

Analysis of Flight Delays and Cancellations

TEAM 3

Team members:

Jialin Liu

Quan Nguyen

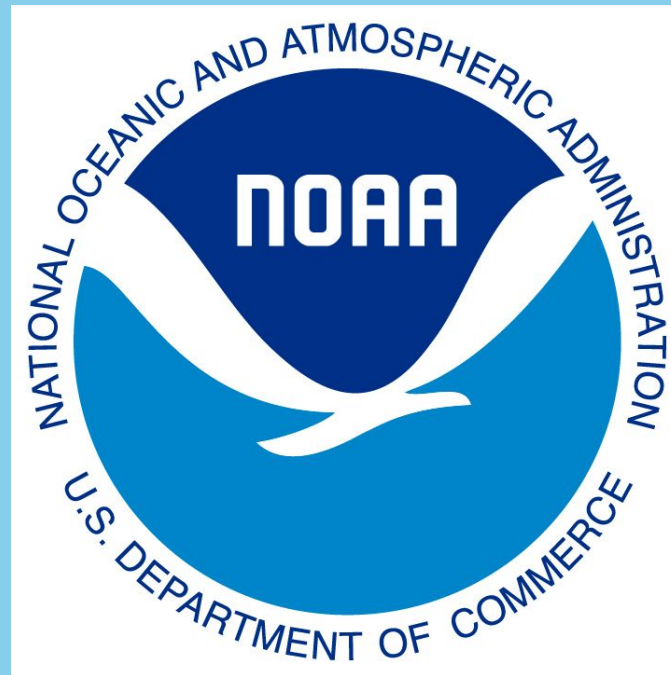
Varun Kaza



Project Proposal & Motivation

- Delays and cancellations in flights disrupt schedules and incur significant costs. This analysis aims to understand these issues and provide actionable insights.
- The weather plays a an important role in commercial travel.

Data Source



Types of Analysis



Exploring Flight Patterns and Insights into Weather Impacts

Data Range: 01/01/2022-07/31/2024

Tracking the count of Arrival and Departure times

Tracking Air time and Distance

Tracking Cancellation and Delays

Cancellation and Delay Reasons



Specific Insights and Deeper Weather Analysis

Data Range: 01/01/2022-07/01/2024

10 busiest domestic airports

Tracking Summary statistics by location and time

Tracking the occurrence of extreme weather conditions
such as Thunder, Fog, and Snow

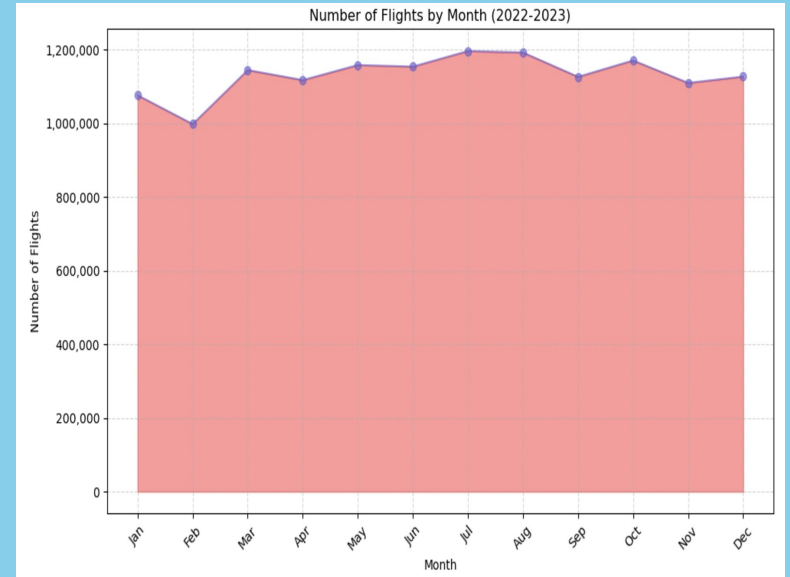
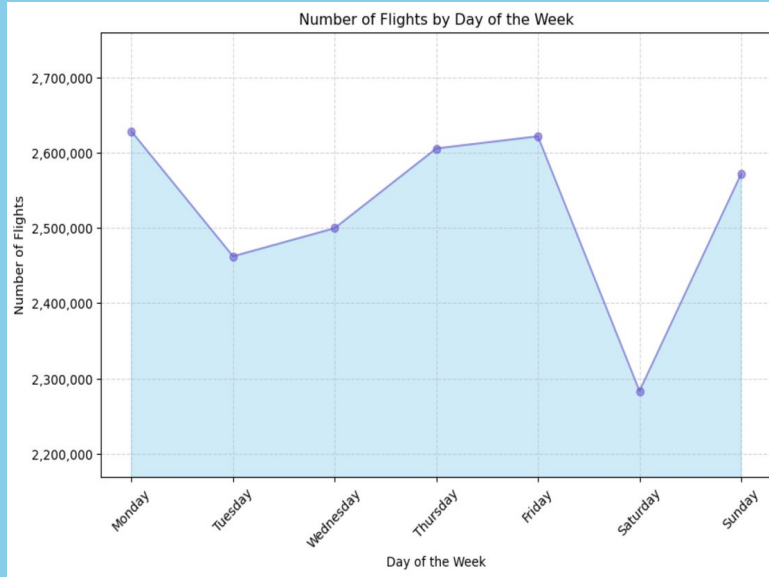
Exploring Weather Data

