Operational Flight Complexity Analysis and Difficulty Scoring

Presented By:

Anshu Rashi

Arihant Sharma

Team Name: Hacks on Crack

TABLE OF CONTENTS

01 Executive Summary

02 Understanding Data

03 Data cleaning

04 Deliverables 1

05 Deliverables 2

06 Deliverables 3

07 Bonus

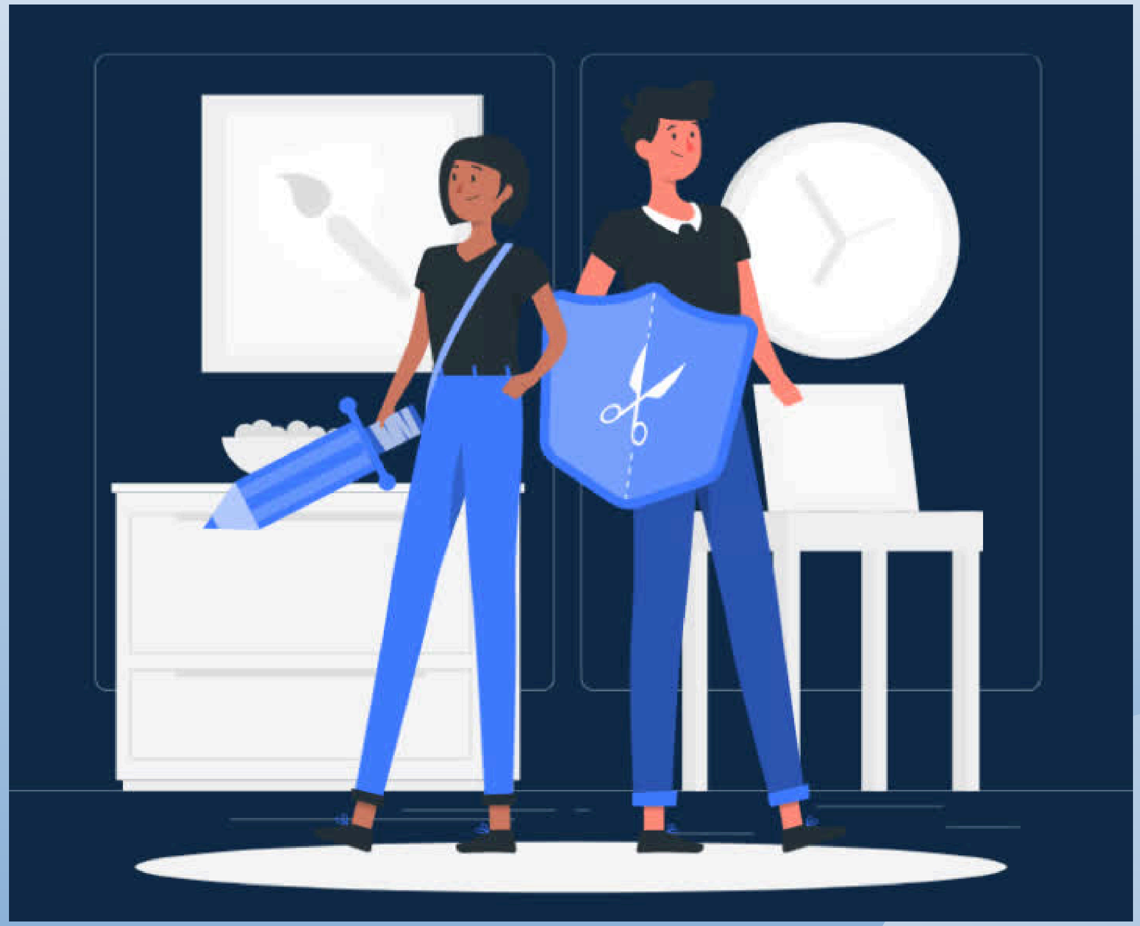
08 References & Appendices

Executive Summary

- **The Problem:** Identifying difficult flights currently relies on inconsistent team experience. This is a reactive approach that risks delays and inefficient resource use.

- **The Solution:** A Flight Difficulty Score. We've developed a single, data-driven score to quantify the operational complexity of every flight. This moves us from guesswork to proactive, data-informed planning.

- **The Vision:** To anticipate challenges and manage difficult flights before they become a problem.



What Makes a Flight Difficult?

Tight Schedules

A tight schedule between landing and the next takeoff puts immense pressure on ground crews to perform all tasks quickly.

High Bagagge Volume

More bags require more time and physical effort for sorting, loading, and unloading, increasing the risk of delays.

Complex Customer Requests

Flights with many passengers requiring assistance (like wheelchairs or unaccompanied minors) demand extra coordination and staff resources.

High Passenger Load

A full flight increases the complexity and time needed for boarding, deplaning, and managing all passenger interactions.

Understanding Data

- Scheduling Appears to Prioritize Aircraft Type Over Actual Load
- The number of children on a plane likely has a near-zero correlation with the minimum time required to service the aircraft.

