Flight Delay Prediction and Analysis

Sai C., Maryann O., Matt J., Jeremiah B., Ricky M.

Flight Delay Prediction And Analysis

Team Members	Emails	Github
Sai C.	schit7@uic.edu	Sschittala
Maryann O.	aolug3@uic.edu	MaryannO45
Matt J.	mwilkj2@uid.edu	Mattwilk1017
Jeremiah B.	jbenj2@uic.edu	Juebenjamin
Ricky M.	fmass3@uic.edu	Fmassa1

https://github.com/juebenjamin/flight-delay-analysis

TABLE OF CONTENTS

How Data Was Analyzed

PROBLEM

STEP 1

INVESTIGATION AND
CHECKING TRENDS

STEP 3

STEP 2

DATA APPROACH

STEP 4

CHALLENGES
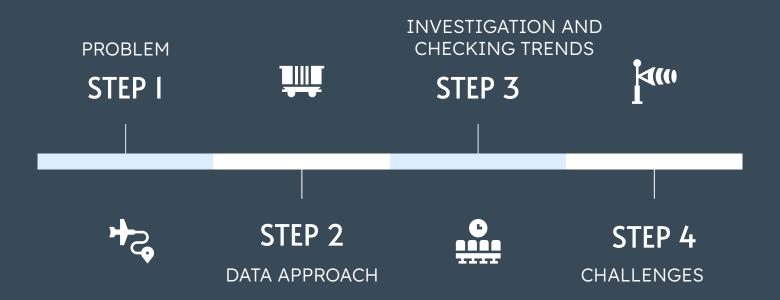
Let's Recap On The Big Idea

Problem - Unpredictable flight delays disrupt travel plans,which increase operational costs, and overall reduce airport efficiency.

2.76%

Why It Matters

In 2024, US had the third-highest cancellation rate among the ten countries considered
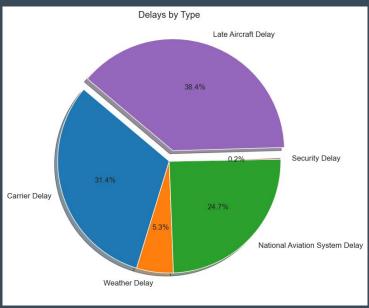
50%

Why Choose

Flight prediction enhance travel efficiency, makes flying more efficient, and use data for smarter aviation operations.
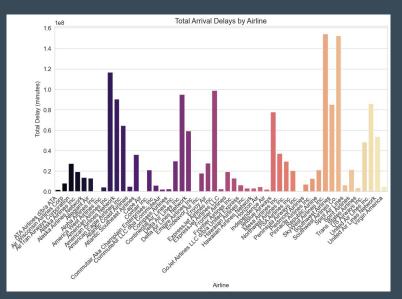
75%

Hypothesis

If airports used a predictive learning model to identify potential delay risks they would be able to better respond, leading to a significant reduction in overall flight delays.
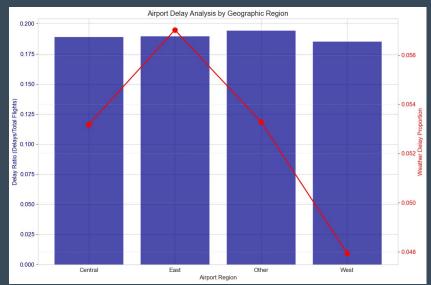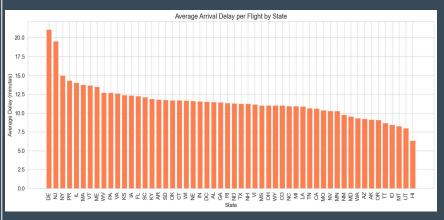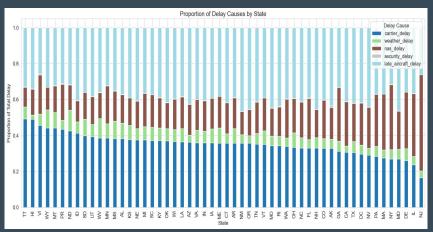
Findings- EDA

Average Delay per Flight by Year (2003–2024)

Delays by Type

Findings- EDA

Findings

Average Arrival Delay per Flight by State

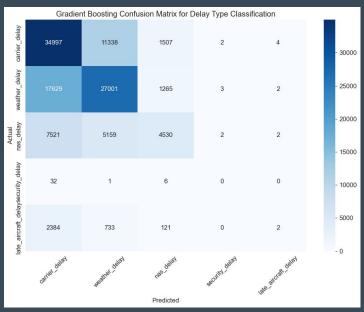

Proportion of Delay Causes by State

ML Models

- Gradient Boosting Classifier: Predicts the primary cause of delays (carrier, weather, NAS), allowing airlines to apply the right operational fixes based on the predicted issue

- PCA Clustering: Grouped States by delay behavior, revealing regional patterns and outliers like New Jersey and Alaska. This can be useful in developing specialized delay reduction strategies tailored to these areas. For example:
    - New Jersey can invest in airspace traffic management systems to reduce congestion bottlenecks
    - Alaska can improve cold- weather operations by improving deicing efficiency. Increased weather forecasting capabilities would help improve this process as well
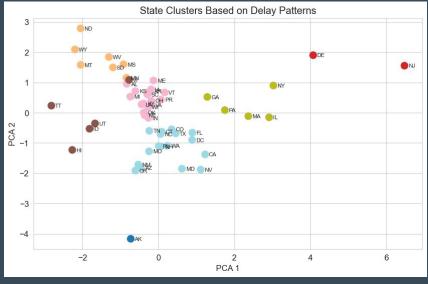
- **Summary of Impactful Applications Utilizing Machine Learning Models**
- As aviation is a very unpredictable industry that is built on top of redundancy to ensure safety and efficiency, using these models can help allocate resources to the appropriate areas before disruptions occur, focus investments where structural delays risks are the highest, and receive better real - time updates to ensure customer satisfaction and keep everyone informed

ML Models Visualizations

Gradient Boosting Confusion Matrix for Delay Type Classification

State Clusters Based on Delay Patterns

Summary & Findings

WHAT WORKED?

- Delay Spikes
- Cause of Delay

WHAT DID NOT WORK?

- Weather data
- Data Leaks

QUESTIONS AND COMMENTS?

Thanks!