	Location	Thunder_Count	Avg_Precipitation	Total_Precipitation
0	Orlando International	258	0.386899	99.82
1	Hartsfield-Jackson Atlanta International	159	0.486604	77.37
2	Dallas/Fort Worth International	150	0.521000	78.15
3	Denver International	140	0.133000	18.62
4	Charlotte Douglas International	125	0.432560	54.07
5	Chicago O'Hare International	109	0.380642	41.49
6	Ronald Reagan Washington National	104	0.359327	37.37
7	Newark Liberty International	96	0.401458	38.54
8	LaGuardia	91	0.554505	50.46
9	Logan International	56	0.446786	25.02

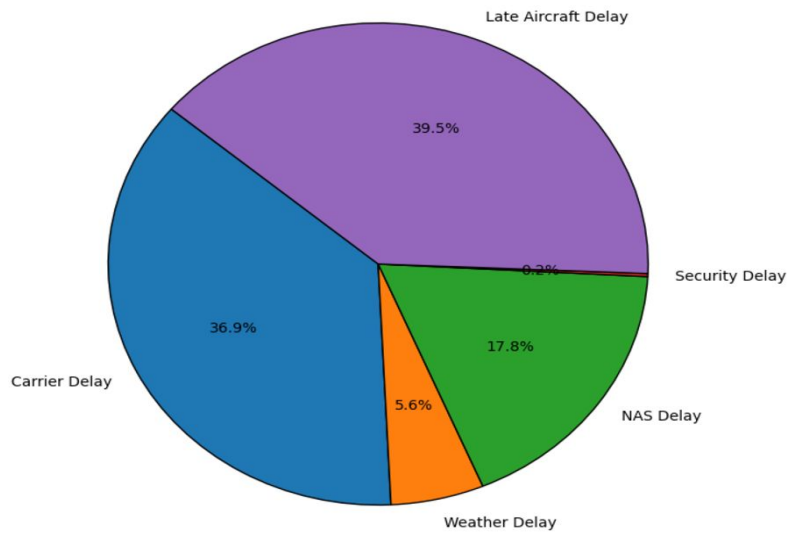
Challenge of Integrating Weather Data

- Integrating weather data posed challenges due to:
 - 1. Data granularity mismatches.
 - 2. Missing timestamps for specific flight records.
 - 3. Different formats in datasets.