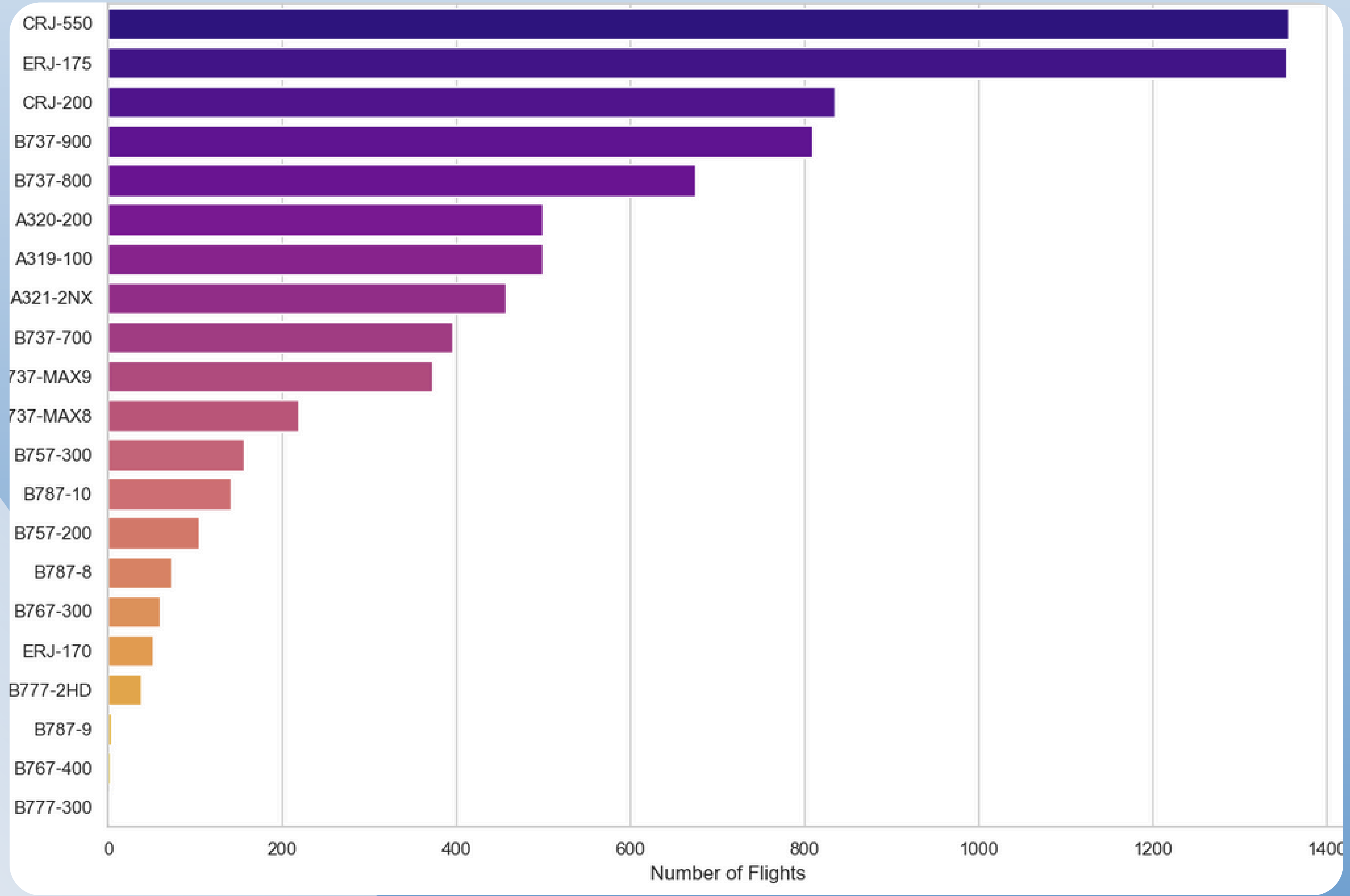


Fig 1: Most popular fleet type

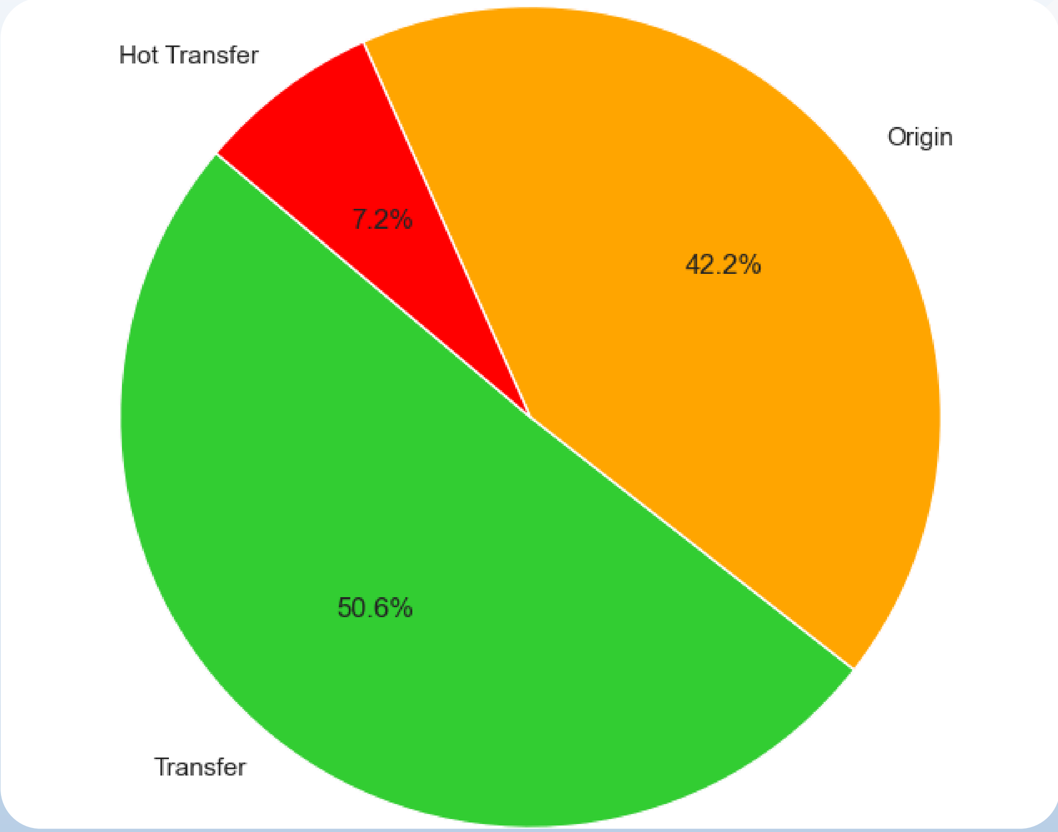


Fig 2: Carrier Distribution

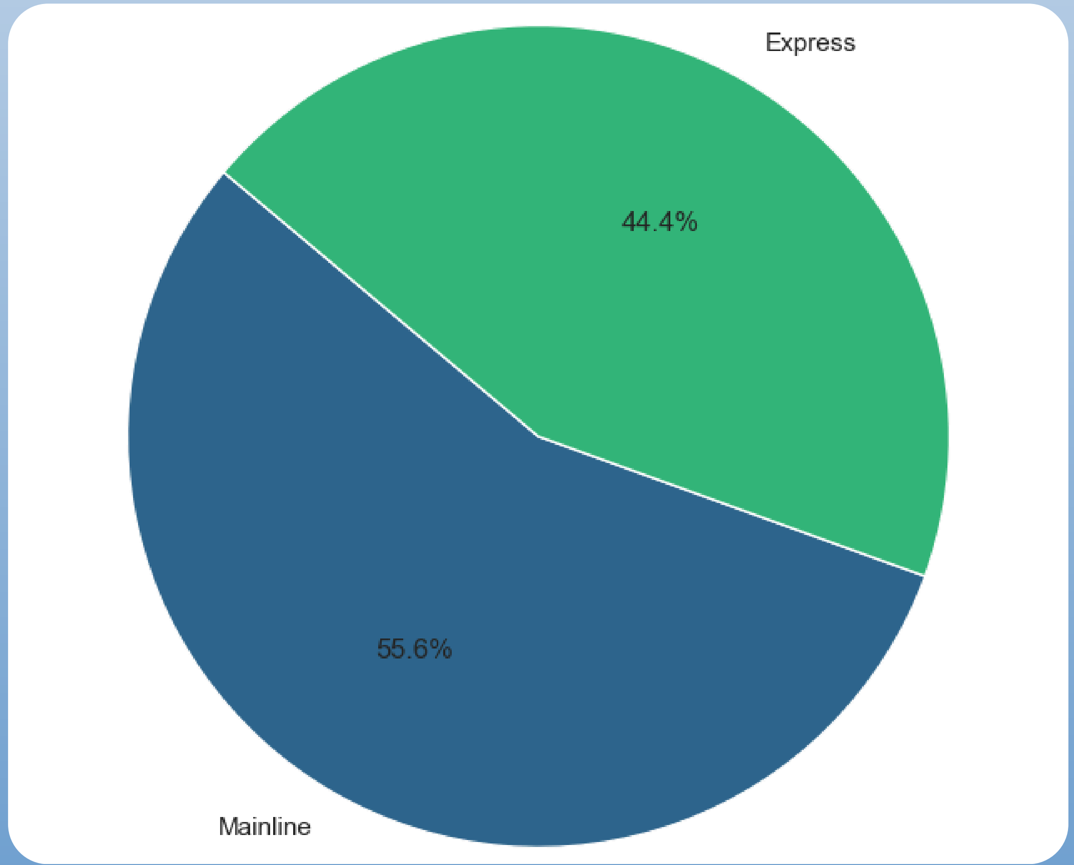
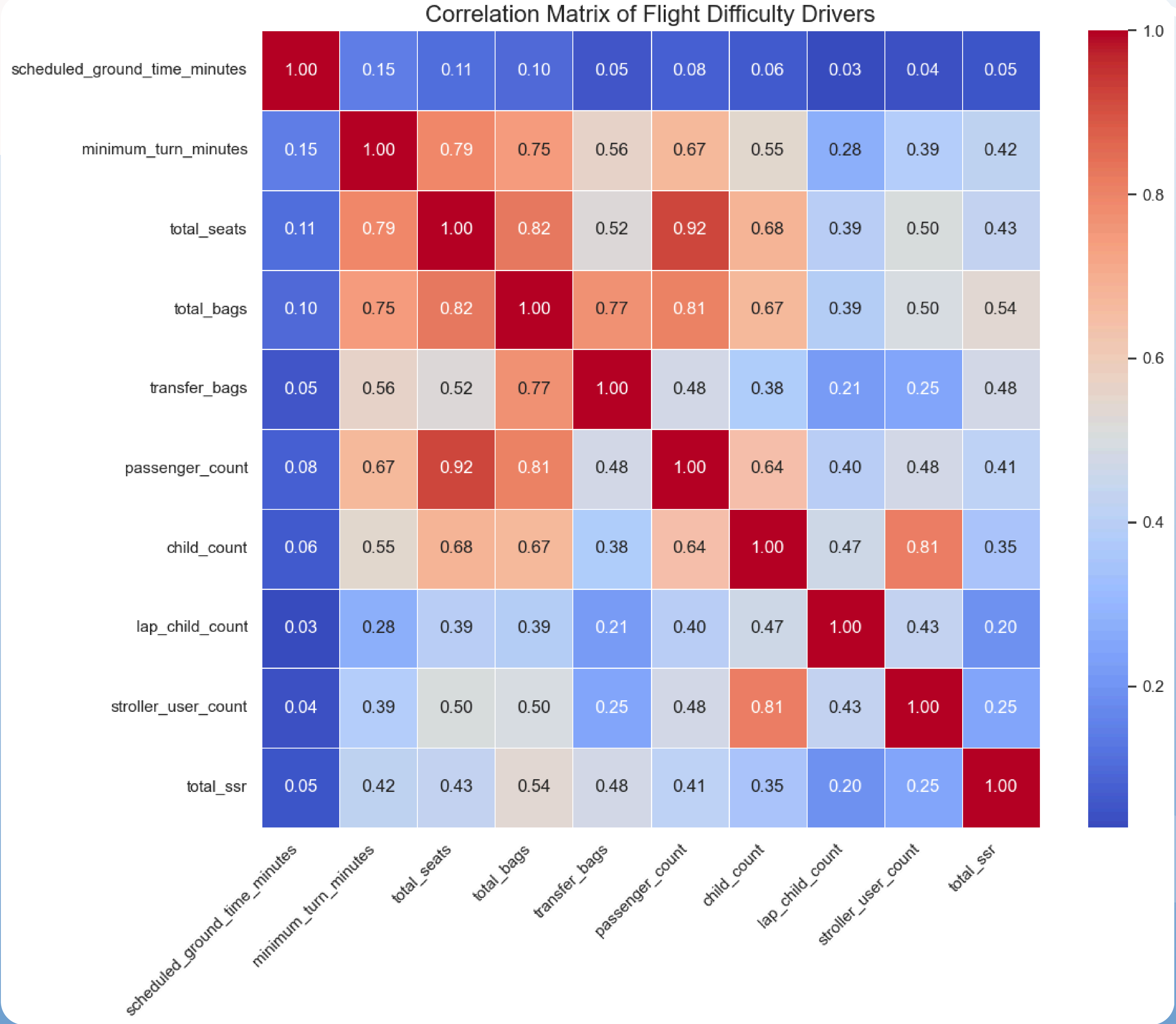
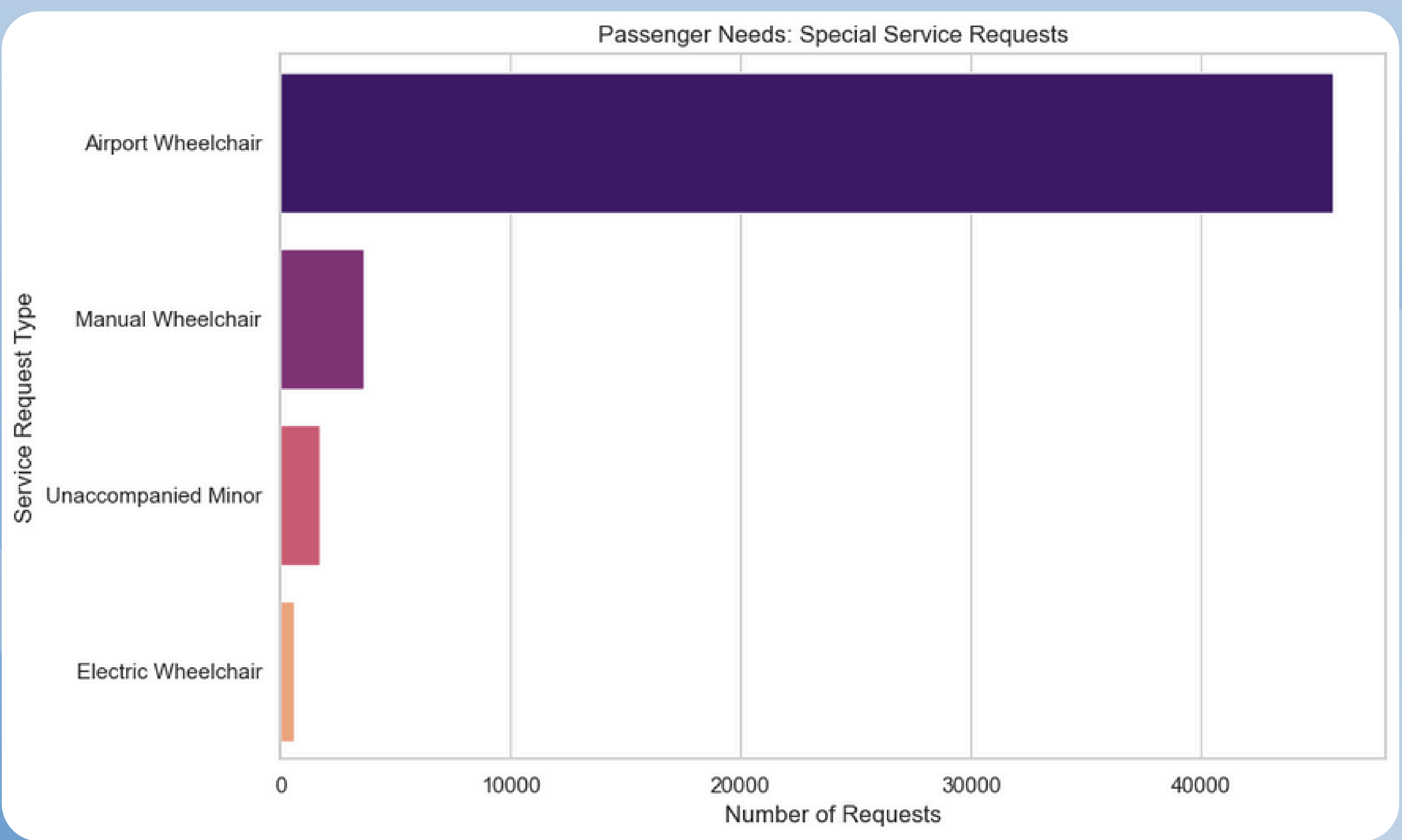


Fig 3: Baggage Demographic



- “minimum_turn_minutes” is influenced by the size of the aircraft, which dictates the scheduled ground time.
- The “total_ssr” (Special Service Requests) count has a strong positive correlation with “passenger_count”.
- There is a very strong positive correlation between “passenger_count”, “total_bags”, and “total_seats”. This is the most obvious but most important finding: flights with more seats tend to have more passengers, who in turn bring more bags.



DATA CLEANING

Date & Time Formatting

Converted all key date/time columns to datetime type for accurate delay calculations and merging.

Duplicate & Missing Data Handling

Converted all key date/time columns to datetime type for delay calculations and merging. Also Dropped flights missing essential time info.

Data Integration

Merged multiple sources (Flights, PNR, Bags, SSR) into a single Master DataFrame for unified analysis.

