Python and SQL assignment

As part of this assignment, you would write a Python script to retrieve publicly available data and perform analysis on that data using SQL.

Publicly available data is available from the World Bank. One dataset which contains information about countries is available through World Bank API[1] and its documentation provides all the necessary details on how to use it. Another dataset which includes Gross Domestic Product (GDP) data in CSV format is available for download[2] from World Bank Data Catalog.

We recommend you keep notes of your design, problems you have encountered as well as steps you have taken to resolve them. We might ask you to present your solution.

Objectives

1. Write Python script to retrieve countries' data from a World Bank Country API as well GDP data from World Bank Data Catalog.
2. Store all data in a database; ideally PostgreSQL locally.
   - Model data in a way to be able to perform data analysis (see below).
3. Write SQL queries to answer questions below.

Analysis

1. List countries with income level of "Upper middle income".
2. List countries with income level of "Low income" per region.
3. Find the region with the highest proportion of "High income" countries.
4. Calculate cumulative/running value of GDP per region ordered by income from lowest to highest and country name.
5. Calculate percentage difference in value of GDP year-on-year per country.
6. List 3 countries with lowest GDP per region.
7. Provide an interesting fact from the dataset.

---

[1] https://datahelpdesk.worldbank.org/knowledgebase/articles/898590-api-country-queries
[2] https://datacatalog.worldbank.org/dataset/global-economic-prospects