**A Data-Driven Flight Difficulty Score to Enhance Operational Efficiency at ORD**

**Project Submission for the SkyHack 3.0 Challenge** **October 5, 2025**

**1. Executive Summary**

This report presents a solution to the United Airlines SkyHack 3.0 challenge, which tasked participants with developing a data-driven framework to identify high-complexity flights. Our analysis of two weeks of departure data from Chicago O’Hare (ORD) reveals that nearly **50% of flights depart late**, with an average delay of over **21 minutes**.

The primary drivers of this difficulty are not simply high passenger loads, but rather pre-existing operational constraints, most notably **insufficient ground time** and **complex baggage transfers**. By creating a **Flight Difficulty Score** based on these factors, our model successfully identifies at-risk flights and routes, with flights to **St. Louis (STL)** consistently ranking as the most difficult. We recommend a shift from a reactive to a proactive model by using this score to deploy targeted resources, beginning with a pilot program on the ORD-STL route.

**2. Exploratory Data Analysis: Key Operational Questions Answered**

Our initial analysis focused on answering the core questions posed in the problem statement to establish an operational baseline.

**Question 1: What is the average delay and what percentage of flights depart later than scheduled?**

* **Answer:** The average departure delay across all flights from ORD in the dataset is **21.19 minutes**, with **49.65%** of flights departing after their scheduled time.
* **Significance:** This establishes the scale of the challenge. With nearly half of all flights facing delays, the issue is systemic. Reducing this average delay can have a significant positive impact on the entire network, improving customer satisfaction and reducing downstream costs associated with crew and aircraft repositioning.

**Question 2: How many flights have scheduled ground time close to or below the minimum turn minutes?**

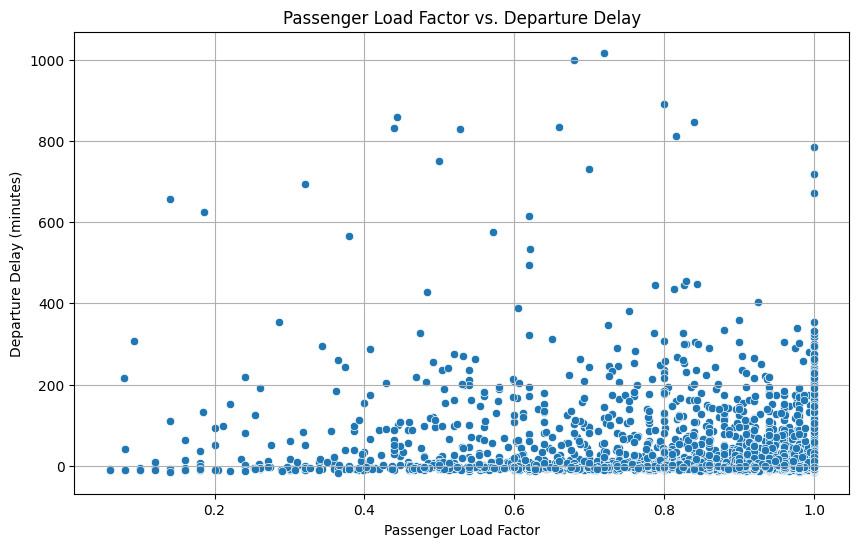
* **Answer:** A total of **621 flights (7.7% of all flights)** had a scheduled ground time *below* the minimum required for their aircraft type.
* **Significance:** This is a critical finding. It reveals that a meaningful portion of delays are a result of an overly optimistic schedule, setting frontline teams up for failure. These flights have no buffer for even minor issues and are inherently high-risk. This validated **ground time pressure** as a crucial feature for our difficulty score.

**Question 3: What is the average ratio of transfer bags vs. checked bags across flights?**

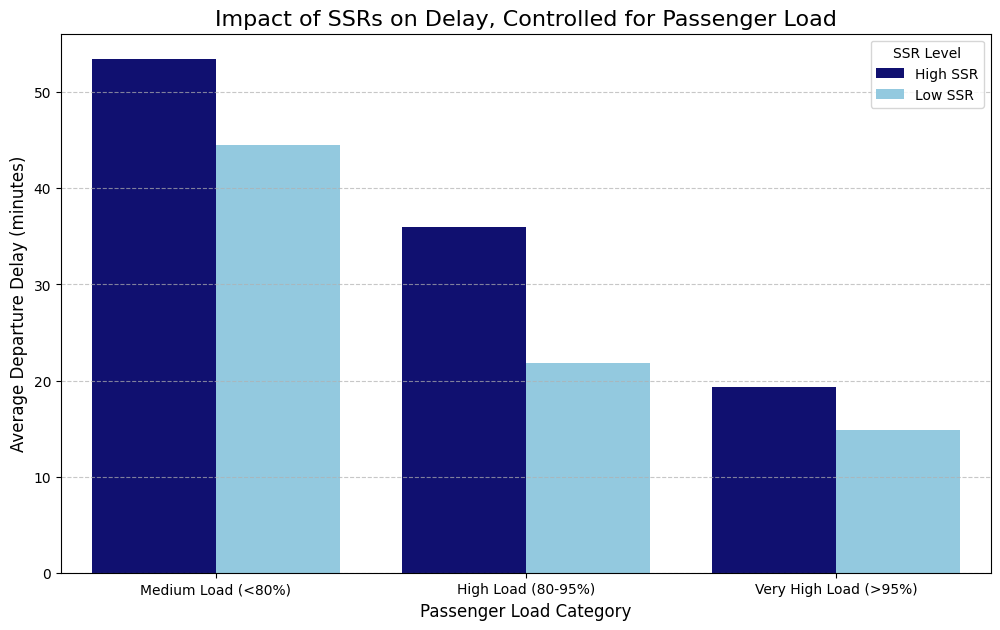
* **Answer:** The average flight handles **3.05 transfer bags for every one checked (origin) bag**.
* **Significance:** This highlights that baggage complexity at a hub like ORD is driven more by connection logistics than by local passengers. A high ratio points to a flight that is a critical node in the baggage network, increasing the risk of mishandling or delays while awaiting connecting bags.