Standardized Column Names

Converted all column headers to lowercase and stripped extra spaces for consistency.

Derived Key Metric

Calculated Departure Delay (minutes) = Actual Departure – Scheduled Departure.

Outcome

- Clean, reliable dataset ready for operational analysis
- 8,063 valid flight records retained.

Deliverables 1

01 Exploratory Data Analysis:

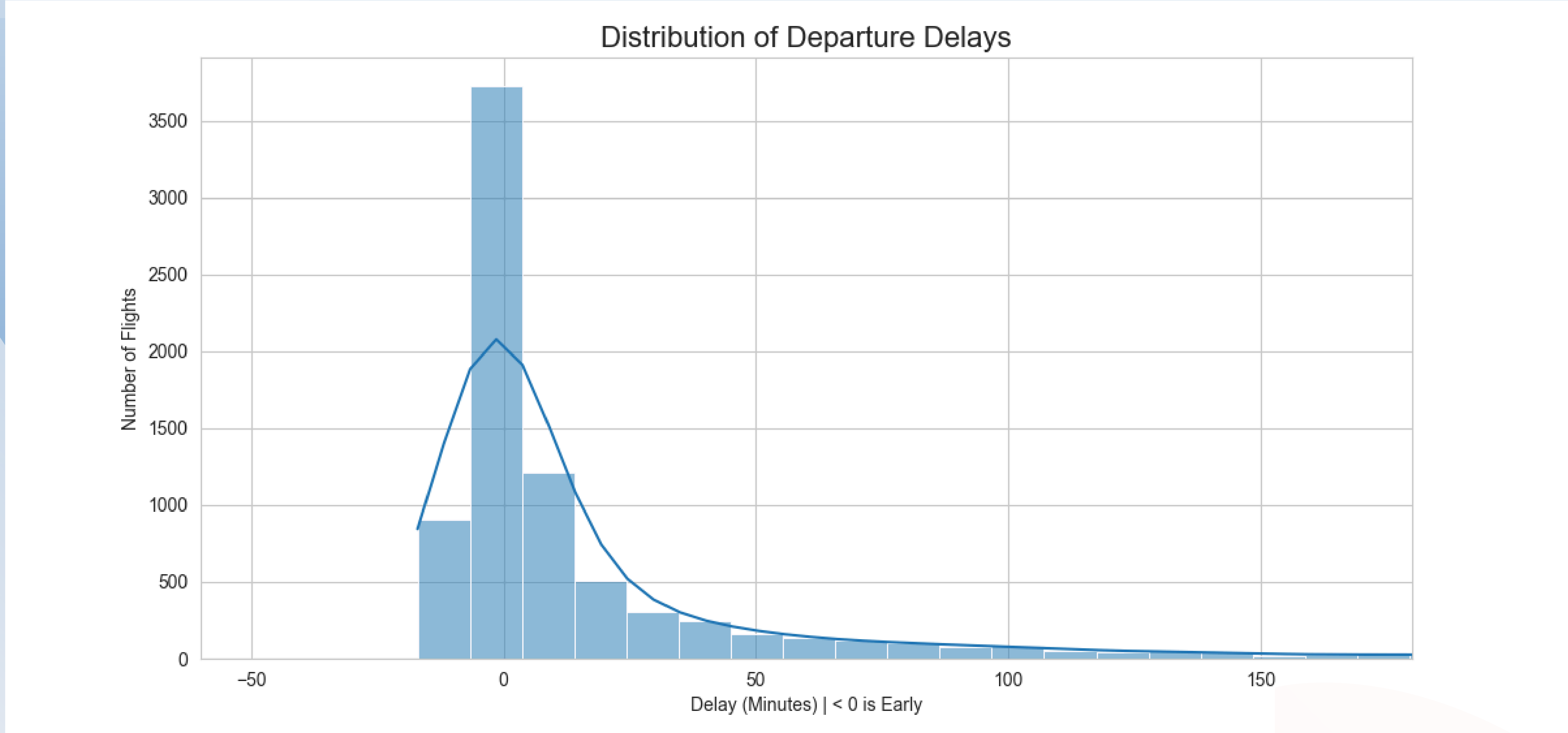
- What is the average delay and what percentage of flights depart later than scheduled?
- How many flights have scheduled ground time close to or below the minimum turn mins?
- What is the average ratio of transfer bags vs. checked bags across flights?
- How do passenger loads compare across flights, and do higher loads correlate with operational difficulty?
- Are high special service requests flights also high-delay after controlling for load?

What is the average delay and what percentage of flights depart later than scheduled?

Key Findings:

- Nearly half of all flights **(49.7%)** depart late.
- As shown in the below graph, when a flight is delayed, the average delay is a substantial **47** minutes.

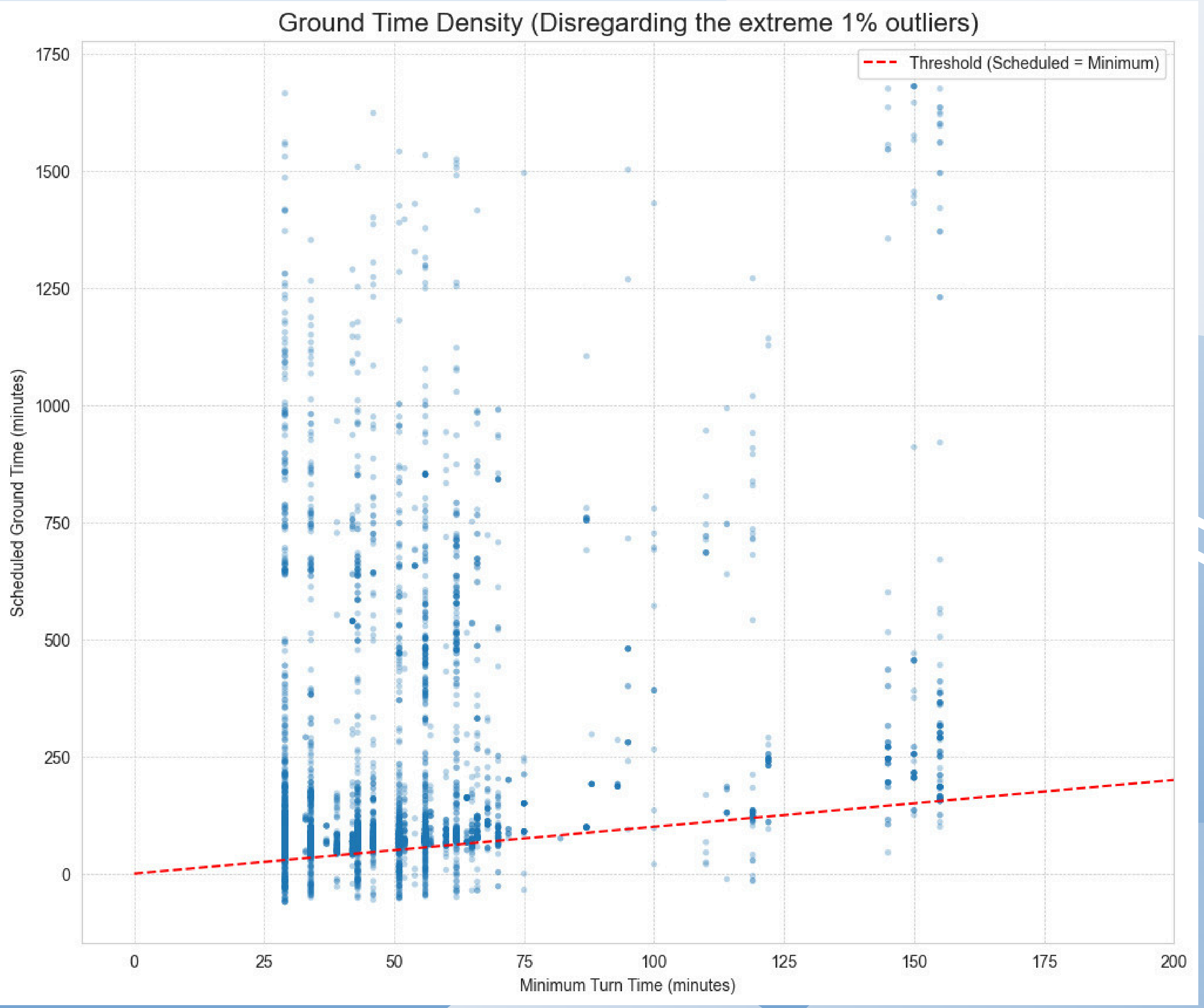
Business Insight: Departure delays are not an exception but a common occurrence, affecting roughly every other flight. The severity of these delays presents a major risk to network integrity, crew scheduling, and customer satisfaction, highlighting a critical need for operational intervention.



