

CS411 Project Proposal - Team 057

Title: SkyTrends

Project Summary: SkyTrends includes a Flight Data Dashboard that aims to provide a comprehensive and interactive platform for users to explore, analyse, and understand key performance indicators related to air travel. Leveraging a rich dataset covering various attributes like flight delays, cancellations, airline efficiency, and more, the dashboard aims to offer valuable insights for travellers, airlines, and policymakers alike.

Description: Air travel involves multiple stakeholders and a myriad of factors contribute to the overall experience. Delays, cancellations, and inefficiencies not only affect passengers but also have financial repercussions for airlines. While there are platforms offering flight status and basic statistics, there's a gap in providing in-depth analytics and predictive insights. Our application aims to fill this gap by providing detailed metrics and trends using an intuitive and interactive dashboard.

Usefulness. While websites like FlightStats and FlightAware provide real-time flight tracking, they lack comprehensive analytical features. Our dashboard stands out by offering data analytics with the ability to customise and filter results. This could be particularly useful for frequent fliers, airline operational analysts, and airport management.

It allows users to access historical flight delay information for specific flights. This information can be valuable for travellers who want to understand the on-time performance of airlines and specific flights before making travel plans.

Realness: The data is sourced from [Kaggle](#), a reputable platform for public datasets. This specific dataset focuses on flight performance and includes a wealth of attributes related to airlines, flight delays, cancellations, and other operational factors.

Attributes and Information:

Temporal Data: YEAR, MONTH, DAY, DAY_OF_WEEK

Airline and Flight Info: AIRLINE, FLIGHT_NUMBER, TAIL_NUMBER

Airport Info: ORIGIN_AIRPORT, DESTINATION_AIRPORT

Schedule and Actual Timing: SCHEDULED_DEPARTURE, DEPARTURE_TIME, SCHEDULED_ARRIVAL, ARRIVAL_TIME

Delay Metrics: DEPARTURE_DELAY, ARRIVAL_DELAY, AIR_SYSTEM_DELAY, SECURITY_DELAY, AIRLINE_DELAY, LATE_AIRCRAFT_DELAY, WEATHER_DELAY

Flight Operation: TAXI_OUT, WHEELS_OFF, ELAPSED_TIME, AIR_TIME, DISTANCE, WHEELS_ON, TAXI_IN

Status Flags: DIVERTED, CANCELLED

Cancellation: CANCELLATION_REASON

The database will store all these attributes to allow for a multifaceted analysis of flight performance. It will include both numerical and categorical data types, allowing for complex SQL queries for detailed insights.

Users can perform the following actions:

Basic Features:

- Search Flights: By airline, date, or airport.
- View KPIs: Such as average delays, cancellation reasons, etc.
- Time Series Analysis: Graphs showing trends over time.

Advanced Features:

- Custom SQL Queries: Users can run their own SQL queries for customised insights.
- Predictive Analytics: Predict the likelihood of a flight delay based on various factors.

Functionality:

The website's main page consists of many buttons on the Navigation Bar shown below, after clicking on them the main area of the page gets refreshed and the corresponding input area is shown.

In the new loaded area, the corresponding input box shown for inserting will provide the input box of flight number, departure, arrival, estimated arrival time, real arrival time etc. Some of these are required, some are not. For search-related functions, after filling up the option and having it submitted, the result then gets displayed under the input box in table form with a certain preference of order.

Creative Component: Predictive Delay Analysis

To enhance our application's functionality, we plan to implement a machine-learning model using past delay data. This model will predict future flight delays based on historical patterns, offering users valuable insights for their travel planning. This simplified approach aligns with our course project's scope while adding a predictive dimension to our Flight Data Dashboard. We will first preprocess and train a machine learning model on Kaggle's past delay dataset. Then, we'll integrate this trained model into our application to provide users with predictions based on historical delay patterns.

Interface:

The interface consists of a header with five tabs: "Delays", "Cancellation", "Air System Issues", "Elapsed time by Distance", and "Custom queries". Below the tabs are four dropdown menus. A table titled "Delays" is shown with columns "AIRLINE", "MONTH", and "AVG DELAY". To the right of the table is a large text area for a custom query, with a "SUBMIT" button below it. Handwritten annotations include: "Tabs to pre defined tables" pointing to the tabs; "Dropdown to apply filters" pointing to the dropdown menus; "Table will change as user clicks on Tab or custom query" pointing to the table; and "Place to write custom query" pointing to the text area.

Tabs to pre defined tables

Dropdown to apply filters

Delays Cancellation Air System Issues Elapsed time by Distance Custom queries

AIRLINE MONTH AVG DELAY

SUBMIT

Table will change as user clicks on Tab or custom query

Place to write custom query

Work distribution:

Ra26: Backend API to run SQL queries, update and fetch from the database.

vporwal3: SQL queries, Setting up Cloud Config and environment, implementing machine learning models for predictive analytics

Zihao10: SQL queries, implement machine learning models for predictive analytics

Heejun2: SQL queries, Front-end functionalities like search, KPI display, and time-series graphs.