Project Description

Tres Amigos
(Mark Strong-Shinozaki, Steven Subianto, Yuuki Matsunari)

Problem Statement

Transportation is an important aspect of our lives, especially with air transportation. A country can have at least tens of flights, some may even hundreds or thousands of flights. If a passenger were to plan a trip, it's unreasonable to expect them to comb through hundreds of flight data just to find their options. This becomes a larger problem if a passenger were to plan multiple trips or a trip that consists of multiple flights; the problem only scales exponentially.

So, given a large set of flight data, a passenger would like to be able to:

- Search for specific airlines and airports
- Search for reasonable flights between locations
- Plan a trip that could consist of multiple flights

These are only some of many applications of the given problem. Thus, a flight search engine is the solution to these problems.

Project Goals

1. **Develop an End-to-End Application:**
   The primary objective of this project is to create a comprehensive and user-friendly airline data search engine. The application should be capable of efficiently retrieving information from the provided datasets and delivering results to users in a user-friendly and accurate manner.

2. **Implement an Efficient Algorithm:**
   To achieve this efficiently, we need to have an algorithm for data processing and retrieval. These algorithms could be MapReduce, SQL, Graph algorithms, this will come down to the different functionalities we would like to have.

3. **User-friendly:**
   The application we are planning to build should have user friendly functionality which would allow users to input complex queries easily, then be able to produce clear and organized results making it easier for the users to access the results they are searching for.

4. **Functionalities and Constraints:**
   The application has to support the following functionalities and constraints, airport and

airline search capabilities, airline aggregation and trip recommendation.

5. **Experimental Evaluation:**
   It is important to make sure our application performs successfully and accurately, it is also key to test and evaluate the performance of the application and the base algorithm. An evaluation could include a measure of response time for specific queries, assessing scalability of the system, as well as efficiency of the algorithms we are choosing to use for specific tasks.

Team Description

・Team members and skills

- **Steven Subianto** - Project Manager
  - C++/C#
  - Python

  Responsibilities: Decide on group direction / Schedule group meetings

- **Mark Strong-Shinozaki** - Development Manager
  - C/C++/C#
  - React.js
  - HTML and CSS

  Responsibilities: Allocate development tasks / Algorithm R&D

- **Yuuki Matsunari** - Task Manager
  - C++
  - Python

  Responsibilities: Keep team on task / Check deadlines

Dataset Description

This link (https://openflights.org/data.html) contains all the datasets that we will be using for the scope of this application. It is located in one centralized location with other datasets such as, planes.dat, countries.dat and routes.dat. It is possible that the two datasets specified (Airlines.dat & Airports.dat) will be combined with other datasets mentioned to assist in completing this application. If more datasets are used then it will be specifically specified within the codebase.

- In the **Airlines.dat** file, it includes data from January 2012, the Dataset contains 5888 airlines and each entry contains the following information: Airline ID, Name, Alias, IATA, ICAO, Callsign, Country and Active.

- In the **Airports.dat** file includes data from January 2017, the Dataset contains over 10,000 airports and each of entry contains the following information: Airport ID, Name, City, Country, IATA, ICAO, Latitude, Longitude, Altitude, Timzone, DST, TZ database time zone Type, Source.
- In the **planes.dat** file a curated selection of 173 passenger aircraft with IATA and/or ICAO codes and the following information: Name, IATA code, ICAO code.
- In the **countries.dat** file a list of ISO 3166-1 country codes and the following information: name, iso_code, and dafif_code.
- In the **routes.dat** file includes data from June 2014, the Dataset contains 67663 routes between 3321 airports on 548 airlines spanning the globe and the following information: Airline, Airline ID, Source airport, Source airport ID, Destination airport, Destination airport ID, Codeshare, Stops, Equipment

Dataset Statistics

- airport.dat (1101KB)
    - 14110 rows, 14 columns

- airlines.dat (388 KB)
    - 21317 rows, 7 columns

- planes.dat (9KB)
    - 246 rows, 3 columns

- countries.dat (6KB)
    - 3 rows, 3 columns

- routes.dat (2322KB)
    - 67663 rows, 9 columns

References

"Airport, airline and route data." OpenFlights.org. https://openflights.org/data.html (retrieved September 12, 2023)