Student ID – AF0410889

Student name – Pragathi G Joshi

Project Title: Power BI Analysis of the Paris Olympics 2024

Objective

- This project aims to leverage Power BI to analyse and visualize various aspects of the Paris Olympics 2024.
- The goal is to provide insightful, interactive dashboards that track key performance metrics such as medal counts, athlete performance, event statistics, and country rankings.
- Integration data from multiple sources to provide a comprehensive overview of event results, athlete statistics, and country performance.
- Delivering insights on medal distribution, performance improvement, and country rankings, assisting in decision-making processes for future event planning and analytics.

Data Sources:

- Paris Olympics 2024 dataset sourced from Kaggle.
- Data imported into Power BI using a Python script.

Methodology

1. **Set Kaggle API Credentials Directory:** Configure the environment to point to the directory where your Kaggle API credentials (kaggle. json) are stored. This is essential for authenticating and interacting with the Kaggle API.
2. **Dataset Identifier and Download Path:** Specify the dataset identifier from Kaggle (e.g., for the "Paris 2024 Olympic Summer Games") and the local directory where the dataset files will be saved.
3. **Clean the Download Folder**: Optionally, delete any pre-existing files in the download folder to prevent any conflicts with old or duplicate files when downloading the new dataset.
4. **Download and Unzip Dataset**: Use the Kaggle API to download the dataset and automatically unzip it into the specified folder.
5. **List of CSV Files to Import:** Define the list of expected CSV files from the dataset. These CSV files contain the data you want to work with.
6. **Load CSV Files into Data Frames:** Load each CSV file into a Pandas DataFrame and store them in a dictionary. The dictionary's keys are the base names of the CSV files (without the .csv extension), which makes it easy to access the datasets.

Key Features:

1. **Medal Tracker:** A dashboard that visualizes the distribution of medals by country, sport, and individual athlete performance.
2. **Event Comparison:** Comparative analysis of events based on historical data, such as the number of participants, countries, and medal distribution.
3. **Athlete Performance:** Track individual athlete progress through their event participation, highlighting key performance indicators.
4. **Geographical Analysis:** A map-based visualization showing country rankings and medal distribution by region.

Expected Outcomes:

- Provide clear visualizations of country rankings and event statistics.
- Highlight trends in medal distribution, athlete performance, and event popularity.
- Help sports analysts, viewers, and stakeholders make informed predictions and decisions.

Tools and Technologies Used:

- Power BI
- CSV/Excel files
- Kaggle API
- Python