SkyHack 3.0: United Airlines

Team Name - Stat Pirates

Team Members - Ritvik Kumar, Rujhan N Sharma

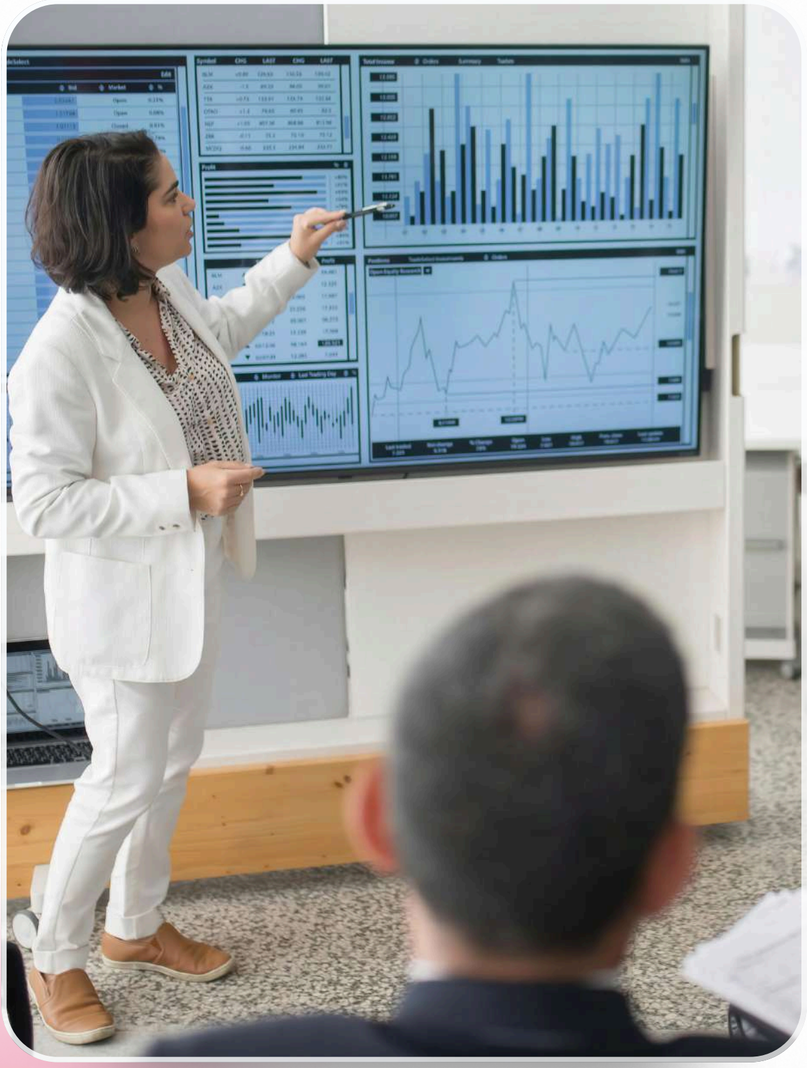


Table of Content

Data Cleaning Process

Data Inspection

- Reviewed dataset structure, column types, and identified issues like missing values or invalid formats.
- Example: Checking for nulls in departure_delay_minutes or invalid date formats in scheduled_departure_date_local.

Handling Missing Data

- Filled or removed null and inconsistent entries to maintain accuracy and reliability.
- Example: Replacing null delay values with 0 or the average delay of similar flights.

Data Type Correction

- Converted columns to proper types (e.g., text → DATETIME, string → INT) for accurate processing.
- Example: Converting scheduled_departure_datetime_local from text to DATETIME.

Removing Duplicates & Irrelevant Entries

- Eliminated repeated or irrelevant records to ensure clean and unique data.
- Example: Deleting repeated flight entries for the same flight_number and scheduled_departure_date_local.

Binary Value Conversion

- Replaced 0/1 with N/Y for better readability and interpretation in reports.
- Example: Changing turn_time_violation_flag = 1 to Y and 0 to N for easier understanding in reports.

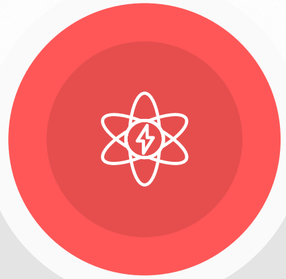
Data Organization Process

Feature Engineering

- Created new columns to derive deeper insights.
- Example: Calculated Flight_Difficulty_Score using Ground_Time_Score, Baggage_Score, Passenger_Service_Score, and Operational_Complexity.

