(CASE WHEN cancelled = 1 THEN 1 ELSE 0 END) AS cancelled_departures,
    ROUND(
        100.0 * SUM(CASE WHEN cancelled = 1 THEN 1 ELSE 0 END) / COUNT(*),
        2
    ) AS cancellation_rate_pct

FROM v_flight_data_enriched
GROUP BY origin_airport
ORDER BY total_departures DESC
LIMIT 20;
-- 📊 Top 20 Origin Airports by Departure Volume
--
+------------+-----------------------------------------------+----------------------
---+--------------+--------------------+-----------------------+---------------
----+-------------------------+
-- | ORIGIN     | AIRPORT NAME                                  | CITY
```

```
    | STATE         | TOTAL DEPARTURES    | ON-TIME DEPARTURES    | OTP %
  | AVG DEPARTURE DELAY     |
--
+------------+-----------------------------------------------+----------------------
---+--------------+---------------------+-----------------------+---------------
----+-------------------------+
-- | ATL        | Hartsfield-Jackson Atlanta Intl             | Atlanta
    | GA           | 66,599              | 53,217                | 81.64%
    | 9.27                    |
-- | ORD        | Chicago O'Hare Intl                         | Chicago
    | IL           | 52,961              | 33,345                | 67.07%
    | 20.09                   |
-- | DFW        | Dallas/Fort Worth Intl                      | Dallas-Fort Worth
    | TX           | 50,933              | 35,491                | 74.95%
    | 13.17                   |
-- | LAX        | Los Angeles Intl                            | Los Angeles
    | CA           | 38,473              | 30,655                | 81.12%
    | 9.34                    |
-- | DEN        | Denver Intl                                 | Denver
    | CO           | 38,254              | 26,761                | 71.43%
    | 15.26                   |
-- | IAH        | George Bush Intercontinental                | Houston
    | TX           | 29,802              | 23,595                | 80.64%
    | 9.33                    |
-- | PHX        | Phoenix Sky Harbor Intl                     | Phoenix
    | AZ           | 29,262              | 23,572                | 81.38%
    | 9.13                    |
-- | SFO        | San Francisco Intl                          | San Francisco
    | CA           | 28,428              | 22,032                | 79.55%
    | 11.70                   |
-- | LAS        | McCarran Intl                               | Las Vegas
    | NV           | 25,806              | 20,164                | 79.13%
    | 10.91                   |
-- | MCO        | Orlando Intl                                | Orlando
    | FL           | 22,575              | 17,418                | 79.24%
    | 11.67                   |
-- | LGA        | LaGuardia (Marine Air Terminal)             | New York
    | NY           | 21,505              | 13,267                | 69.93%
    | 18.64                   |
-- | DTW        | Detroit Metro                               | Detroit
    | MI           | 21,328              | 15,853                | 76.83%
    | 14.44                   |
-- | CLT        | Charlotte Douglas Intl                      | Charlotte
    | NC           | 20,434              | 16,380                | 82.93%
    | 8.20                    |
-- | BOS        | Logan Intl                                  | Boston
    | MA           | 20,193              | 13,794                | 75.66%
    | 14.42                   |
-- | MSP        | Minneapolis-Saint Paul Intl                 | Minneapolis
    | MN           | 20,073              | 16,190                | 82.00%
```