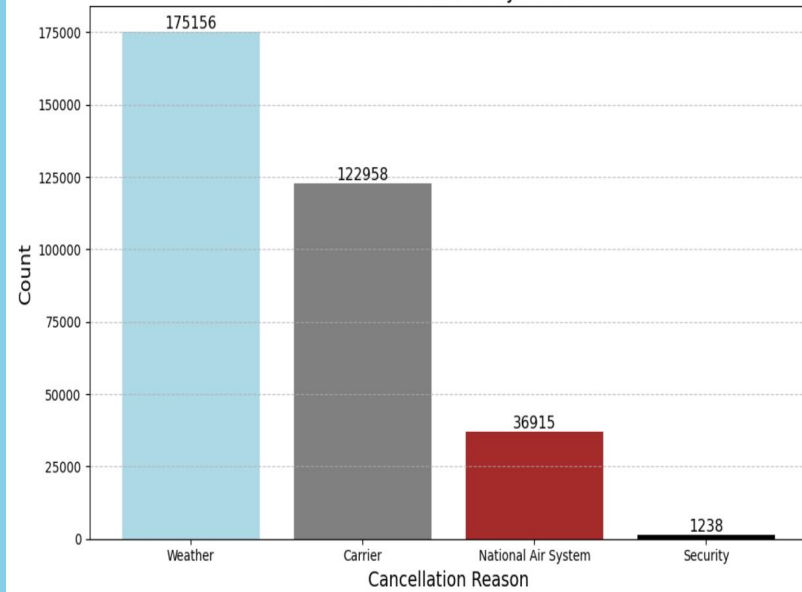
Exploring Flight Patterns

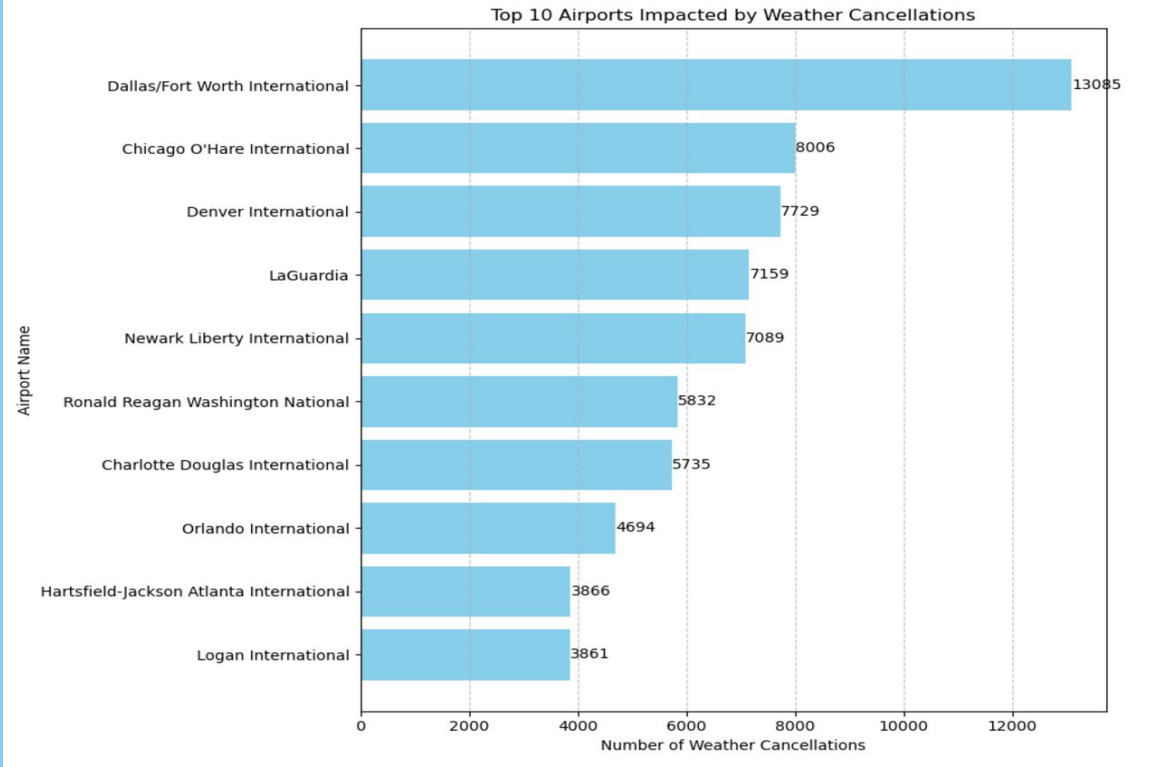