Grouping Related Attributes

- Combined similar columns into logical categories for easier analysis.
- Example: Grouped TOTAL_BAGS, TRANSFER_BAGGAGE, and HOT_TRANSFER_BAGGAGE under Baggage Metrics.



Creating Derived Levels or Categories

- Converted continuous scores into categorical levels for clarity.
- Example: Classified Flight_Difficulty_Score into levels — Low, Medium, High.

Joining Tables

- Merged related datasets to enrich the analysis.
- Example: Joined FINAL_DATA with airport info table using Airport_IATA_Code.

Data Standardization

- Ensured consistent units, formats, and naming conventions across all columns.
- Example: Standardized all airport codes to uppercase and aligned datetime formats.

Flight Difficulty Score

It systematically quantifies the relative complexity of each flight using the datasets provided

Ground Time Score

Reflects how efficiently an aircraft is prepared between flights; shorter or delayed ground times increase operational difficulty.

Baggage Score

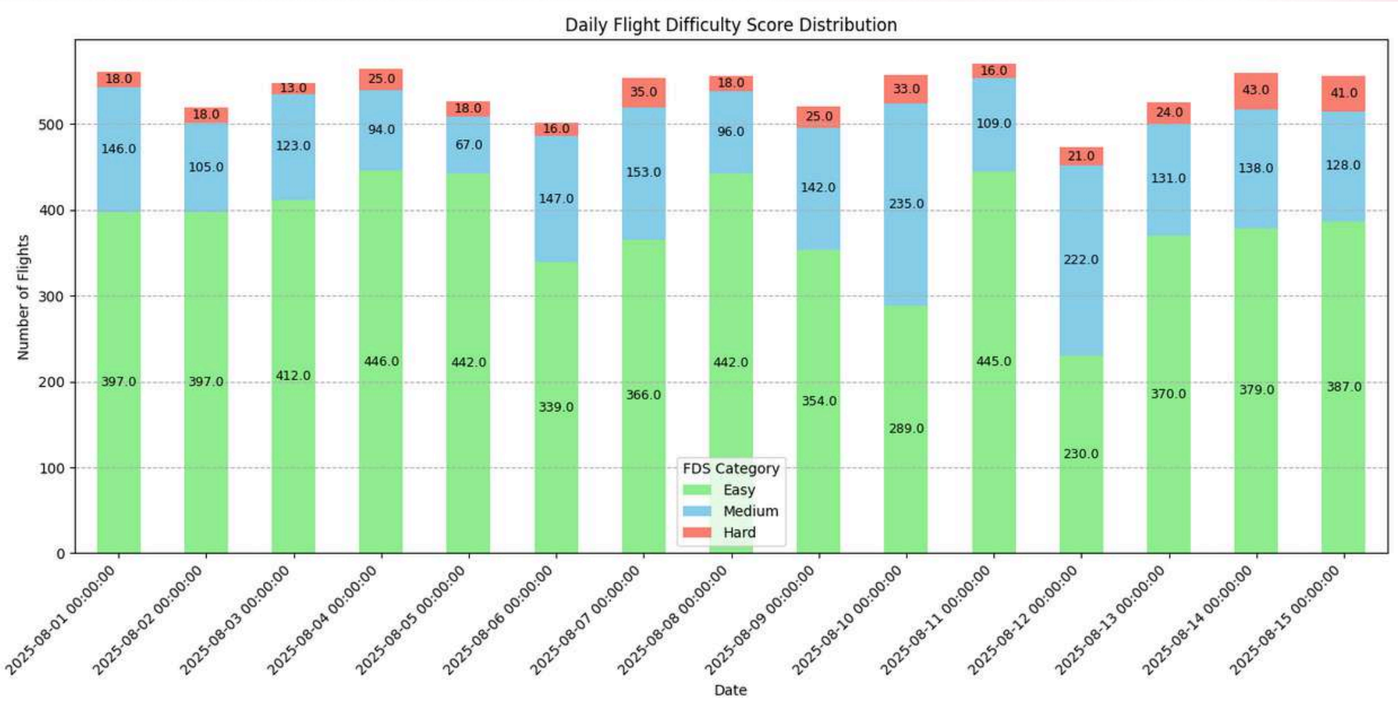
Measures the handling complexity and volume of baggage, where higher loads or transfer bags raise coordination challenges.

Pax Service Score

indicates passenger-related operational strain, such as special assistance requests or last-minute bookings.

Operational Complexity

reflects how much a flight's performance deviates from schedule — higher delays imply higher operational strain.

