Project – 3

Purpose Statement

The purpose of this project was to assess and evaluate airplane crashes from 1908-2019 in order to visualise them. The intention was that once ETL was complete we could analyse the data to show safety among airlines, number of incidents/crashes across decades, uncover the airline with the most crashes and inspect the percentage of fatalities to passengers per crash.
The two sources of data that are presented are connected by one common column that is the Airline name or "Operator". We can then decipher the information based upon this certain variable. There has also been an analysis to understand which airline by country has the most crashes, by filtering this information which country of origin is the most unreliable.

Team

1) Kelly
2) Warren Villarosa
3) Lily
4) Bimal
5) Divya

Data source - https://www.kaggle.com/chadmunger/plane-crashes-1908-2019

Objective : Exploratory Data analysis for plane crash between 1908 and 2019, underlying reasons, locations of the events, and finding trends.

Point of Interests:

1) Mapping the route and point of crash?
2) Countries/Region has had most of crash.
3) Worst Crashes-Highest fatality rate
4) Finding the trends - has the crashes increased over the time?
5) Categorizing the flights - military, passenger, private flights?
6) Crashes over the period? -pre 1940s and post 1940s

Methodology:

1) Cleaning the data and warehousing the data in PostgreSQL Database.
2) Creating Python-Flask API webpage and project the findings in the form of dashboard and analysis.
    a) Create Three Views (Index, Dashboard 1, Dashboard 2)
    b) Create API Endpoints that return JSON

Requirements:

1. Python Flask - powered API:
2. HTML/CSS: Web design
3. Javascript: Maps, graphs, drop-down menus etc.
4. At least one database (SQL, MongoDB, SQLite, etc.): Postgres

5. At least one JS library that we did not cover.
6. ~~Dataset must have at least 100 records.~~
7. Project must include some level of user-drive interaction (e.g. menus, dropdowns, textboxes).
8. Final visualisation should include at least three views.