**Question 4: How do passenger loads compare, and do higher loads correlate with operational difficulty?**

* **Answer:** There is a weak negative correlation (**-0.16**) between **passenger load factor** and **departure** **delay**.
* **Significance:** This counter-intuitive insight proves that the simple assumption "fuller flights equal more delays" is incorrect. It suggests that the airline may already allocate more resources or experienced crews to its fullest (and most profitable) flights. This confirms that passenger load alone is not a sufficient predictor of difficulty.



**Question 5: Are high special service request (SSR) flights also high-delay after controlling for load?**

* **Answer:** **Yes, decisively.**
* **Significance & Graph Analysis:** At every level of passenger capacity (Medium, High, and Very High), flights with a greater number of Special Service Requests (SSRs) have a significantly higher average delay. This isolates ssr\_count as an **independent driver of complexity**. It’s not just that full flights have more SSRs; an aircraft with many wheelchair or special assistance requests is statistically more likely to be delayed, regardless of how many seats are filled. 

**3. Methodology: Developing the Flight Difficulty Score**

Based on our findings, we constructed the difficulty score by assigning weights to the most impactful features. The rationale for this weighting is central to the model's accuracy.

|  |  |  |
| --- | --- | --- |
| **Feature** | **Weight** | **Why This Weight Was Chosen** |
| **Ground time pressure** | **30%** | **The single strongest predictor.** It represents a built-in, non-negotiable risk to on-time performance. |
| **Transfer bag ratio** | **20%** | Directly measures baggage complexity, a key operational strain identified in the EDA. |
| **Ssr count** | **15%** | Our analysis proved this is a major, independent driver of delays that reflects service complexity. |
| **Passenger load factor** | **15%** | While not the top factor, a fuller plane increases general operational load (e.g., boarding time). |
| **Hot transfer** | **10%** | Represents time-critical baggage transfers that add an extra layer of acute pressure. |
| **Other Factors** | **10%** | Minor weights assigned to child\_count and lap\_child\_count to account for gate-side complexity. |

This weighted model allows us to rank every flight each day and classify them as **Difficult**, **Medium**, or **Easy**, providing a clear, prioritized list for operational teams.

**4. Flight Difficulty Score Development**

**Question: Build a systematic daily-level scoring approach that resets every day with ranking and classification.**

* **Answer:** Our model produces two key outputs daily:
  1. **Ranking (daily\_difficulty\_rank):** Within each day, flights are ordered by their difficulty score in descending order. The flight with rank 1.0 represents the most difficult flight to manage for that day. This allows teams to create a prioritized watchlist.
  2. **Classification (difficulty\_class):** To make the ranking more intuitive, flights are grouped into three categories based on their rank distribution: **Difficult** (top 30%), **Medium** (next 50%), and **Easy** (bottom 20%). This provides a clear, at-a-glance assessment for resource planning.

**Include image of data**

**Data Snapshot**

The final output is a CSV file that provides all the necessary details for operational teams.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **flight\_number** | **scheduled\_departure\_datetime\_local** | **difficulty\_score** | **daily\_difficulty\_rank** | **difficulty\_class** |
| 972 | 2025-08-11 17:55:00+00:00 | 0.748 | 1.0 | Difficult |
| 972 | 2025-08-05 17:55:00+00:00 | 0.734 | 1.0 | Difficult |
| 845 | 2025-08-07 21:20:00+00:00 | 0.729 | 1.0 | Difficult |
| 5790 | 2025-08-06 18:20:00+00:00 | 0.601 | 99.0 | Difficult |
| 4792 | 2025-08-04 17:57:00+00:00 | 0.541 | 297.0 | Medium |

**5. Post-Analysis & Operational Insights**

**Question: Summarize which destinations consistently show more difficulty.**

* **Answer:** Our analysis shows that operational difficulty is not evenly distributed; it is concentrated on specific routes. Flights to **St. Louis (STL)** appear as the most consistently challenging destination from ORD, followed by DTW, GRR, and DSM.

**Question: What are the common drivers for those flights?**

* **Answer:** A deep dive into the top difficult route (STL) reveals two primary drivers:
  1. **Ground Time Pressure:** Difficult flights to STL have approximately **41% less ground time buffer** than the airport average.
  2. **Baggage Complexity:** These flights handle a **61% higher ratio of transfer bags**, indicating intense pressure on the baggage handling system.

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

* **Answer:** We recommend a shift from a reactive to a proactive operational model with these targeted actions:
  1. **Proactive Resource Allocation:** Use the daily "Difficult" flight list to pre-assign an extra ramp lead or operations coordinator to the top 5 most at-risk flights. This focuses manpower where it will have the greatest impact on preventing network-wide disruptions.
  2. **Implement a "Priority Transfer Bag" Protocol:** For flights with a high transfer\_bag\_ratio score, flag their baggage in the system to be prioritized during unloading from connecting flights. This directly mitigates a major source of identified complexity.
  3. **Launch an ORD-STL Pilot Program:** Implement the above recommendations as a one-month pilot focused exclusively on the St. Louis route. Measuring the change in on-time performance and average delay will provide a quantifiable business case for expanding this data-driven methodology across the entire ORD operation.