1.PROBLEM

The business problem addressed in this portfolio project revolves around understanding and extracting insights from Netflix's movie dataset.

2.DATA CONNECTING TO EXCEL

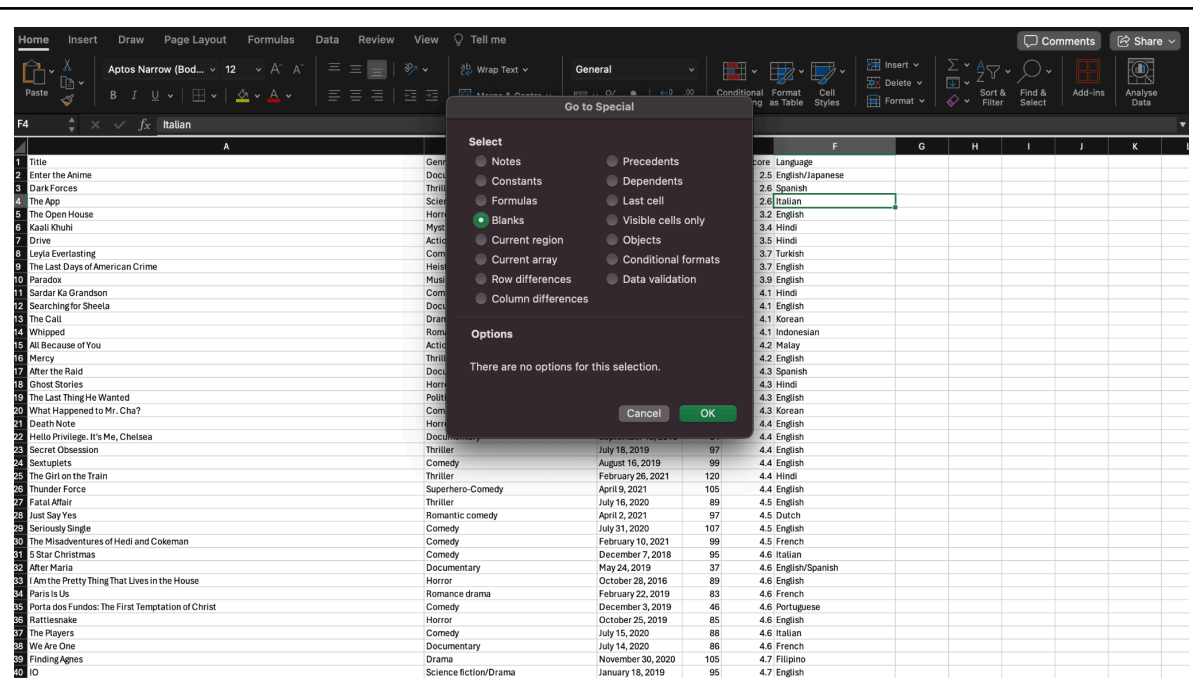To connect the dataset to Excel, you can follow these steps:

1. Download the Dataset: If you haven't already, download the Netflix movies dataset from the Kaggle link you provided or from any other source.

2. Open Excel: Launch Microsoft Excel on your computer.

3. Import External Data:
   - Go to the "Data" tab in Excel.
   - Click on "Get Data" or "From File" (the exact wording may vary depending on your Excel version).
   - Select "From CSV" if your dataset is in CSV format, or choose the appropriate option for other file formats.

4. Select the Dataset File: Browse to the location where you saved the downloaded dataset file and select it.

5. Configure Import Options:
   - Excel will open a dialog box where you can configure import options.
   - Verify that the delimiter and file origin settings are correct for your dataset. If needed, adjust these settings to ensure the data is imported correctly.
   - You may also preview the data to ensure it looks correct before importing.

6. Load the Data:
   - Once you've configured the import options, click on the "Import" button to import the dataset into Excel.

7. Data Refresh (Optional):
   - If your dataset is regularly updated or if you want to keep the Excel workbook connected to the original dataset, you can enable data refresh.
   - To do this, go to the "Data" tab, select "Queries & Connections" or "Refresh All," and choose the appropriate refresh options.