```
    | 10.23                    |
-- | EWR        | Newark Liberty Intl                          | Newark
    | NJ           | 19,608              | 13,211                | 73.53%
    | 15.06                    |
-- | SLC        | Salt Lake City Intl                          | Salt Lake City
    | UT           | 19,325              | 16,601                | 86.50%
    | 5.46                     |
-- | JFK        | JFK Intl (New York Intl)                     | New York
    | NY           | 18,873              | 12,862                | 72.71%
    | 18.84                    |
-- | SEA        | Seattle-Tacoma Intl                          | Seattle
    | WA           | 18,839              | 15,861                | 84.80%
    | 6.93                     |
-- | FLL        | Fort Lauderdale-Hollywood Intl               | Ft. Lauderdale
    | FL           | 16,187              | 12,270                | 77.92%
    | 11.66                    |
--
+------------+-----------------------------------------------+----------------------
---+--------------+---------------------+-----------------------+----------------
----+-------------------------+
-- ☑ Interpretation:
--  ◇ ATL (Atlanta) had the highest total departures (66,599), with a strong OTP
of 81.64% and moderate delays.
--  ◇ ORD (Chicago O’Hare) and DFW (Dallas/Fort Worth) had lower OTPs (67.07% and
74.95%) and high average delays (20.09 mins at ORD).
--  ◇ SLC (Salt Lake City) and SEA (Seattle) demonstrated excellent operational
performance, with OTPs above 84% and very low delay averages (5.46 mins and 6.93
mins).
--  ◇ LGA (LaGuardia) and JFK struggled with longer delays (~18+ mins) and OTPs
under 73%, indicating likely airspace congestion challenges.
--  ◇ Cancellation rates were highest at LGA (11.78%) and EWR (8.36%), possibly
due to poor weather or operational bottlenecks in the NY/NJ corridor.
-- ✈ These metrics highlight the best- and worst-performing airports and help
identify where delay-reduction strategies could be prioritized.
-- ========================================
SELECT
    origin_airport || ' → ' || destination_airport AS route,
    airline_name,
    COUNT(*) AS total_flights,

    -- On-time arrivals: arrival delay ≤ 15 mins
    SUM(CASE WHEN cancelled = 0 AND arrival_delay <= 15 THEN 1 ELSE 0 END) AS
on_time_arrivals,
    ROUND(
        100.0 * SUM(CASE WHEN cancelled = 0 AND arrival_delay <= 15 THEN 1 ELSE 0
END) /
        NULLIF(SUM(CASE WHEN cancelled = 0 THEN 1 ELSE 0 END), 0), 2
    ) AS otp_rate_pct,

    -- Avg arrival delay (only for non-cancelled)
```

```
    ROUND(AVG(CASE WHEN cancelled = 0 THEN arrival_delay END), 2) AS
avg_arrival_delay,

    -- Cancellation stats
    SUM(CASE WHEN cancelled = 1 THEN 1 ELSE 0 END) AS cancelled_flights,
    ROUND(
        100.0 * SUM(CASE WHEN cancelled = 1 THEN 1 ELSE 0 END) / COUNT(*), 2
    ) AS cancellation_rate_pct

FROM v_flight_data_enriched
GROUP BY route, airline_name
ORDER BY total_flights DESC
LIMIT 20;
-- ==========================================
-- ☑ Phase 3 - Block 7: Route-Level Performance Summary
-- Project: U.S. Airline Performance & Delay Analysis
-- Dataset: v_flight_data_enriched
-- Author: Shruti Sumadhur Ghosh
-- ==========================================

-- 📊 Route-Level Performance Summary:
--
+------------------+----------------------------+--------------+---------------
---+----------------+--------------------+--------------------+----------------
--------+
-- | Route            | Airline Name               | Total Flights | On-Time
Arrivals  | OTP Rate (%)   | Avg Arrival Delay   | Cancelled Flights   |
Cancellation Rate (%)    |
--
+------------------+----------------------------+--------------+---------------
---+----------------+--------------------+--------------------+----------------
--------+
-- | HNL → OGG        | Hawaiian Airlines Inc.     | 1801         | 1608
     | 89.58          | 2.87               | 6                  | 0.33
          |
-- | OGG → HNL        | Hawaiian Airlines Inc.     | 1798         | 1501
     | 83.95          | 6.12               | 10                 | 0.56
          |
-- | DAL → HOU        | Southwest Airlines Co.     | 1340         | 1051
     | 80.97          | 4.53               | 42                 | 3.13
          |
-- | HOU → DAL        | Southwest Airlines Co.     | 1330         | 1067
     | 83.23          | 4.53               | 48                 | 3.61
          |
-- | HNL → KOA        | Hawaiian Airlines Inc.     | 1279         | 1157
     | 90.53          | 2.62               | 1                  | 0.08
          |
-- | KOA → HNL        | Hawaiian Airlines Inc.     | 1276         | 1091
     | 85.70          | 5.02               | 3                  | 0.24
          |
```