How many flights have scheduled ground time close to or below the minimum?

Key Findings:
A staggering **616** flights were scheduled with a turnaround time below the operational minimum. An additional **147** flights operated on a razor-thin margin with **less than a 10%** time buffer.

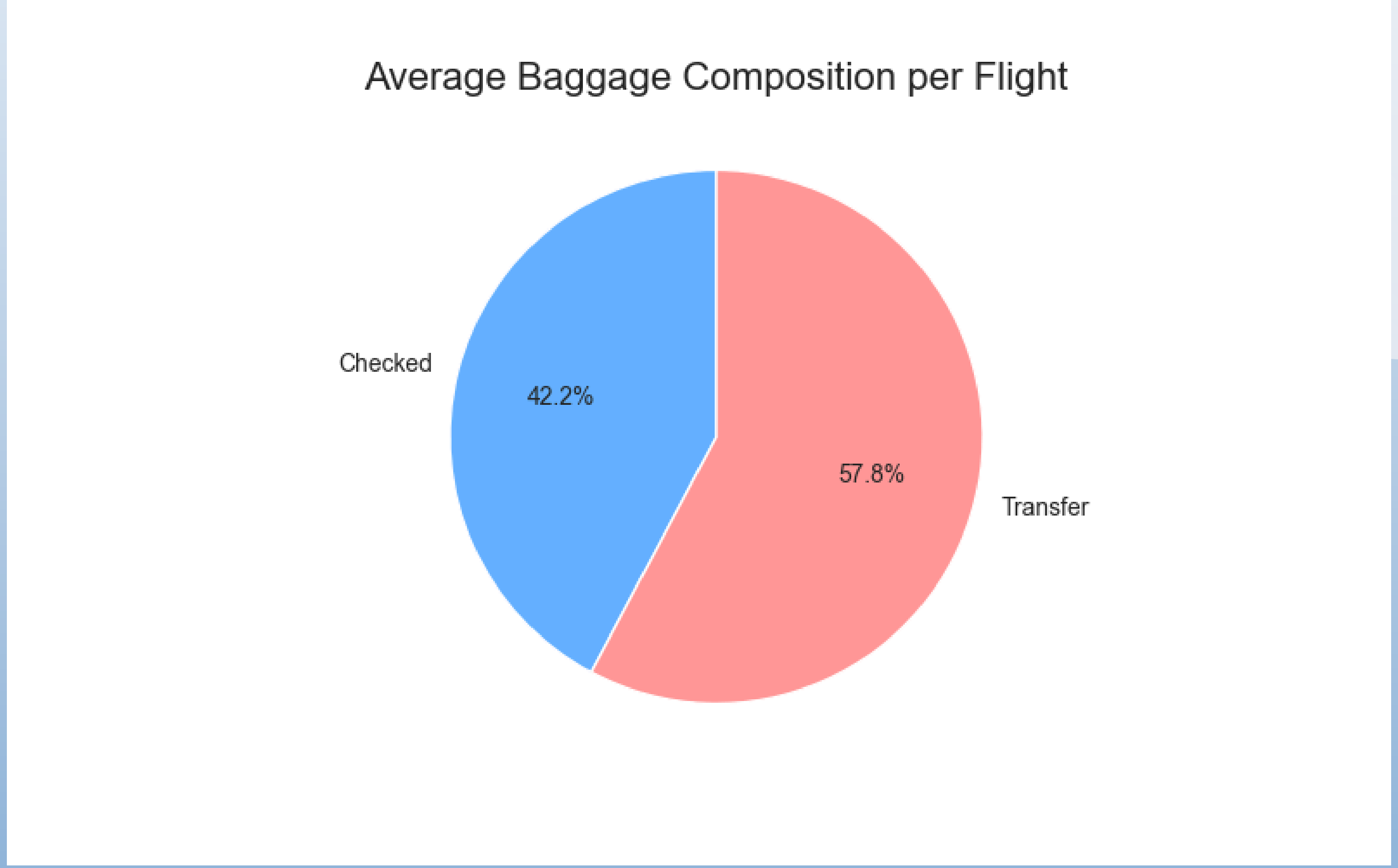
Business Insight: Many flights are operationally at risk before the day even begins. This lack of scheduled buffer time creates a fragile system where minor disruptions can easily cascade into significant delays, highlighting a major structural challenge to on-time performance.



What is the average ratio of transfer bags vs. checked bags across flights?

Key Findings:
An overwhelming majority, two-thirds **(66.8%)** of all bags handled are transfer bags, not bags originating from Chicago.

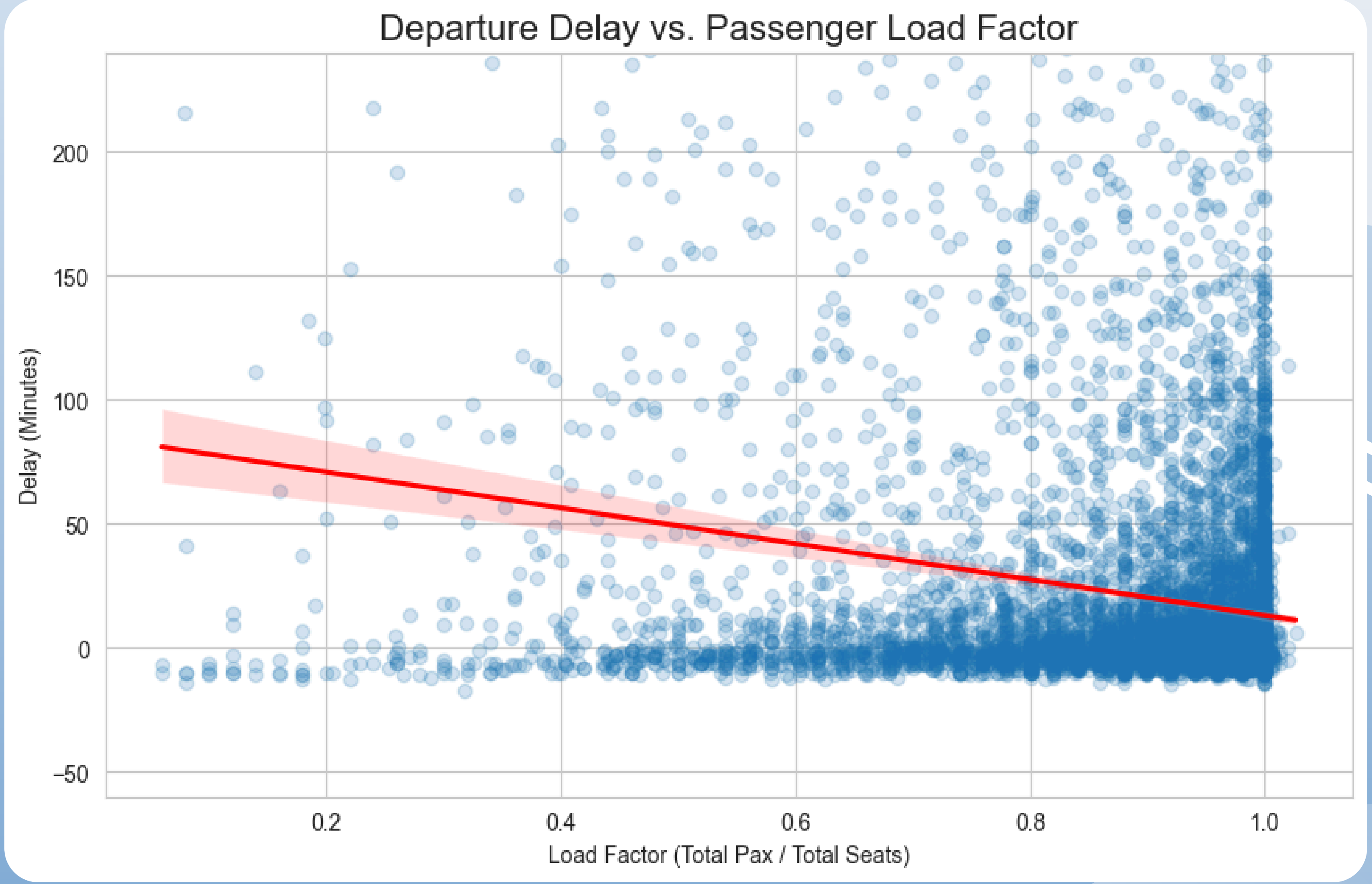
Business Insight: ORD operates fundamentally as a hub, not just an origin/destination airport. The extreme volume of transfer baggage is a primary driver of ground-handling complexity. This makes efficient bag sorting and transfer not just a routine task, but a critical factor for the entire network's on-time performance.



