

# Global Airline Operations

## Exploratory Data Analysis

---

Name: Youssef

### Problem Statement:

The airline industry is a vital component of the global economy. This project investigates patterns and factors that influence flight delays across major global airline carriers. The objective is to identify key causes of delays, seasonal trends, and how these factors vary across airlines and airports. Using this dataset, we aim to explore correlations between delay types and total delay times, and investigate if certain delay causes can be mitigated by proactive measures. Insights from this analysis can help improve airline operational efficiency and customer satisfaction.

### Business Impact:

Understanding delay patterns enables airline companies and airport authorities to optimize scheduling, reduce costs, and enhance passenger experience. By identifying the primary contributors to delays, resources can be allocated more effectively. Additionally, improved on-time performance could positively impact airline reputation, leading to increased customer retention and profit margins.

### General Dataset Information:

File Name: airline\_delay\_causes.csv

Description: Flight delay data from 2003 to 2023 including causes such as weather, security, carrier, and NAS delays.

Dataset Details: ~1,200,000 Rows & 12 Columns

Size: 48MB

Source: [Kaggle – Global Airline Operations Dataset](#)

### Target Features:

- Airline - Carrier code
- Date - Date of flight
- CarrierDelay - Delay caused by airline

- WeatherDelay - Delay caused by weather conditions
- NASDelay - Delay due to National Aviation System
- SecurityDelay - Delay due to security protocols
- LateAircraftDelay - Delay from late-arriving aircraft
- ArrDelay - Total arrival delay
- Month, Year - Derived time features for trend analysis

### **Analysis #1 - Total Records**

To provide scope, the dataset contains over 1.2 million records of flight operations globally. This wide span allows us to identify patterns across time, locations, and airlines.

### **Analysis #2 - Records by Year**

Grouping by Year shows clear seasonality and disruptions in flight schedules, especially around the COVID-19 period. Delays sharply reduced in 2020-2021 due to the pandemic and have increased again with post-pandemic recovery.

### **Analysis #3 - Delay Cause Breakdown**

Analyzing each delay category individually reveals that Late Aircraft Delay and NAS Delay are consistently high contributors. Weather delays vary seasonally while Security Delays are low and infrequent.

### **Analysis #4 - Basic Statistics**

The mean delay is around 15 minutes with a long tail of extreme values. Maximum delays surpass 1000 minutes. Outlier removal was tested but retained due to realistic worst-case events (e.g., hurricanes, security threats).

### **Analysis #5 - Delay Trends by Airline**

Grouped by Airline, the data shows major differences in delay types. Some carriers consistently have higher LateAircraftDelay, suggesting internal inefficiencies. Low-cost carriers experience more delays related to turnaround times.

### **Analysis #6 - Delay Distribution by Airport**

Airports were analyzed to find hubs with chronic delay problems. Certain major hubs like ATL and ORD show higher delay concentrations. This is expected due to high volume and complex airspace.

### **Analysis #7 - Monthly Trends**

Using Month as a derived feature, it was found that July and December exhibit higher average delays. These months coincide with holiday and peak travel periods. Airlines may benefit from extra resources in these months.

**Conclusion:**

This EDA provided key insights into the patterns and factors contributing to global airline delays. Late Aircraft Delay and NAS Delay are primary contributors, with seasonal and airport-specific trends clearly visible. The findings validate some assumptions while also highlighting the importance of contextual data such as weather and operational volume. These insights can support further modeling and decision-making in airline operations and logistics.