```
-- | SAN → LAX         | Skywest Airlines Inc.       | 1213          | 1009
      | 84.44          | 3.93                | 18                | 1.48
            |
-- | LAX → SAN         | Skywest Airlines Inc.       | 1206          | 969
      | 81.29          | 6.03                | 14                | 1.16
            |
-- | LAX → DFW         | American Airlines Inc.      | 1164          | 917
      | 81.58          | 7.35                | 40                | 3.44
            |
-- | DFW → LAX         | American Airlines Inc.      | 1158          | 856
      | 76.70          | 7.88                | 42                | 3.63
            |
-- | HNL → LIH         | Hawaiian Airlines Inc.      | 1155          | 1061
      | 91.94          | 1.73                | 1                 | 0.09
            |
-- | LIH → HNL         | Hawaiian Airlines Inc.      | 1150          | 1000
      | 87.03          | 3.59                | 1                 | 0.09
            |
-- | MCO → ATL         | Delta Air Lines Inc.        | 1062          | 974
      | 92.59          | -2.77               | 10                | 0.94
            |
-- | ATL → MCO         | Delta Air Lines Inc.        | 1058          | 900
      | 85.88          | 2.21                | 10                | 0.95
            |
-- | LGA → ATL         | Delta Air Lines Inc.        | 1049          | 688
      | 70.56          | 13.03               | 74                | 7.05
            |
-- | ATL → LGA         | Delta Air Lines Inc.        | 1047          | 695
      | 70.85          | 15.16               | 66                | 6.30
            |
-- | DFW → ORD         | American Airlines Inc.      | 1003          | 694
      | 73.13          | 13.37               | 54                | 5.38
            |
-- | ORD → DFW         | American Airlines Inc.      | 1001          | 688
      | 72.57          | 13.31               | 53                | 5.29
            |
-- | SAT → DFW         | American Airlines Inc.      | 990           | 787
      | 84.71          | 3.99                | 61                | 6.16
            |
-- | DFW → SAT         | American Airlines Inc.      | 988           | 698
      | 74.81          | 11.14               | 55                | 5.57
            |
--
+------------------+---------------------------+--------------+---------------
---+---------------+--------------------+--------------------+-----------------
--------+

-- ☑ Interpretation: Route-Level Performance Summary
-- ◇ Hawaiian Airlines dominates the list, especially for intra-Hawaii routes
(e.g., HNL → OGG, HNL → LIH) with exceptional OTP (≥90%) and extremely low
```

cancellation rates (<0.1%).
-- ◇ Delta’s MCO → ATL route shows a negative average arrival delay (-2.77 mins), indicating early arrivals — a sign of schedule efficiency.
-- ◇ Southwest's DAL–HOU shuttle routes maintain high frequency with solid OTP (~81–83%), though cancellation rates are slightly higher (~3%).
-- ◇ American Airlines’ DFW → LAX and DFW → ORD routes show lower OTP (~73–76%) and significant delays (avg 7–13 mins), possibly due to congestion or operational strain.
-- △ Routes involving LGA and ORD show high cancellation rates (5–7%), likely impacted by weather or airspace congestion.
-- 📌 This summary highlights top-performing routes and identifies segments where operational improvements could be targeted.
-- ▨ Next Step: Destination Airport-Level Analysis (Arrivals)
-- ==
-- ☑ Phase 3 - Block 7: Route-Level Performance Summary
-- Project: U.S. Airline Performance & Delay Analysis
-- Dataset: v_flight_data_enriched
-- Author: Shruti Sumadhur Ghosh
-- ==

-- 📊 Route-Level Performance Summary:
--
+------------------+----------------------------+--------------+---------------+----------------+--------------------+---------------------+-----------------+
-- | Route | Airline Name | Total Flights | On-Time Arrivals | OTP Rate (%) | Avg Arrival Delay | Cancelled Flights | Cancellation Rate (%) |
--
+------------------+----------------------------+--------------+---------------+----------------+--------------------+---------------------+-----------------+
-- | HNL → OGG | Hawaiian Airlines Inc. | 1801 | 1608 | 89.58 | 2.87 | 6 | 0.33 |
-- | OGG → HNL | Hawaiian Airlines Inc. | 1798 | 1501 | 83.95 | 6.12 | 10 | 0.56 |
-- | DAL → HOU | Southwest Airlines Co. | 1340 | 1051 | 80.97 | 4.53 | 42 | 3.13 |
-- | HOU → DAL | Southwest Airlines Co. | 1330 | 1067 | 83.23 | 4.53 | 48 | 3.61 |
-- | HNL → KOA | Hawaiian Airlines Inc. | 1279 | 1157 | 90.53 | 2.62 | 1 | 0.08 |
-- | KOA → HNL | Hawaiian Airlines Inc. | 1276 | 1091 | 85.70 | 5.02 | 3 | 0.24 |