How do passenger loads compare across flights, and do higher loads correlate with operational difficulty?

Key Findings:
The data shows a **(-0.174)** correlation between between how full a flight is (load factor) and its departure delay.

Business Insight:
The passenger load factor is a critical predictor of complexity, but the difficulty escalates non-linearly. As a flight approaches 100% capacity, it hits a "tipping point" where challenges like slow boarding and bin space competition spike dramatically. This makes the load factor a powerful leading indicator to proactively allocate resources to the specific flights most at risk of operational friction and delay.



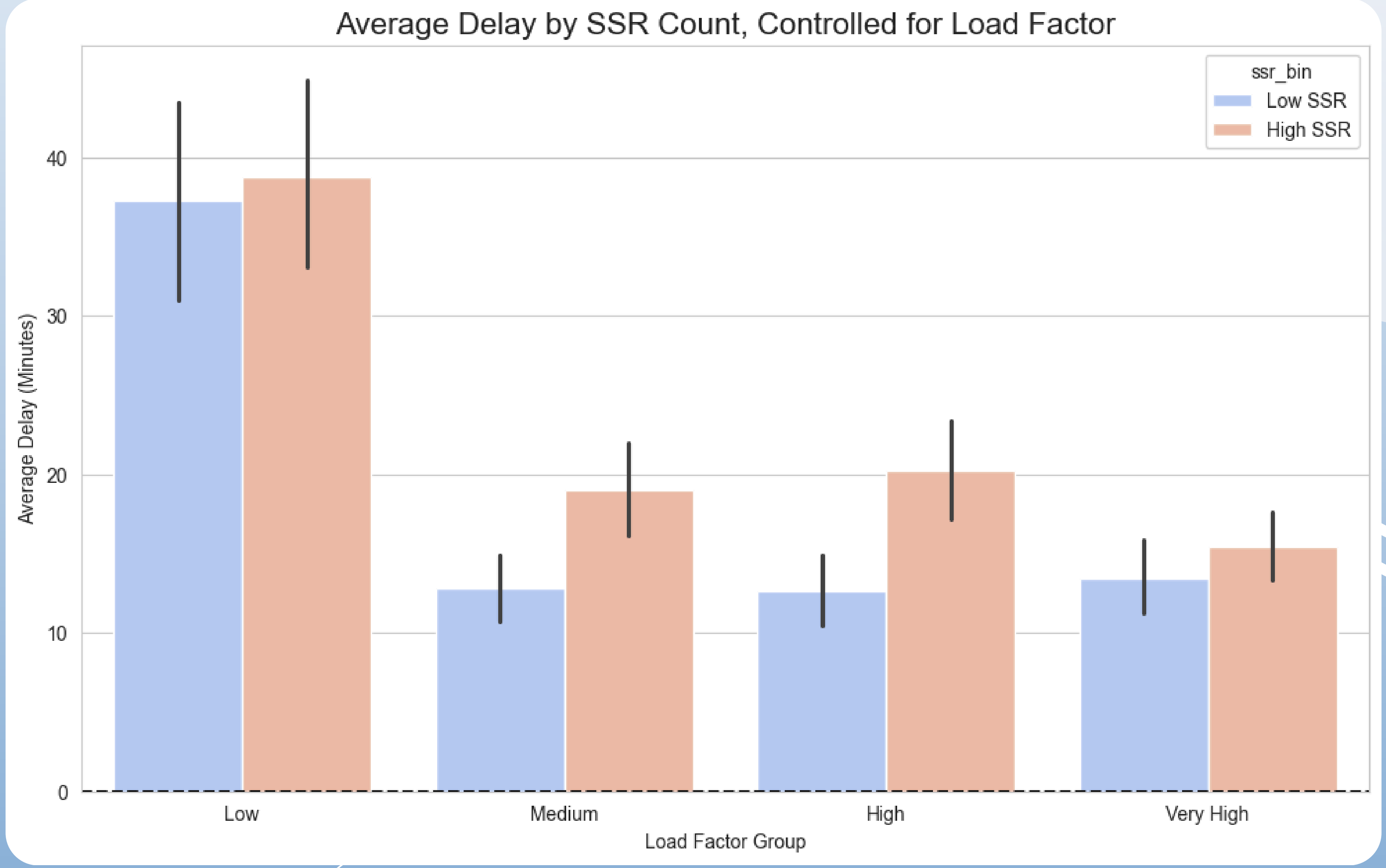
Are high special service requests flights also high-delay after controlling for load?

Key Findings:

Flights with a disproportionately high number of SSRs for their size (e.g., a small regional jet with numerous wheelchair requests) are statistically more likely to depart late than similarly loaded flights with fewer service requests.

Business Insight:

Special service requests are an independent risk factor for delays, not just a byproduct of high passenger loads. Each request, particularly for mobility assistance, introduces a fixed time cost at the gate for coordination and handling. This insight allows for a more nuanced risk assessment, enabling teams to flag flights where the composition of passengers—not just the count—poses a direct threat to on-time performance, allowing for proactive resource allocation.



Deliverables 2

02 Flight Difficulty Score Development:

Build a systematic daily-level scoring approach that resets every day. Below are the outputs required:

- Ranking: Within each day, order flights by their difficulty score so that the highest-ranked flights represent the most difficult to manage.
- Classification: Group flights into three categories — Difficult, Medium, or Easy — based on rank distribution.

Methodology for Flight Difficulty Score

Our 3-Step Scoring Framework Description:

1. **Feature Engineering:**
   - Created a comprehensive set of metrics from all five datasets to measure different aspects of flight complexity.
   - Features are grouped into categories like Ground Time Pressure, Baggage Complexity, and Passenger Service Needs.
2. **Daily Scoring & Ranking:**
   - The core of the model: flights are processed on a day-by-day basis.
   - Features are normalized daily, meaning a flight's score is relative to the other flights on the same day.
   - A weighted sum of these normalized features produces the final Difficulty Score.
3. **Classification:**

Based on its daily rank, each flight is classified into one of three categories:
Difficult (Top 20%), Medium (Middle 50%), or Easy (Bottom 30%).