8. Verify Data: After importing, verify that the data is displayed correctly in Excel. You should see the dataset's contents in a new worksheet or in the Power Query Editor window if you used Power Query to import the data.

9. Save Your Excel Workbook: Once the data is imported and verified, save your Excel workbook to retain the connection to the dataset.

Following these steps will allow you to connect the dataset to Excel and begin analysing it using Excel's features and functions. If you encounter any issues during the import process or need further assistance, feel free to ask!
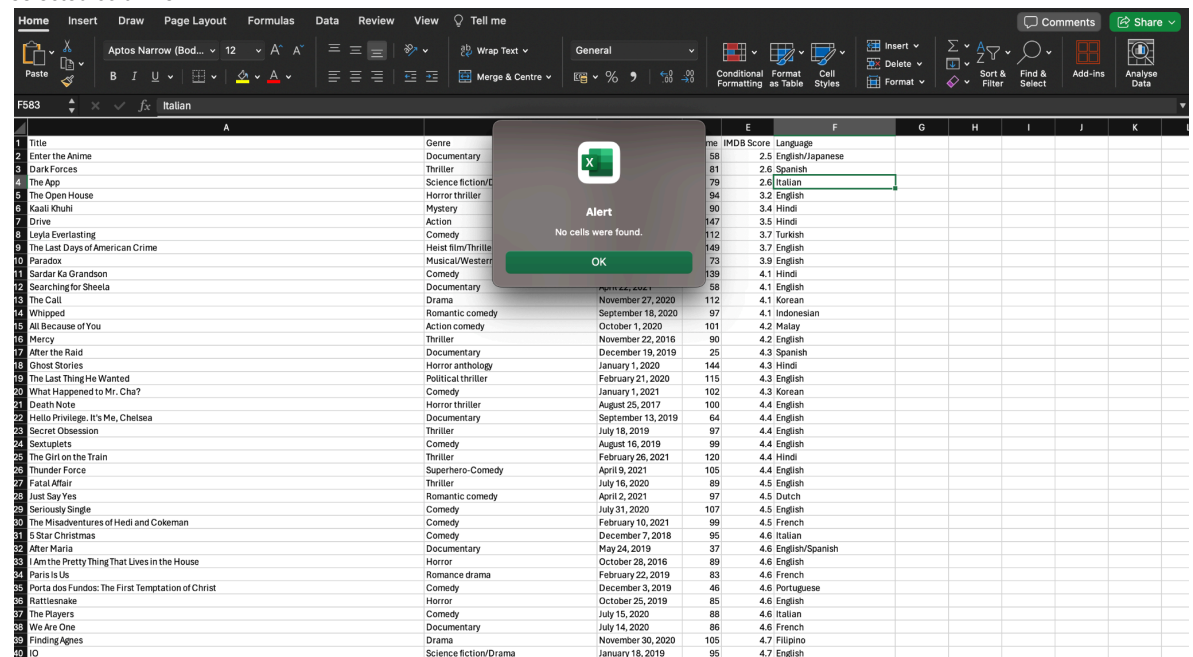
3.DATA CLEANING

Basic cleaning steps typically involve handling missing values, removing duplicates, and correcting any inconsistencies in the data. Here's how you can perform these steps in Excel:

1. Handling Missing Values:
   - Identify columns with missing values. Click on Find and Select option.
   -Click Go to special
   -click on the "Blanks" and OK.