```
-- | SAN → LAX         | Skywest Airlines Inc.       | 1213          | 1009
      | 84.44          | 3.93                | 18                | 1.48
            |
-- | LAX → SAN         | Skywest Airlines Inc.       | 1206          | 969
      | 81.29          | 6.03                | 14                | 1.16
            |
-- | LAX → DFW         | American Airlines Inc.      | 1164          | 917
      | 81.58          | 7.35                | 40                | 3.44
            |
-- | DFW → LAX         | American Airlines Inc.      | 1158          | 856
      | 76.70          | 7.88                | 42                | 3.63
            |
-- | HNL → LIH         | Hawaiian Airlines Inc.      | 1155          | 1061
      | 91.94          | 1.73                | 1                 | 0.09
            |
-- | LIH → HNL         | Hawaiian Airlines Inc.      | 1150          | 1000
      | 87.03          | 3.59                | 1                 | 0.09
            |
-- | MCO → ATL         | Delta Air Lines Inc.        | 1062          | 974
      | 92.59          | -2.77               | 10                | 0.94
            |
-- | ATL → MCO         | Delta Air Lines Inc.        | 1058          | 900
      | 85.88          | 2.21                | 10                | 0.95
            |
-- | LGA → ATL         | Delta Air Lines Inc.        | 1049          | 688
      | 70.56          | 13.03               | 74                | 7.05
            |
-- | ATL → LGA         | Delta Air Lines Inc.        | 1047          | 695
      | 70.85          | 15.16               | 66                | 6.30
            |
-- | DFW → ORD         | American Airlines Inc.      | 1003          | 694
      | 73.13          | 13.37               | 54                | 5.38
            |
-- | ORD → DFW         | American Airlines Inc.      | 1001          | 688
      | 72.57          | 13.31               | 53                | 5.29
            |
-- | SAT → DFW         | American Airlines Inc.      | 990           | 787
      | 84.71          | 3.99                | 61                | 6.16
            |
-- | DFW → SAT         | American Airlines Inc.      | 988           | 698
      | 74.81          | 11.14               | 55                | 5.57
            |
--
+------------------+---------------------------+--------------+---------------
---+---------------+--------------------+--------------------+-----------------
--------+

-- ☑ Interpretation: Route-Level Performance Summary
-- ◇ Hawaiian Airlines dominates the list, especially for intra-Hawaii routes
(e.g., HNL → OGG, HNL → LIH) with exceptional OTP (≥90%) and extremely low
```

```
cancellation rates (<0.1%).
-- ◇ Delta’s MCO → ATL route shows a negative average arrival delay (-2.77 mins), indicating early arrivals — a sign of schedule efficiency.
-- ◇ Southwest's DAL–HOU shuttle routes maintain high frequency with solid OTP (~81–83%), though cancellation rates are slightly higher (~3%).
-- ◇ American Airlines’ DFW → LAX and DFW → ORD routes show lower OTP (~73–76%) and significant delays (avg 7–13 mins), possibly due to congestion or operational strain.
-- ⚠ Routes involving LGA and ORD show high cancellation rates (5–7%), likely impacted by weather or airspace congestion.
-- 📌 This summary highlights top-performing routes and identifies segments where operational improvements could be targeted.
-- ▨ Next Step: Destination Airport-Level Analysis (Arrivals)
PRAGMA table_info(new_airports);
PRAGMA table_info(v_flight_data_enriched);
-- ============================================
-- ☑ Phase 3 - Block 8: Destination Airport-Level Arrival Performance
-- Project: U.S. Airline Performance & Delay Analysis
-- Dataset: v_flight_data_enriched + new_airports
-- Author: Shruti Sumadhur Ghosh
-- ============================================