Business Insight: This daily-resetting approach ensures the model is adaptive and highlights the most challenging flights for that specific day's operation, rather than using a fixed, absolute scale.

Classification

Business Insight:

- A consistent 20% of flights are identified as 'Difficult'.
- The vast majority of flights fall into the 'Medium' (50%) category, representing the standard operational baseline, while the remaining 30% are classified as 'Easy', requiring minimal proactive oversight.

- This 20/50/30 split provides a clear, data-driven starting point for resource allocation.

- Instead of treating all flights equally, the ground team can now concentrate their attention and experienced personnel on the 20% of flights that pose the highest risk, allowing for more efficient and targeted interventions to prevent delays before they happen.

Key Drivers - Feature Engineering

Feature Category	Engineered Features	Rationale
Ground Time Pressure	**f_ground_time_pressure**	Measures the ratio of required vs. scheduled turn time. A higher value indicates less buffer and higher risk.
Baggage Complexity	**f_transfer_bag_ratio, f_bags_per_pax**	High volumes of transfer bags and bags per passenger significantly increase ground handling workload and complexity.
Passenger Service Needs	**f_ssr_count, f_child_ratio**	High counts of Special Service Requests (e.g., wheelchairs) and children require extra staff coordination and time.
Flight Characteristics	**f_is_wide_body, f_load_factor, f_haul_category**	Larger aircraft, fuller flights, and longer hauls all contribute to increased operational resource requirements.

Technical Approach: Features were derived by aggregating and merging all five datasets into a single, flight-level master table. Ratios and flags were created to normalize for factors like aircraft size and passenger count.

The Scoring Model in Action

How the daily score is calculated and applied.

Scoring Logic:
For each day, every feature's value is scaled from **0 to 1 (MinMaxScaler).**
The final score is a weighted sum: **Score = (w1 * feat1_norm) + (w2 * feat2_norm) + ...**Weights are assigned based on the feature's operational impact discovered during the EDA.

Final Output: A single CSV file containing every flight, its unique feature values, its final scaled score (1-100), its rank for that day, and its final classification ('Difficult', 'Medium', 'Easy').
This provides a daily "watch list" of the most challenging flights.

A	B	C	D	E	F	G	H	I
flight_number	arrival_station	f_ground_time_p	f_transfer_bag_r	f_ssr_count	f_load_factor	difficulty_score	daily_rank	difficulty_class
1776	DEN	1.35	0.85	12	0.98	98.2	1	Difficult
920	LHR	0.62	0.72	15	0.95	95.5	2	Difficult
2494	MCI	1.1	0.55	2	0.88	78.4	15	Medium
5262	LNK	0.83	0.21	0	0.85	45.1	48	Easy
4792	ROA	1	0.15	1	0.91	41.7	55	Easy

Snippet of the CSV file (Doesn’t include all the columns, shown in xlsx format here just for visualization.)

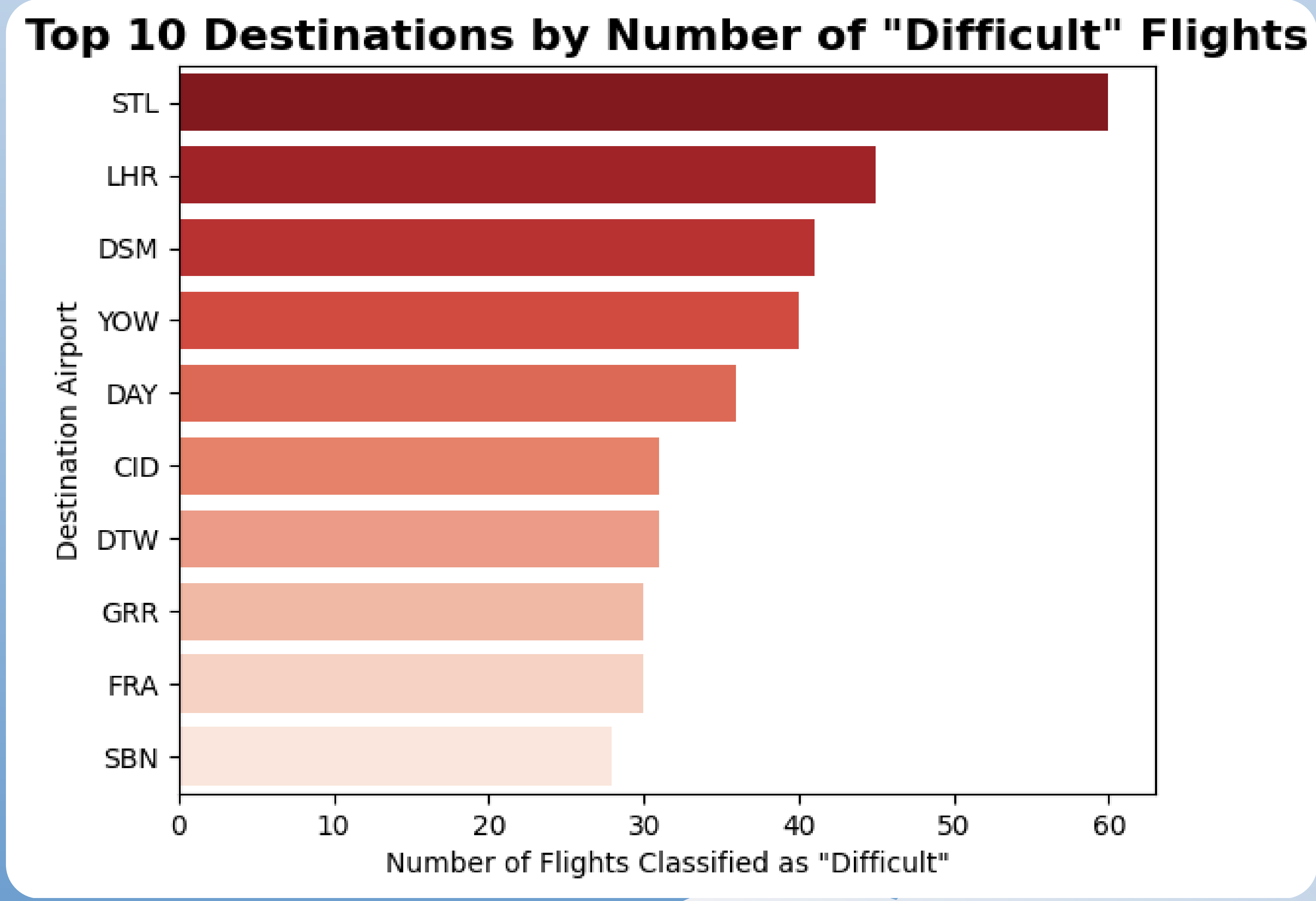
Deliverables 3

03 Post Analysis and Operational Insights:

- Summarize which destinations consistently show more difficulty.
- What are the common drivers for those flights?
- What specific actions would you recommend based on the findings for better operational efficiency?