Title	Genre			Score	Language
Enter the Anime	Docu...			2.5	English/Japanese
Dark Forces	Thrill...			2.6	Spanish
The App	Scie...			2.6	Italian
The Open House	Horr...			3.2	English
Kaali Khuhi	Myst...			3.4	Hindi
Drive	Actio...			3.5	Hindi
Leyla Everlasting	Com...			3.7	Turkish
The Last Days of American Crime	Heis...			3.7	English
Paradox	Musi...			3.9	English
Sardar Ka Grandson	Com...			4.1	Hindi
Searching for Sheela	Docu...			4.1	English
The Call	Dram...			4.1	Korean
Whipped	Rom...			4.1	Indonesian
All Because of You	Actio...			4.2	Malay
Mercy	Thrill...			4.2	English
After the Raid	Docu...			4.3	Spanish
Ghost Stories	Horr...			4.3	Hindi
The Last Thing He Wanted	Polit...			4.3	English
What Happened to Mr. Cha?	Com...			4.3	Korean
Death Note	Horr...			4.4	English
Hello Privilege. It's Me, Chelsea	Documentary	September 13, 2019		4.4	English
Secret Obsession	Thriller	July 18, 2019	97	4.4	English
Sextuplets	Comedy	August 16, 2019	99	4.4	English
The Girl on the Train	Thriller	February 26, 2021	120	4.4	Hindi
Thunder Force	Superhero-Comedy	April 9, 2021	105	4.4	English
Fatal Affair	Thriller	July 16, 2020	89	4.5	English
Just Say Yes	Romantic comedy	April 2, 2021	97	4.5	Dutch
Seriously Single	Comedy	July 31, 2020	107	4.5	English
The Misadventures of Hedi and Cokeman	Comedy	February 10, 2021	99	4.5	French
5 Star Christmas	Comedy	December 7, 2018	95	4.6	Italian
After Maria	Documentary	May 24, 2019	37	4.6	English/Spanish
I Am the Pretty Thing That Lives in the House	Horror	October 28, 2016	89	4.6	English
Paris Is Us	Romance drama	February 22, 2019	83	4.6	French
Porta dos Fundos: The First Temptation of Christ	Comedy	December 3, 2019	46	4.6	Portuguese
Rattlesnake	Horror	October 25, 2019	85	4.6	English
The Players	Comedy	July 15, 2020	88	4.6	Italian
We Are One	Documentary	July 14, 2020	86	4.6	French
Finding Agnes	Drama	November 30, 2020	105	4.7	Filipino
IO	Science fiction/Drama	January 18, 2019	95	4.7	English

2. Removing Duplicates:

- Identify columns where duplicates may exist. This is often relevant for categorical data such as movie titles or genres.
- Select the range of data containing potential duplicates.
- Go to the "Data" tab and click on "Remove Duplicates."
- Choose the columns you want to check for duplicates and click "OK." Excel will remove duplicate rows based on the selected columns.

Title	Genre			IMDB Score	Language
Enter the Anime	Documentary		58	2.5	English/Japanese
Dark Forces	Thriller		81	2.6	Spanish
The App	Science fiction/D...		79	2.6	Italian
The Open House	Horror thriller		94	3.2	English
Kaali Khuhi	Mystery		90	3.4	Hindi
Drive	Action		147	3.5	Hindi
Leyla Everlasting	Comedy		112	3.7	Turkish
The Last Days of American Crime	Heist film/Thriller		149	3.7	English
Paradox	Musical/Western		73	3.9	English
Sardar Ka Grandson	Comedy		139	4.1	Hindi
Searching for Sheela	Documentary	April 22, 2021	58	4.1	English
The Call	Drama	November 27, 2020	112	4.1	Korean
Whipped	Romantic comedy	September 18, 2020	97	4.1	Indonesian
All Because of You	Action comedy	October 1, 2020	101	4.2	Malay
Mercy	Thriller	November 22, 2016	90	4.2	English
After the Raid	Documentary	December 19, 2019	25	4.3	Spanish
Ghost Stories	Horror anthology	January 1, 2020	144	4.3	Hindi
The Last Thing He Wanted	Political thriller	February 21, 2020	115	4.3	English
What Happened to Mr. Cha?	Comedy	January 1, 2021	102	4.3	Korean
Death Note	Horror thriller	August 25, 2017	100	4.4	English
Hello Privilege. It's Me, Chelsea	Documentary	September 13, 2019	64	4.4	English
Secret Obsession	Thriller	July 18, 2019	97	4.4	English
Sextuplets	Comedy	August 16, 2019	99	4.4	English
The Girl on the Train	Thriller	February 26, 2021	120	4.4	Hindi
Thunder Force	Superhero-Comedy	April 9, 2021	105	4.4	English
Fatal Affair	Thriller	July 16, 2020	89	4.5	English
Just Say Yes	Romantic comedy	April 2, 2021	97	4.5	Dutch
Seriously Single	Comedy	July 31, 2020	107	4.5	English
The Misadventures of Hedi and Cokeman	Comedy	February 10, 2021	99	4.5	French
5 Star Christmas	Comedy	December 7, 2018	95	4.6	Italian
After Maria	Documentary	May 24, 2019	37	4.6	English/Spanish
I Am the Pretty Thing That Lives in the House	Horror	October 28, 2016	89	4.6	English
Paris Is Us	Romance drama	February 22, 2019	83	4.6	French
Porta dos Fundos: The First Temptation of Christ	Comedy	December 3, 2019	46	4.6	Portuguese
Rattlesnake	Horror	October 25, 2019	85	4.6	English
The Players	Comedy	July 15, 2020	88	4.6	Italian
We Are One	Documentary	July 14, 2020	86	4.6	French
Finding Agnes	Drama	November 30, 2020	105	4.7	Filipino
IO	Science fiction/Drama	January 18, 2019	95	4.7	English