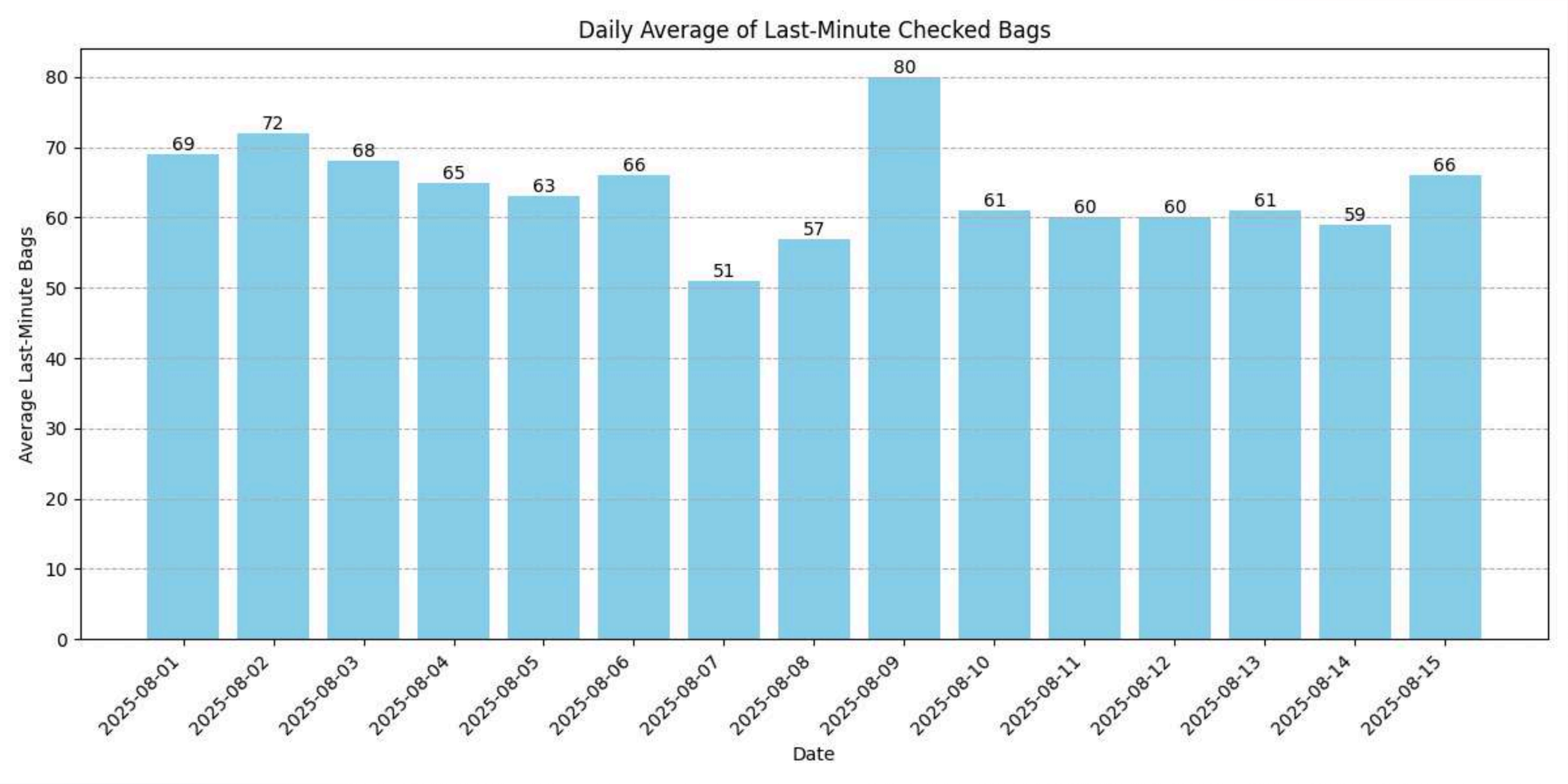


Flight Difficulty Score

Flight Difficulty Score = 0.25 * Ground Time Score + 0.25 * Baggage Score + 0.25 * Pax Service Score + 0.25 * Operational Complexity