Proportion of Flight Delays by Cause

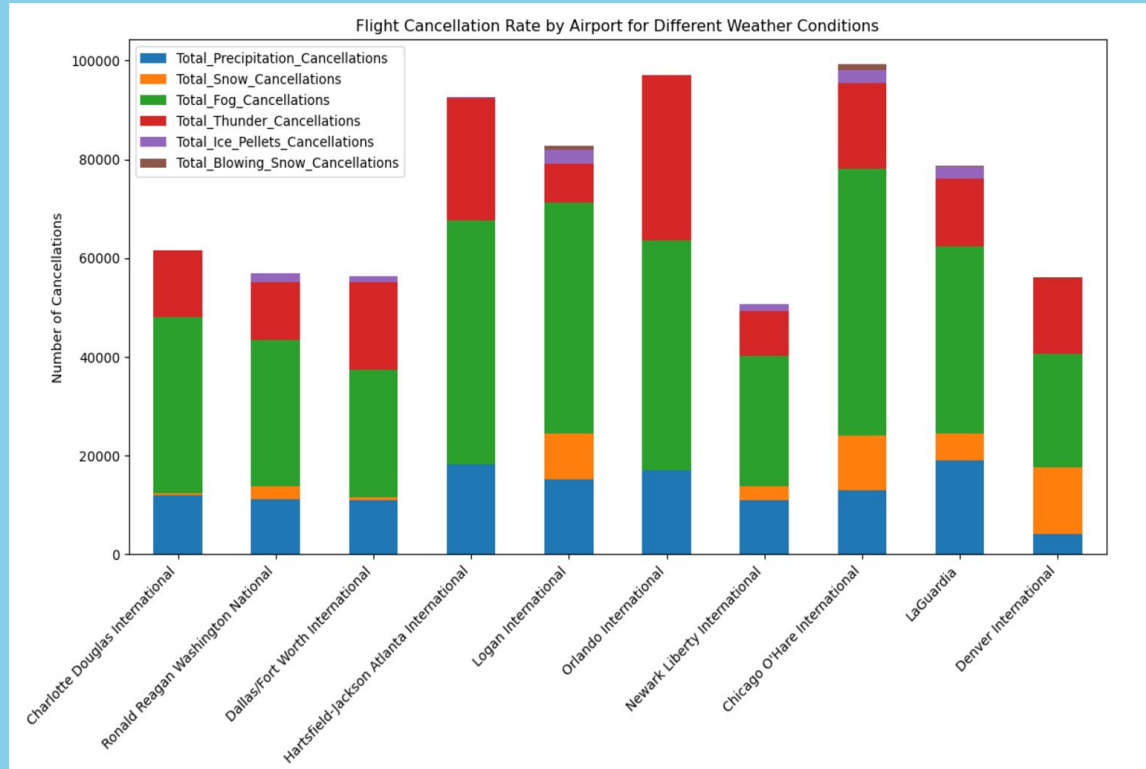


Cancellation Counts by Reason

