Using web scraping to build a database of movie related information from: The Movie Database (TMDB) movie data

Problem statement:

A common business requirement in the context of information gathering is to extract and filter relevant data from web pages that host this information. However, access to information spread over several web pages, hosted potentially on multiple websites is a cumbersome process and we cannot rely on manual procedures to execute this task. In this project, you will employ a programmatic approach to access, parse and extract relevant information from a website of interest.

Objective:

The project's goal is to extract data (from a chosen number of pages) from The Movie Database website (https://www.themoviedb.org/) into a tabular data format so that further analysis (e.g., details about a movie's genre, cast, and user rating) can be facilitated.

To execute this project, you will have to read the documentation links provided against each task in the assignment and adapt the code examples provided in the documentation for the task at hand.

Pre-requisites :

Tools:
Jupyter Notebook or Google Colab or Microsoft Visual Studio IDE
Languages: Python, HTML
Libraries: requests, beautifulSoup, pandas