Practice Exercise: EDA With Python

The following is a post-class exercise for practicing exploratory data analysis using Python.

Note: This is neither a graded assessment nor has any time restraints for completion.

Case Study Number & Title	3. Analyzing the IMDB movie data for drawing inferences on the movie ratings
Background Information	IMDB dataset contains the top rated movies and television shows with information on the cast and crew, and the gross earnings from the box office or runtime on broadcasting channels.
Problem Statement/ Business objectives	Analyze the data to identify top grossing movies and shows, and explore using Python to extract relevant insights.
Data, Information for case analysis	Data is provided as a csv file. Below is the source and attribute information. Source link: https://www.kaggle.com/datasets/harshitshankhdhar/imdb-dataset-of-top-1000-movies-and-tv-shows Data Description **Series_Title:** Name of the movie **Released_Year:** Year of movie release **Certificate:** Certificate earned by the movie **Runtime:** Total runtime of the movie **Genre:** Genre of the movie **IMDB_Rating:** Rating of the movie as on IMDB website **Overview:** Summary of the movie **Meta_score:** Score earned by the movie **Director:** Name of the Director **Star1, Star2, Star3, Star4:** Name of the Stars **Noofvotes:** Total number of votes **Gross:** Gross collection earned by the movie
Questions	1. How many movies directed by Hayao Miyazaki the meta score of which lies between 85 and 100 have occupied this list? 2. What is the average rating of movies that have been PG-13 certified?

	3. Create a pivot table citing the median scores of A certified movies released post 2015. 4. Determine the least rating given for Francis Ford Coppola movies. 5. How many movies with actors Al Pacino and Robert De Niro as the first and second leads are present in the data?
Solution	A sample solution also provided with the dataset
Deliverables for Solution and Rubric	Non-graded assessment
Key Takeaways/Results	Exploring and analyzing data using Python and deriving meaningful insights.