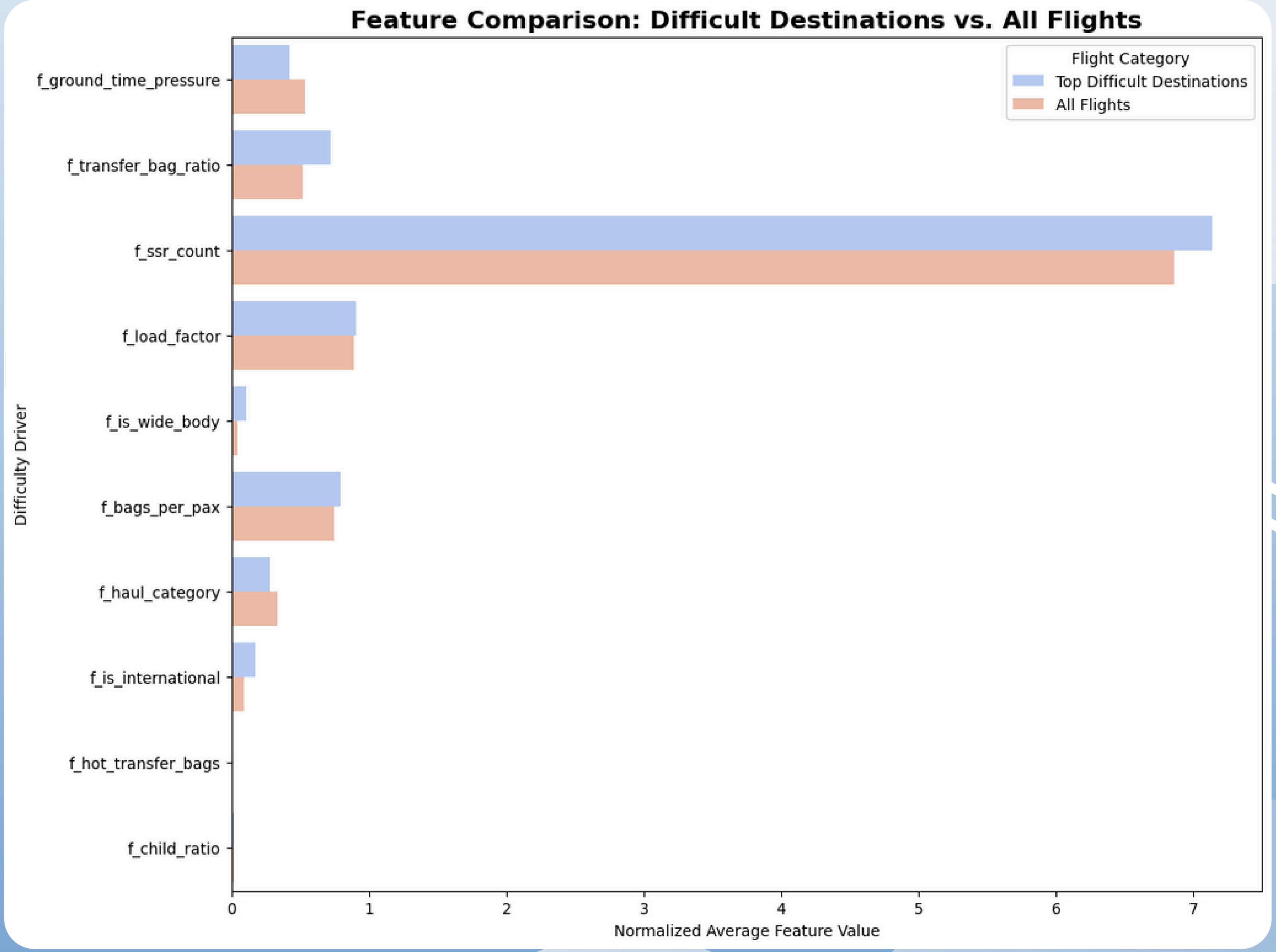
Summarize which destinations consistently show more difficulty.

The analysis identifies the top 10 destination airports that most frequently receive flights classified as 'Difficult' departing from ORD. The results clearly show that flight complexity is not random but is heavily concentrated on major domestic hubs and key international routes. Destinations like **St. Louis (STL), Des Moines (DSM), Ottowa (YOW), and London (LHR)** consistently appear at the top of the list, indicating that flights to these major airports are systematically more challenging to operate.



What are the common drivers for those flights?

- The data reveals that a "difficult" flight is a result of compounding pressures from three distinct areas: **scheduling, volume, and complexity.**
- The most challenging flights are defined by **intense ground time pressure** and the logistical burden of a **high ratio of transfer bags**. This operational strain is then amplified by sheer density-these flights are not just full (load factor), but also heavy with luggage per person (bags per pax), maximizing the physical workload.
- Finally, International, **long-haul flights** (is_international, haul_category) and those with a **high number of special service requests (ssr_count)** demand specialized handling, from customs protocols to dedicated passenger assistance.



What specific actions would you recommend based on the findings for better operational efficiency?

1. Proactive Passenger Service Staffing at the Gate

- **Data-Driven Finding:** Flights to high-volume leisure destinations (e.g., Orlando) and international hubs consistently show a high number of Special Service Requests (f_ssr_count), particularly for mobility assistance.
- **Proposed Action:**
  - Using the daily difficulty report, flag all flights with a high SSR count.
  - Proactively station additional customer service agents and wheelchair assistants at the departure gates for these flights, beginning 30 minutes before boarding. This preempts boarding delays caused by last-minute assistance needs.

2. Dynamic Monitoring of High-Pressure Turns

- **Data-Driven Finding:** A primary driver of difficulty across all routes is intense f_ground_time_pressure, where a flight's scheduled ground time is dangerously close to the minimum required.
- **Proposed Action:**
  - The daily difficulty report should create a "watchlist" of the top flights with the highest ground time pressure.
  - The airport operations manager should actively monitor the inbound arrival status of these specific aircraft. If an inbound delay occurs, they can dynamically re-assign ground crew from a lower-priority 'Easy' flight to protect the on-time performance of the 'Difficult' flight at risk.

Bonus: Consider These Factors

- **Introduce a "Weather Impact" Factor:**
  - The current model doesn't account for weather, which is a massive driver of ground operation complexity. We could integrate Terminal Aerodrome Forecast (TAF) data for ORD. A flight scheduled during a forecasted thunderstorm, high winds, or de-icing conditions is inherently more difficult, and adding a weather variable would dramatically increase the model's predictive accuracy.
- **Consider "Crew Experience" as a Variable:**
  - Not all crews are created equal. A flight staffed with a junior captain, a new gate agent, and an inexperienced ground crew will find a "Medium" difficulty flight far more challenging than a veteran team would. If crew data were available, adding a feature for the average experience level of the assigned crew would capture a critical human element of operational difficulty.
- **Factor in "Downstream Network Impact":**
  - Some flights are more critical to the airline's network than others. A delay on a flight to a major hub, which has dozens of connecting passengers heading to other destinations, has a much larger ripple effect than a delay on a flight to a small regional airport. You could create a "Network Criticality Score" for each destination, giving more weight to flights that, if delayed, would cause the most significant downstream disruption.

References and Appendices

- **Special Service Request (SSR) Operations**
  - **Citation:** WheelchairTravel.org. (2024). Special Service Request (SSR) Codes.
  - **Relevance:** Details how SSRs are direct triggers for allocating specific staff and equipment, validating them as a resource planning metric.
  - **Link:** WheelchairTravel.org Resource Page

- **Baggage Handling Systems Analysis**
  - **Citation**: Fok, K. K., & Lirn, T. C. (2018). Analyzing the efficiency of baggage handling systems at international airports. Journal of Air Transport Management, 68, 110-117.
  - **Relevance:** Provides academic backing that transfer baggage is inherently more complex and time-consuming to handle than origin baggage.
  - **Link:** ScienceDirect Article

- **Primary Data Sources**
  - **Relevance:** This collection of files (Flight Level Data.csv, Bag+Level+Data.csv, PNR+Flight+Level+Data.csv, PNR Remark Level Data.csv, Airports Data.csv) serves as the foundation for this entire analysis.

Thank you United, for giving our ideas the clearance for take-off!

UNITED