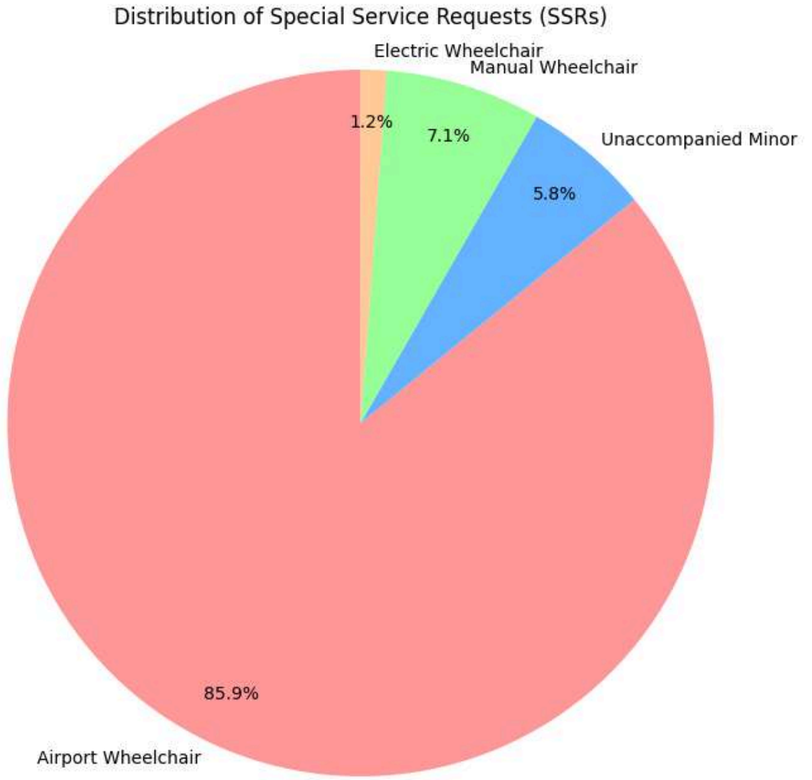
Last-Minute Checked bags

It tells us about the average bags checked-in at the last moment each day



Special Service Request

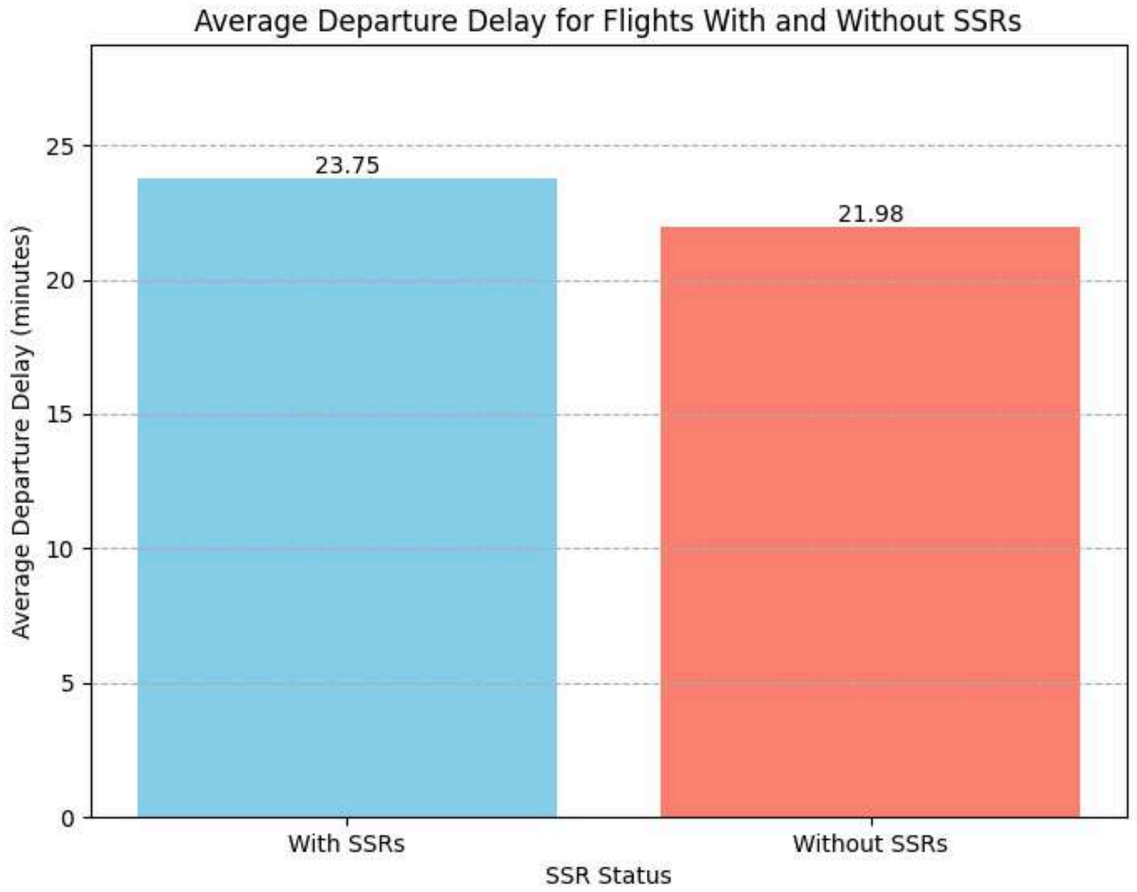
The analysis shows that the majority of special service requests (85.9%) are for airport wheelchair assistance, indicating that mobility support is the most common passenger need. Requests for manual wheelchairs (7.1%) and unaccompanied minors (5.8%) form smaller segments, while electric wheelchair requests (1.2%) are minimal.