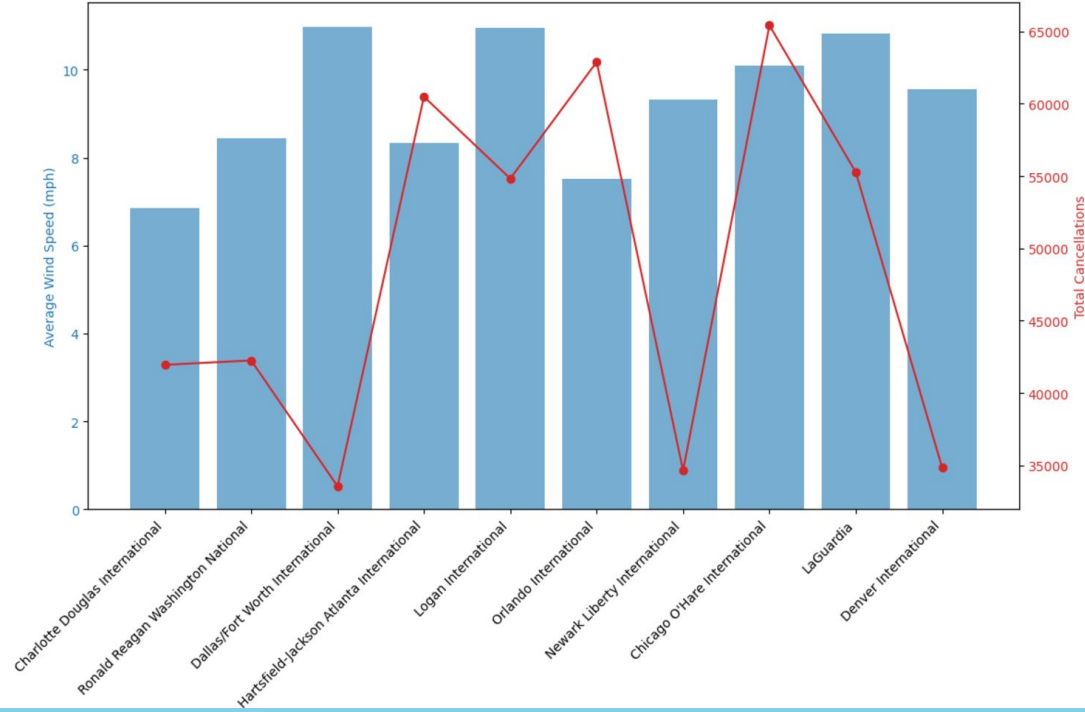




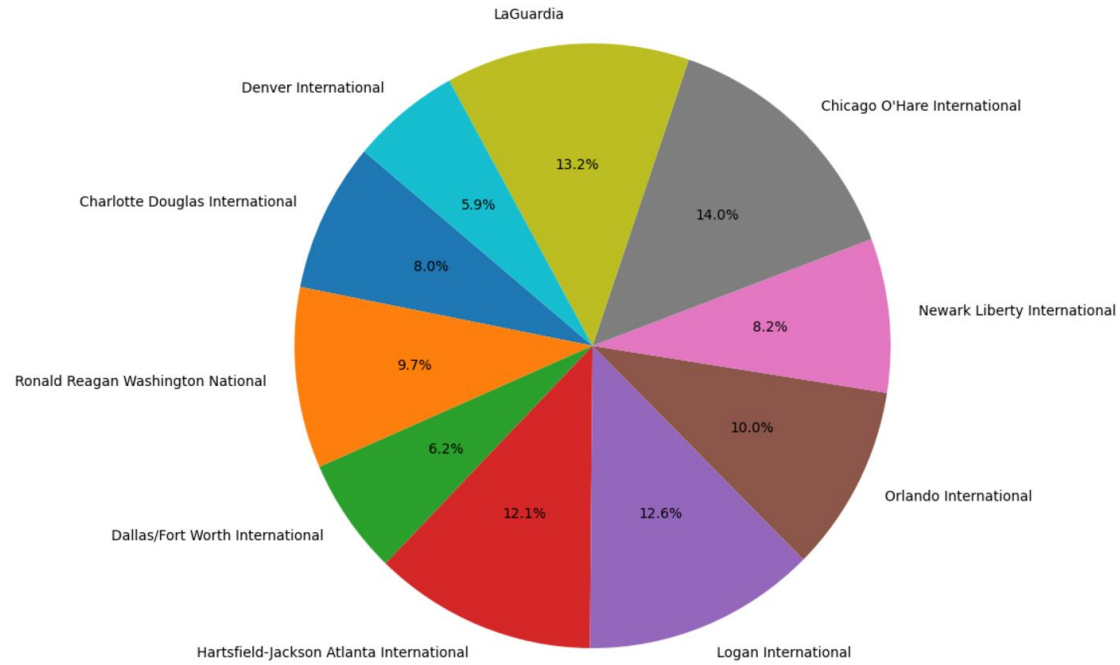
Deeper Weather Analysis