-- 🎯 Objective:
-- Evaluate airport performance as *destinations* based on:
-- ◇ Total Arrivals
-- ◇ On-Time Performance (OTP %)
-- ◇ Average Arrival Delay
-- ◇ Cancellation Rate

SELECT
    v.DESTINATION_AIRPORT AS DEST,
    a.AIRPORT AS DEST_AIRPORT_NAME,
    a.CITY AS DEST_CITY,
    a.STATE AS DEST_STATE,

    COUNT(*) AS total_arrivals,

    -- On-time arrivals: arrival delay ≤ 15 mins
    SUM(CASE WHEN v.CANCELLED = 0 AND v.ARRIVAL_DELAY <= 15 THEN 1 ELSE 0 END) AS on_time_arrivals,
    ROUND(
        100.0 * SUM(CASE WHEN v.CANCELLED = 0 AND v.ARRIVAL_DELAY <= 15 THEN 1 ELSE 0 END) /
        NULLIF(SUM(CASE WHEN v.CANCELLED = 0 THEN 1 ELSE 0 END), 0), 2
    ) AS otp_arrival_pct,

    -- Average arrival delay (excluding cancelled)
    ROUND(AVG(CASE WHEN v.CANCELLED = 0 THEN v.ARRIVAL_DELAY END), 2) AS avg_arrival_delay,
```

```
    -- Cancellation metrics
    SUM(CASE WHEN v.CANCELLED = 1 THEN 1 ELSE 0 END) AS cancelled_arrivals,
    ROUND(
        100.0 * SUM(CASE WHEN v.CANCELLED = 1 THEN 1 ELSE 0 END) / COUNT(*),
        2
    ) AS cancellation_rate_pct

FROM v_flight_data_enriched v
JOIN new_airports a ON v.DESTINATION_AIRPORT = a.IATA_CODE

GROUP BY v.DESTINATION_AIRPORT
ORDER BY total_arrivals DESC
LIMIT 20;
-- 📊 Destination Airport-Level Performance Summary:
--
+------+-----------------------------------------------+-------------------------+--
------+---------------+------------------+---------------+--------------------+
-------------------------+
-- | DEST | DEST_AIRPORT_NAME                             | DEST_CITY
| STATE  | Total Arrivals | On-Time Arrivals  | OTP Rate (%)   | Avg Arrival Delay
 | Cancellation Rate (%)    |
--
+------+-----------------------------------------------+-------------------------+--
------+---------------+------------------+---------------+--------------------+
-------------------------+
-- | ATL  | Hartsfield-Jackson Atlanta Intl               | Atlanta
| GA     | 66,741         | 54,381           | 83.38          | 3.5
 | 2.28                     |
-- | ORD  | Chicago O'Hare Intl                           | Chicago
| IL     | 53,060         | 34,926           | 70.36          | 15.41
 | 6.45                     |
-- | DFW  | Dallas/Fort Worth Intl                        | Dallas-Fort Worth
| TX     | 51,037         | 35,947           | 75.99          | 10.72
 | 7.32                     |
-- | LAX  | Los Angeles Intl                              | Los Angeles
| CA     | 38,463         | 29,915           | 79.28          | 5.19
 | 1.89                     |
-- | DEN  | Denver Intl                                   | Denver
| CO     | 38,300         | 28,536           | 76.15          | 8.58
 | 2.16                     |
-- | IAH  | George Bush Intercontinental                  | Houston
| TX     | 29,820         | 23,221           | 79.59          | 6.16
 | 2.16                     |
-- | PHX  | Phoenix Sky Harbor Intl                       | Phoenix
| AZ     | 29,250         | 23,635           | 81.81          | 4.35
 | 1.23                     |
-- | SFO  | San Francisco Intl                            | San Francisco
| CA     | 28,437         | 21,029           | 76.02          | 7.78
 | 2.73                     |
-- | LAS  | McCarran Intl                                 | Las Vegas
```

