Details about the solution

The task is divided into 5 parts, I have made directory for each part of the deliverables.

- Part1:
    - Single python file for the basic ETL described in the **Data Pipeline Development Section (A very basic version)**

- Part2:
    - SQL File for DB design ant the required queries, it actually loads the data from the same table loaded in Part1

- Part3 and Part 4 (Single Folder):
    - Advanced ETL Pipeline (python file) with data quality checks and batch processing to load the json data.
    - SQL file with additional indexes and an optimized query

- Part5:
    - PBIX file of the dashboard design
    - PDF Document explaining the dashboard and dashboard screenshots.

- Additional
    - I scaled the dataset using a python script that I am attaching in the dataset creator folder as well.
    - Data folder contains the dataset I used.

I have tried adding comments and section in codes and queries for explanations as much as possible and I would be glad to explain if any part still needs explanation.

Thank you.

Wisal Mukhtiar