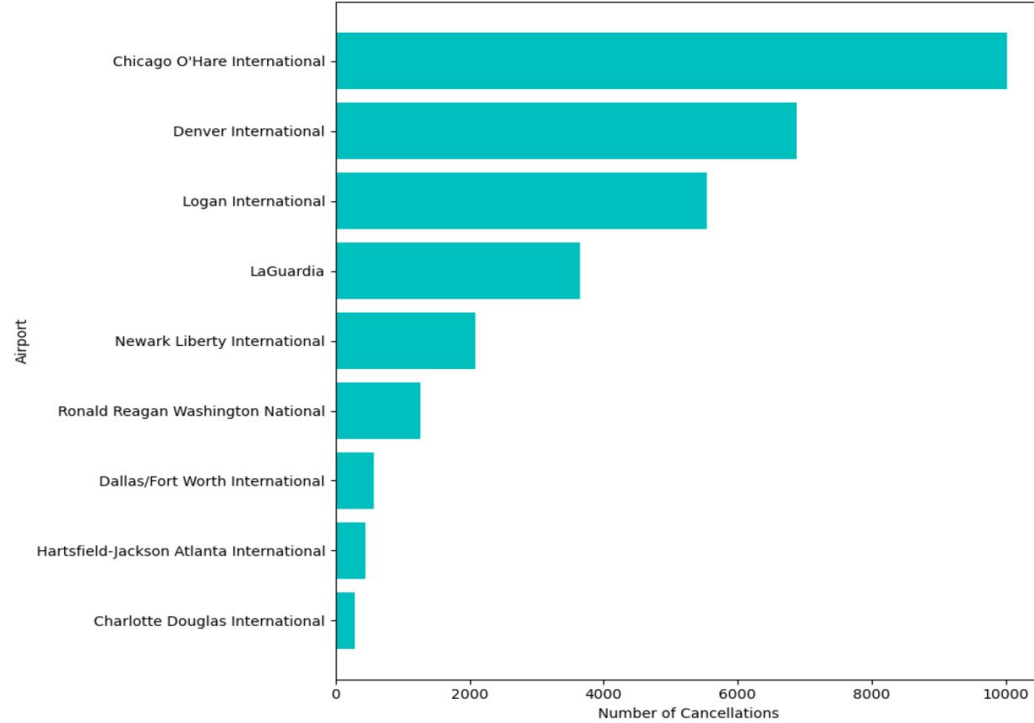
Impact of Average Wind Speed on Flight Cancellations by Airport