Special-Services Delay Analysis

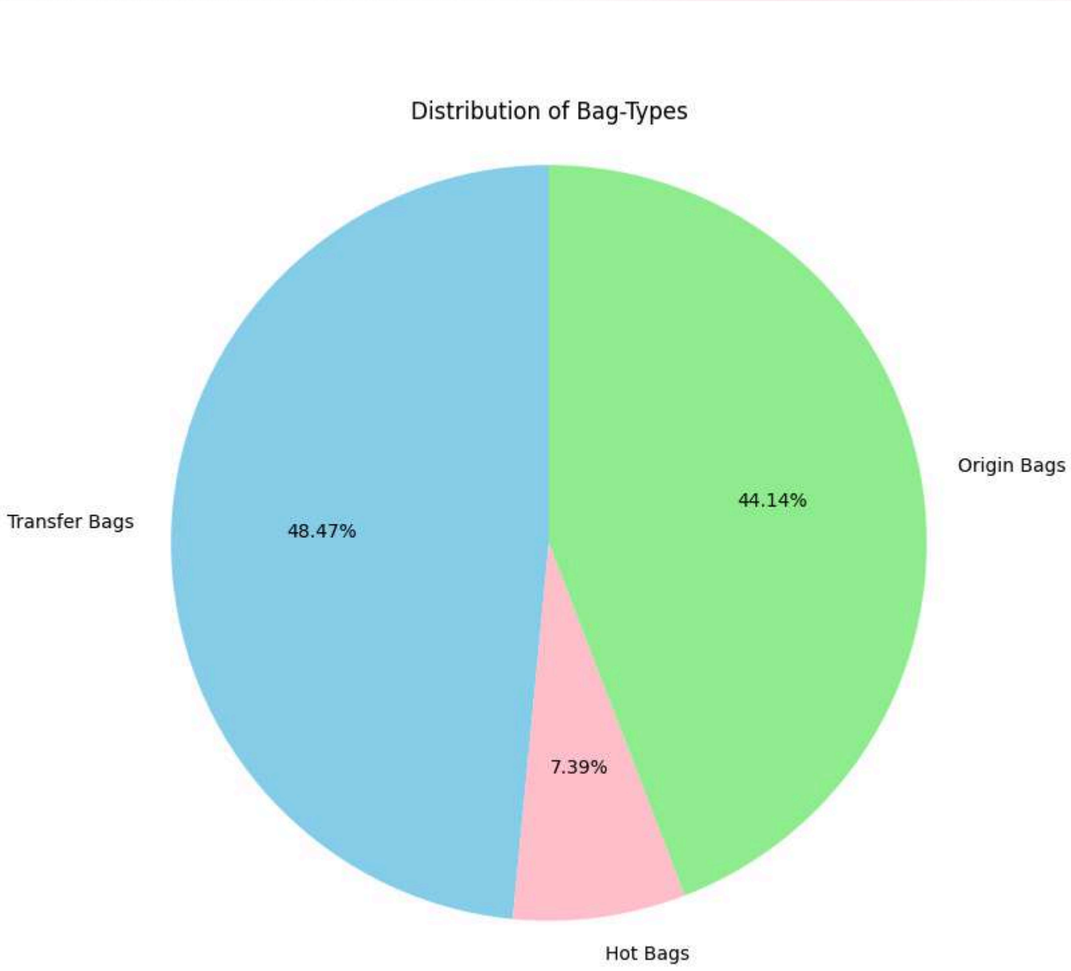
It helps us in analyzing delay caused due special-service demands

Flights that included Special Service Requests (SSRs) experienced a slightly higher average departure delay (23.75 minutes) compared to flights without SSRs (21.98 minutes).



Baggage Types

The analysis reveals that transfer bags (48.47%) constitute the largest share of total baggage, closely followed by origin bags (44.14%), while hot bags (7.39%) represent a small portion.



Thank You

Link : https://github.com/Rujhan052/skyhack