4.Splitting Date:

To split the "Premiere" date column into separate columns for date, month, and year in Excel, you can follow these steps:

Insert New Columns:

- Insert three new columns next to the "Premiere" column to accommodate the split date, month, and year.

Split the Date:

- Select the first cell under the "Date" column (the one you just inserted).
- Go to the "Data" tab.
- Click on "Text to Columns" in the Data Tools group.

- In the Convert Text to Columns Wizard, choose the "Delimited" option and click "Next."
- Choose the delimiter option that separates the date components in your "Premiere" column (e.g., if dates are separated by slashes (/), select "Delimited" and then check the "Other" box and enter "/" as the delimiter). Click "Next."
- In the next step, you can choose the data format for each column. For the "Date" column, select "Do not import column (skip)" since you're splitting the date into separate columns. For the "Month" and "Year" columns, choose the appropriate date format (e.g., "MDY" for Month-Day-Year) and click "Finish."

Adjust Column Widths (Optional):
- Once the date, month, and year are split into separate columns, you may need to adjust the column widths to make sure all the data is visible.

Verify the Results:
- Verify that the date, month, and year are split correctly across the three columns. Check a few random rows to ensure accuracy.

Your "Premiere" date column should now be split into separate columns for date, month, and year in Excel. You can now use these columns for further analysis or formatting as needed.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Title	Genre	Month	Date	Year	Runtime	IMDB Score	Language				
2	Enter the Anime	Documentary	August	5	2019	58	2.5	English/Japanese				
3	Dark Forces	Thriller	August	21	2020	81	2.6	Spanish				
4	The App	Science fiction/Drama	December	26	2019	79	2.6	Italian				
5	The Open House	Horror thriller	January	19	2018	94	3.2	English				
6	Kaali Khuhi	Mystery	October	30	2020	90	3.4	Hindi				
7	Drive	Action	November	1	2019	147	3.5	Hindi				
8	Leyla Everlasting	Comedy	December	4	2020	112	3.7	Turkish				
9	The Last Days of American Crime	Heist film/Thriller	June	5	2020	149	3.7	English				
10	Paradox	Musical/Western/Fantasy	March	23	2018	73	3.9	English				
11	Sardar Ka Grandson	Comedy	May	18	2021	139	4.1	Hindi				
12	Searching for Sheela	Documentary	April	22	2021	58	4.1	English				
13	The Call	Drama	November	27	2020	112	4.1	Korean				
14	Whipped	Romantic comedy	September	18	2020	97	4.1	Indonesian				
15	All Because of You	Action comedy	October	1	2020	101	4.2	Malay				
16	Mercy	Thriller	November	22	2016	90	4.2	English				
17	After the Raid	Documentary	December	19	2019	25	4.3	Spanish				
18	Ghost Stories	Horror anthology	January	1	2020	144	4.3	Hindi				
19	The Last Thing He Wanted	Political thriller	February	21	2020	115	4.3	English				
20	What Happened to Mr. Cha?	Comedy	January	1	2021	102	4.3	Korean				
21	Death Note	Horror thriller	August	25	2017	100	4.4	English				
22	Hello Privilege. It's Me, Chelsea	Documentary	September	13	2019	64	4.4	English				
23	Secret Obsession	Thriller	July	18	2019	97	4.4	English				
24	Sextuplets	Comedy	August	16	2019	99	4.4	English				
25	The Girl on the Train	Thriller	February	26	2021	120	4.4	Hindi				
26	Thunder Force	Superhero-Comedy	April	9	2021	105	4.4	English				
27	Fatal Affair	Thriller	July	16	2020	89	4.5	English				