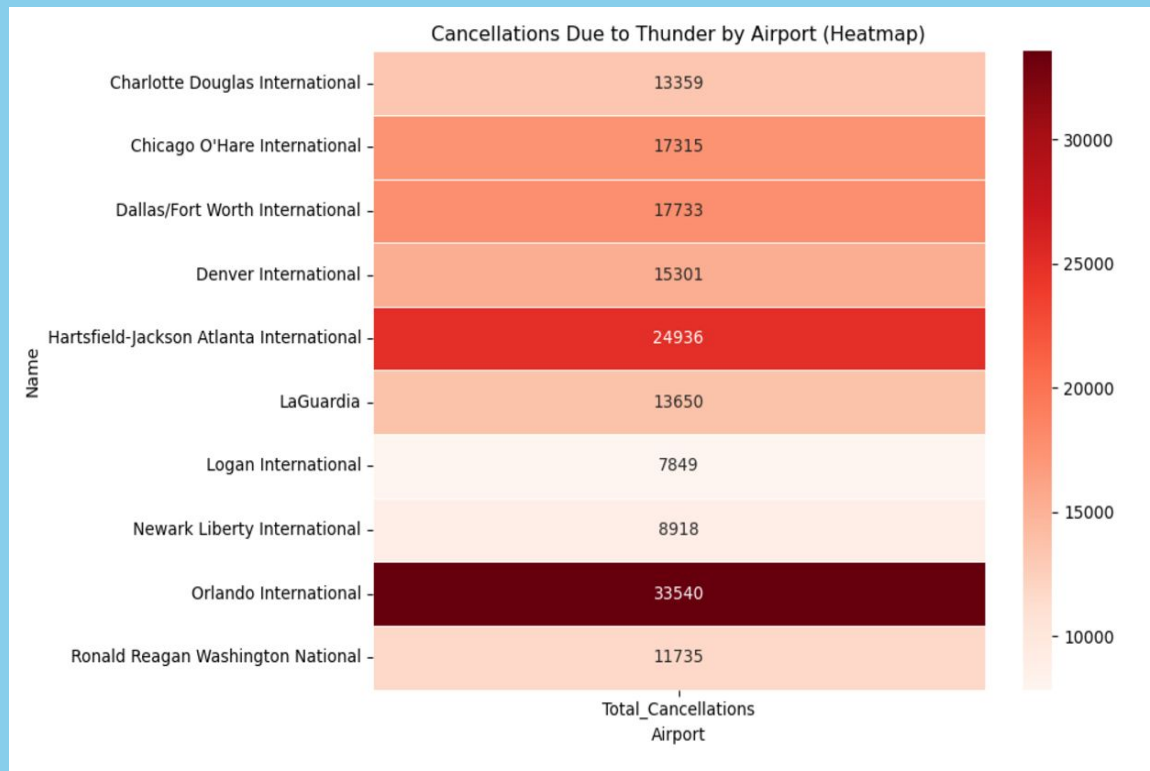


Impact of Precipitation on Flight Cancellations by Airport

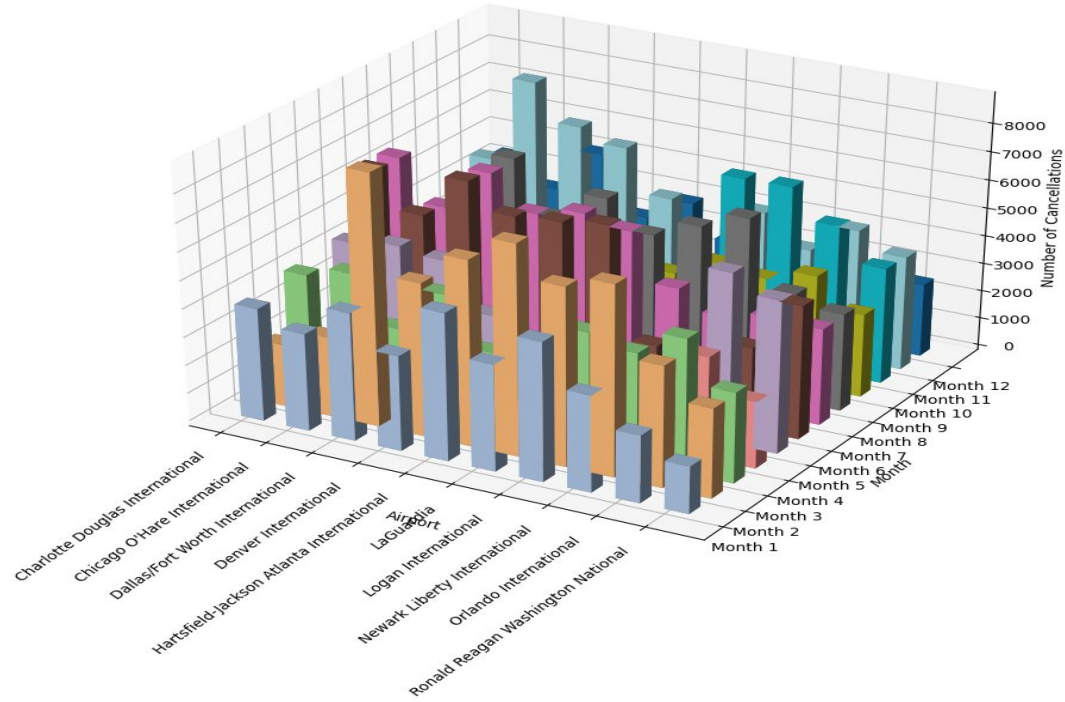


Impact of Snow on Flight Cancellations by Airport (Sorted by Number of Cancellations)





Flight Cancellations Due to Severe Weather by Airport and Month (3D)



Machine Learning

- Purpose: Predicting Flight Cancellations
- Random Forest Classifier
- Feature Engineering:
 - **Weather Severity Score:** simple score combining precipitation, snow, and wind speed
 - **Adverse Weather Flag:** binary feature that flags days with weather events like thunderstorms and fog.
 - **Departure Hour**