```
| NV     | 25,804         | 20,602            | 80.84          | 4.32
 | 1.24                    |
-- | MCO  | Orlando Intl                                  | Orlando
| FL     | 22,586         | 16,996            | 77.23          | 7.34
 | 2.56                    |
-- | LGA  | LaGuardia (Marine Air Terminal)               | New York
| NY     | 21,513         | 12,393            | 65.0           | 17.6
 | 11.38                   |
-- | DTW  | Detroit Metro                                 | Detroit
| MI     | 21,318         | 16,209            | 78.81          | 7.83
 | 3.52                    |
-- | CLT  | Charlotte Douglas Intl                        | Charlotte
| NC     | 20,474         | 16,215            | 82.24          | 3.93
 | 3.7                     |
-- | BOS  | Logan Intl                                    | Boston
| MA     | 20,167         | 12,866            | 70.62          | 13.48
 | 9.66                    |
-- | MSP  | Minneapolis-Saint Paul Intl                   | Minneapolis
| MN     | 20,096         | 15,991            | 81.05          | 5.54
 | 1.82                    |
-- | EWR  | Newark Liberty Intl                           | Newark
| NJ     | 19,621         | 13,681            | 76.36          | 8.57
 | 8.68                    |
-- | SLC  | Salt Lake City Intl                           | Salt Lake City
| UT     | 19,342         | 16,452            | 85.77          | 1.42
 | 0.83                    |
-- | JFK  | JFK Intl (New York Intl)                      | New York
| NY     | 18,858         | 12,797            | 72.45          | 13.91
 | 6.33                    |
-- | SEA  | Seattle-Tacoma Intl                           | Seattle
| WA     | 18,830         | 15,569            | 83.38          | 2.12
 | 0.83                    |
-- | FLL  | Fort Lauderdale-Hollywood Intl                | Ft. Lauderdale
| FL     | 16,184         | 12,172            | 77.41          | 6.22
 | 2.84                    |
--
+------+-----------------------------------------------+-------------------------+--
------+---------------+------------------+---------------+--------------------+
-------------------------+

-- ☑ Interpretation: Destination Airport Arrival Performance
-- ◇ ATL (Atlanta), ORD (Chicago), and DFW (Dallas-Fort Worth) are the top 3
destination airports by volume — all over 50,000 arrivals in 2015.
-- ◇ SLC and SEA show outstanding on-time performance (OTP ≥ 83%) and minimal
delays (avg < 3 mins), indicating operational efficiency.
-- ◇ LGA (LaGuardia) has the **lowest OTP** (65%) and **highest cancellation
rate** (11.38%) — likely due to congestion and weather impacts.
-- △ ORD and JFK exhibit long average delays (13–15 mins) and relatively high
cancellation rates (6–6.5%), requiring capacity and flow management.
-- 📌 This analysis helps identify top-performing arrival airports and those
```

```
needing intervention to improve delay and cancellation metrics.
-- ============================================
-- ☑ Phase 3 - Block 9: Carrier vs Airport Delay Responsibility
-- Project: U.S. Airline Performance & Delay Analysis
-- Dataset: v_flight_data_enriched
-- Author: Shruti Sumadhur Ghosh
-- ============================================

-- 🤔 Objective:
-- Analyze the share of delay causes by airline:
--  ◇ Carrier Delay (airline's responsibility)
--  ◇ External Delay (airport/system/weather/security/late aircraft)

SELECT
    airline_name,
    COUNT(*) AS total_flights,

    -- Total Delayed Flights (departure delay > 15 mins and not cancelled)
    SUM(CASE WHEN CANCELLED = 0 AND DEPARTURE_DELAY > 15 THEN 1 ELSE 0 END) AS
delayed_flights,

    -- % of flights delayed
    ROUND(
        100.0 * SUM(CASE WHEN CANCELLED = 0 AND DEPARTURE_DELAY > 15 THEN 1 ELSE 0
END) /
        NULLIF(SUM(CASE WHEN CANCELLED = 0 THEN 1 ELSE 0 END), 0), 2
    ) AS delay_rate_pct,

    -- Average Delay by Type (only delayed flights)
    ROUND(AVG(CASE WHEN DEPARTURE_DELAY > 15 THEN AIRLINE_DELAY ELSE NULL END), 2)
AS avg_carrier_delay,
    ROUND(AVG(CASE WHEN DEPARTURE_DELAY > 15 THEN LATE_AIRCRAFT_DELAY ELSE NULL
END), 2) AS avg_late_aircraft_delay,
    ROUND(AVG(CASE WHEN DEPARTURE_DELAY > 15 THEN WEATHER_DELAY ELSE NULL END), 2)
AS avg_weather_delay,
    ROUND(AVG(CASE WHEN DEPARTURE_DELAY > 15 THEN AIR_SYSTEM_DELAY ELSE NULL END),
2) AS avg_air_system_delay,
    ROUND(AVG(CASE WHEN DEPARTURE_DELAY > 15 THEN SECURITY_DELAY ELSE NULL END), 2)
AS avg_security_delay

