CS418 Group #18

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

Let’s Talk 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

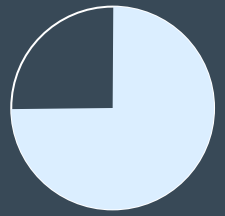
In 2024, US had the third-highest cancellation rate among the ten countries considered

50%

Why Choose

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

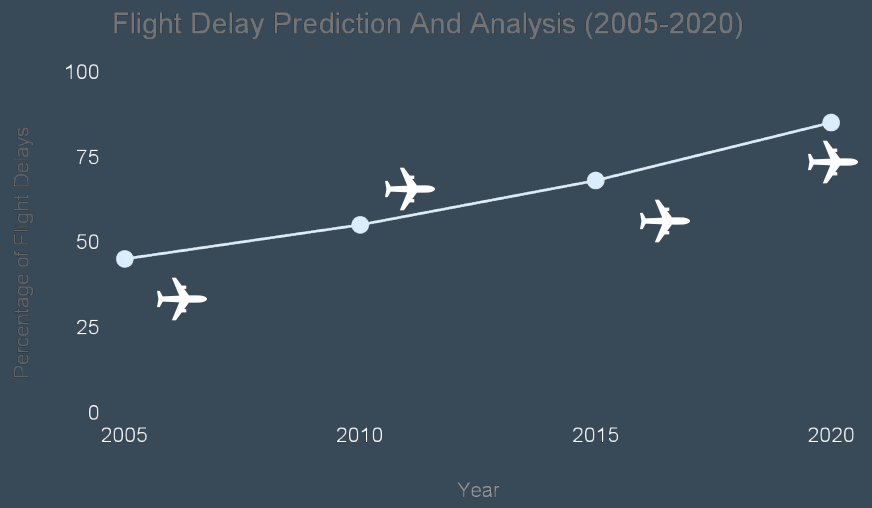
75%



Hypothesis

According to the FAA, weather conditions are the strongest predictor of flight delays, over 75% of all significant flight delays; weather data, like visibility, wind speed, and thunderstorms impact flight Delay.

Data Resources



https://www.transtats.bts.gov/ontime/Cancellation.aspx

Sources

BTS On-Time Performance: Historical flight schedules and delay records.

OpenSky Network / Flightradar24: Real-time flight tracking and status updates.

NOAA / OpenWeatherMap: Weather data (temperature, precipitation, wind) for departure/arrival airports.

Data Size & Types:

- **Historical Records**: Potentially millions of flight entries over several years.
- **Features**: Flight times, airport codes, delay durations, weather indicators.
- **Formats**: CSV files, JSON (API responses), time-series data.

Solutions

We plan to approach this problem through four major factors listed below.

- **Exploratory Data Analysis (EDA)**: Identify key factors influencing flight delays (e.g., weather, time of day, airline).
- **Feature Engineering:** Construct meaningful variables such as airport congestion levels and historical delay trends.
- **Machine Learning Models:** Implement multiple models to predict flight delays. Some examples include decision trees, logistic regression, etc.
- **Assess Airport Patterns:** Identify specific airport functions that add to delays such as capacity issues, taxi times, airport structure, etc

Scope and Next Steps

Short Term (Until Progress Report):

- Complete data collection and cleaning
- Conduct EDA to identify patterns
- Create visualizations for any major trends that show up in data collection
- Implement at least one predictive model

Long Term (Final deliverable):

- Optimize feature selection for model improvement.
- Develop a very accurate model for prediction.
- Develop an interactive system/model that could predict delays based on any given flight
- Create a report highlighting recommendations airlines can take to mitigate these issues

Thanks!