**Predicting and Analyzing U.S. Flight Delays: Identifying Key Factors and Improving Forecast Accuracy**

# Proposal

Flight delays are one of the most serious challenges in the US aviation industry, causing huge inconvenience and coupled with losses to airline companies. This work examines the key patterns, causes, and trends in US flights taken from January 2017 to July 2022 that have an impact on flight punctuality. This will help airlines understand the root cause of the delay problem and thus enable them to optimize their flight schedule further for better passenger satisfaction.

The dataset from Kaggle contains flight details, including the departure and arrival time, airline operating, airport from which and where it landed, duration of delay, and finally consolidated weather data. This dataset has over millions of values, thereby providing an adequate overview of flight operations and further delays in the US over a considerable period.

**Objectives:**

1. **Causes of Delays:** Analyze different causes of flight delays, including but not limited to weather conditions, airline performance, and airport performance, seasonal causes, among others, for actionable insights helpful in mitigating delays.

2. **Trend Analysis**: Analyze the trends of flight delays with respect to time. This will include high-traffic time, holidays, and seasonal impacts.

3. **Airline & Airport Performance:** Ascertain how different airlines and different airports perform in causing or being caused by delays and, later, assess their performances against these measures.

4. **Predictive Modeling:** Utilize the insights to construct machine learning models that can predict, using airline data, flight delays resulting from weather, time of day, and other factors.

This analysis will offer valuable insights to airlines, passengers, and policymakers, leading to improved operational efficiency, better planning, and enhanced customer satisfaction in the aviation sector.

# SMART questions –

1. "Can we predict the total delay time for flights in the U.S. with the highest accuracy using historical data (2017–2022), based on factors such as airline efficiency, airport traffic, weather conditions, and National Aviation System, to provide valuable insights for airlines, airports, and regulatory bodies to optimize scheduling and reduce delays for future flight operations?"
2. "How can we leverage real-time data and machine learning to accurately predict flight delays in the U.S. based on chronicled data, considering factors like weather conditions, air traffic, intervene of National Aviation System, and airline operations, to enhance scheduling and minimize disruptions across the aviation network?"

**Dataset link -**

# <https://www.kaggle.com/datasets/jawadkhattak/us-flight-delay-from-january-2017-july-2022/data>

# GitHub repository -

<https://github.com/HemaG39635070/JamesRiverhouse>