Case Study

To create an end-to-end flow for executing the queries related to Airlines Analysis using Jupyter Notebook and generating visualizations.

We have two files:

1. Airports.csv
2. Flight_Data.csv

Files attached here with all schema and data creation script:

Airlines Data.7z

Tasks to be performed:

- Run the script “Script.sql” attached in the zipped folder “Airlines Data.7z” to create following tables:
  - Hub_Airport,
  - Hub_Flight1,
  - Link_Flight_Airport1
  - Sat_Airport and
  - Sat_Flight1.
- Load that dataset into the SQL DB.
- Clean the dataset, remove any non-required columns.
- Load that data into the Jupyter Notebook.
- Execute the following queries:
  1. Total No. of different flights running.
  2. Flights going to particular country
  3. Flights going to state
  4. Flights going to a city.
  5. Perform calculation to Identify Regional/International Airport.
  6. Flights going to every State.
  7. Find Different Airlines Available
  8. Which Airline has the maximum running flights
  9. No of Airports in the state AZ, DE, and NY.
- Get possible Visualization of data.