FROM v_flight_data_enriched
GROUP BY airline_name
ORDER BY delay_rate_pct DESC;
--
+-----------------------------+--------------+-----------------+---------------
--+--------------------+-------------------------+--------------------+---------
--------------+------------------------+
-- | Airline Name                | Total Flights | Delayed Flights  | Delay Rate
(%)   | Avg Carrier Delay   | Avg Late Aircraft Delay | Avg Weather Delay   | Avg
Air System Delay  | Avg Security Delay     |
```

```
--
+------------------------------+--------------+------------------+---------------
--+--------------------+-------------------------+--------------------+---------
---------------+------------------------+
-- | Frontier Airlines Inc.       | 14,669        | 4,671            | 32.36
    | 19.44              | 39.44                   | 1.50               | 25.70
              | 0.00                   |
-- | American Eagle Airlines Inc. | 65,513        | 17,835           | 30.86
    | 19.74              | 33.33                   | 8.17               | 12.73
              | 0.14                   |
-- | Spirit Air Lines             | 19,612        | 5,159            | 26.94
    | 16.40              | 22.90                   | 1.88               | 32.44
              | 0.13                   |
-- | JetBlue Airways              | 48,157        | 12,194           | 26.79
    | 22.94              | 34.30                   | 3.51               | 13.91
              | 0.22                   |
-- | United Air Lines Inc.        | 87,606        | 22,173           | 26.02
    | 24.36              | 24.37                   | 5.38               | 11.71
              | 0.00                   |
-- | Southwest Airlines Co.       | 221,586       | 45,085           | 20.97
    | 17.44              | 28.81                   | 2.11               | 5.11
              | 0.04                   |
-- | Atlantic Southeast Airlines  | 111,206       | 21,713           | 20.69
    | 27.06              | 31.79                   | 2.49               | 11.32
              | 0.00                   |
-- | Skywest Airlines Inc.        | 107,099       | 21,235           | 20.40
    | 22.75              | 38.09                   | 4.06               | 9.60
              | 0.07                   |
-- | American Airlines Inc.       | 97,549        | 18,633           | 20.06
    | 29.35              | 30.13                   | 5.35               | 8.41
              | 0.07                   |
-- | Virgin America               | 10,403        | 1,758            | 17.36
    | 14.60              | 30.25                   | 3.03               | 29.21
              | 0.10                   |
-- | Delta Air Lines Inc.         | 147,486       | 24,125           | 16.68
    | 29.08              | 23.04                   | 9.87               | 11.71
              | 0.04                   |
-- | US Airways Inc.              | 73,942        | 11,582           | 16.36
    | 26.31              | 22.85                   | 3.45               | 11.86
              | 0.18                   |
-- | Alaska Airlines Inc.         | 29,614        | 3,468            | 11.79
    | 22.11              | 31.14                   | 3.84               | 9.03
              | 0.10                   |
-- | Hawaiian Airlines Inc.       | 14,133        | 1,195            | 8.48
    | 25.16              | 24.61                   | 2.68               | 0.51
              | 0.16                   |
--
+------------------------------+--------------+------------------+---------------
--+--------------------+-------------------------+--------------------+---------
---------------+------------------------+
```

```
-- ☑ Interpretation: Delay Responsibility Analysis
-- 🔺 Frontier Airlines and American Eagle Airlines have the highest delay rates (30–32%),
--    with substantial average delays due to late aircraft and carrier-related issues —
--    indicating significant operational inefficiencies.
-- ⚠ Spirit Airlines and JetBlue also exhibit high delay rates (~27%), primarily driven by
--    late aircraft and air traffic control (air system) delays.
-- 🔸 American Airlines and Atlantic Southeast Airlines show moderate delay rates (~20%),
--    but high average carrier delays (~27–29 minutes), hinting at airline-level process gaps.
-- ☑ Delta, US Airways, and Alaska Airlines maintain relatively low delay rates (12–17%),
--    reflecting stronger internal operations and scheduling efficiency.
-- ▨ Hawaiian Airlines stands out with the lowest delay rate (8.5%) and minimal external
--    delay impact — consistent with its industry-leading on-time performance.