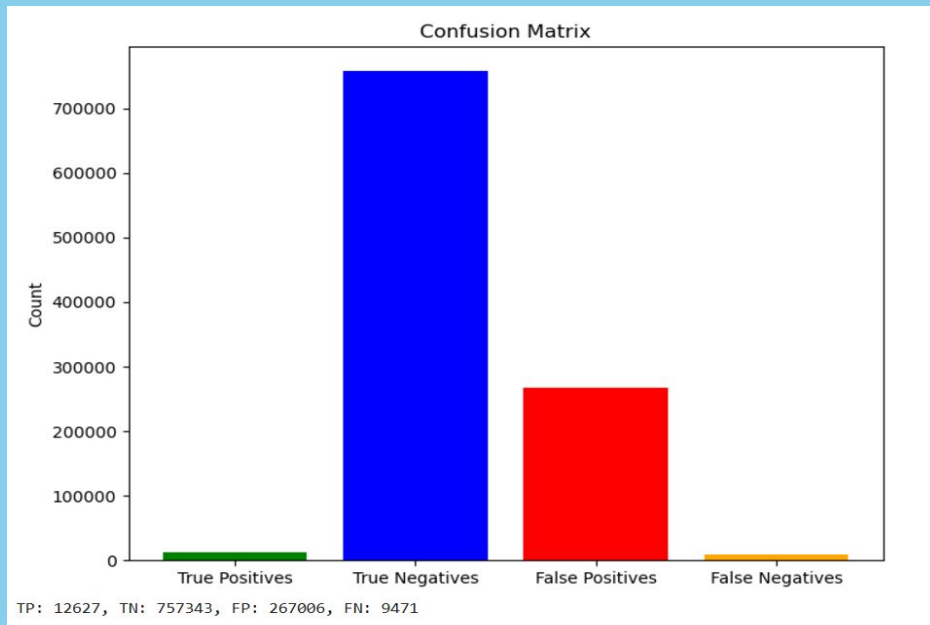
Model Performance

Baseline Model:

- AUC: 0.5
- Precision: 0.97

Improved Model:

- AUC (Area Under Curve): 0.704
- Precision: 0.968
- Recall: 0.722
- F1 Score: 0.820



Recommendations

- Airlines can use the predictions to proactively manage resources
- Strong influence of weather conditions, we recommend leveraging real-time weather data to feed into the model

Summary and Conclusion

- Key takeaways:
- 1. Weather is Not the most important factor in delays, however, it does for cancellations.
- 2. Machine learning models provide actionable insights.
- 3. Recommendations aim to minimize disruptions.

Thank You!