-- 📌 Insight:
-- This breakdown helps distinguish between delays caused by the airline's own operations
-- (carrier/late aircraft) versus external factors (weather, air traffic control, security).
-- 🔸 Airlines with high self-induced delays should focus on:
--     - Improving ground operations and turnaround processes
--     - Optimizing fleet scheduling and crew management
--     - Minimizing cascading delays from previous flights
-- 🔸 Airports and regulators can use this view to:
--     - Address recurrent system-level delays (e.g., airspace congestion)
--     - Coordinate with airlines to streamline scheduling at peak hours
-- 🎯 Actionable next step: Deep-dive delay analysis by time of day, route, and region
-- in subsequent blocks.

-- ==========================================
-- ☑ Phase 3: Exploratory Data Analysis (EDA)
-- Project: U.S. Airline Performance & Delay Analysis
-- Author: Shruti Sumadhur Ghosh
-- ==========================================

-- 🔸 Block 1: Flight Volume Summary
-- ☑ Interpretation: ATL is the busiest origin with 66K+ departures. ORD, DFW, and LAX follow. High-frequency airports are potential hubs and key for operational focus.

-- 🔸 Block 2: Monthly Trends
-- ☑ Interpretation: July is the busiest month, followed by June and August—likely due to summer travel. February has the fewest flights.
```

```
-- ◇ Block 3: Weekday Trends
-- ☑ Interpretation: Fridays have the highest flight volumes, followed by Thursdays. Saturdays and Sundays see the least traffic, suggesting lower business travel.

-- ◇ Block 4: Time of Day Trends
-- ☑ Interpretation: Most flights are scheduled between 6 AM and 7 PM. Early morning (6–9 AM) is peak departure time. Red-eye flights are limited.

-- ◇ Block 5: Airline-Level Performance Summary
-- ☑ Interpretation: Alaska and Hawaiian Airlines lead with >85% OTP and <1% cancellations. American Eagle and Frontier show weak performance with high delay and cancellation rates. Delta has balanced performance with high volume and good OTP.

-- ◇ Block 6: Origin Airport-Level Departure Performance
-- ☑ Interpretation: ATL leads in departures with strong OTP (81.6%). ORD shows weakest OTP (67%) and highest avg delay (20 mins). SLC and SEA are top performers with >85% OTP and <7 min delays.

-- ◇ Block 7: Route-Level Performance Summary
-- ☑ Interpretation: Hawaiian Airlines dominates intra-island routes with ≥90% OTP and negligible cancellations. Delta's MCO → ATL shows early arrivals. LGA and ORD routes face high cancellations (5–7%).

-- ◇ Block 8: Destination Airport-Level Arrival Performance
-- ☑ Interpretation: ATL has highest arrivals with 83% OTP. ORD and DFW have weaker performance with high delays (10–15 mins) and cancellations (6–7%). LGA and BOS show OTP challenges and high cancellation rates.

-- ◇ Block 9: Delay Causes by Airline
-- ☑ Interpretation: Frontier and American Eagle have the highest delay rates (~30%). Frontier suffers from systemic and turnaround delays. Hawaiian Airlines is the most reliable (8.48% delay rate). Delta and Southwest manage delays well at scale.
-- ==========================================
-- ☑ Phase 4: Dashboard Development (Power BI)
-- Project: U.S. Airline Performance & Delay Analysis
-- Author: Shruti Sumadhur Ghosh
-- ==========================================

-- 📌 All cleaned and enriched data views (e.g., v_flight_data_enriched, new_airports, new_airlines) are now ready for Power BI import.

-- 🧭 Next Step:
-- ➤ Load final views/datasets into Power BI
-- ➤ Build visuals based on key metrics:
--    - Airline/Route/Airport performance
--    - Delay & Cancellation trends
```

```
--    - OTP vs delay type comparison
-- ➤ Design an interactive, insight-driven dashboard
```