28	Just Say Yes	Romantic comedy	April	2	2021	97	4.5	Dutch				
29	Seriously Single	Comedy	July	31	2020	107	4.5	English				
30	The Misadventures of Hedi and Cokeman	Comedy	February	10	2021	99	4.5	French				
31	5 Star Christmas	Comedy	December	7	2018	95	4.6	Italian				
32	After Maria	Documentary	May	24	2019	37	4.6	English/Spanish				
33	I Am the Pretty Thing That Lives in the House	Horror	October	28	2016	89	4.6	English				
34	Paris Is Us	Romance drama	February	22	2019	83	4.6	French				
35	Porta dos Fundos: The First Temptation of Christ	Comedy	December	3	2019	46	4.6	Portuguese				
36	Rattlesnake	Horror	October	25	2019	85	4.6	English				
37	The Players	Comedy	July	15	2020	88	4.6	Italian				
38	We Are One	Documentary	July	14	2020	86	4.6	French				
39	Finding Agnes	Drama	November	30	2020	105	4.7	Filipino				
40	IO	Science fiction/Drama	January	18	2019	95	4.7	English				
41	Sentinelle	Action	March	5	2021	80	4.7	French				
42	Sol Levante	Anime / Short	April	2	2020	4	4.7	English				
43	The Binding	Drama	October	2	2020	93	4.7	Italian				
44	We Can Be Heroes	Superhero	December	25	2020	100	4.7	English				
45	Christmas Crossfire	Thriller	December	4	2020	106	4.8	German				
	Coin Heist	Heist	January	6	2017	97	4.8	English				

5. Data Format Conversion:
   - Ensure that data is in the correct format for analysis. For example, dates should be in the date format, numerical data should be in numeric format, and text data should be in text format.
   - Use Excel's formatting options to convert data to the desired format. Right-click on the column header, choose "Format Cells," and select the appropriate format.

By following these basic cleaning steps, you can prepare your dataset for further analysis in Excel. Remember to save your cleaned dataset as a separate file to avoid overwriting the original data.

4. DATA ANALYSING QUESTIONS

Certainly! Here are some questions you could explore with the given dataset:

1. Genre Analysis:
   - Which genres are most prevalent in the Netflix dataset?

2. Premiere Analysis:
   - Are there any notable trends in premieres over time?

3. IMDb Score Analysis:
   - Are there any trends in IMDb scores over the years?

4. Language Analysis:
   - Which languages are most common in Netflix movies?

5. Top Movies Analysis:
   - What are the top-rated movies on Netflix based on IMDb score?

These questions should give you a good starting point for analyzing the Netflix movie dataset. Depending on your interests and objectives, you can delve deeper into specific aspects of the data.

5.DATA ANALYSIS STEPS

Here's how you can use pivot tables in Excel to analyze the Netflix movie dataset based on the provided questions:

1. Genre Analysis:
   - Insert a pivot table.
   - Drag the "Genre" field to the Rows area.
   - Drag any numerical field (e.g., "Title") to the Values area to see counts by genre.
   - Optionally, you can add filters for other fields like "Year" to analyze genre trends over time.
   -Select the Pivot Table
   -Go to the Insert Menu and choose the appropriate chart as mentioned in the **6.DATA VISUALIZATION**

2. Year Analysis:
   - Insert a pivot table.
   - Drag the "Year" field to the Rows area.
   - Drag the "Title" field to the Values area and set it to count to see the number of movies premiered each year.
   - Optionally, you can group years by quarters or months to identify any seasonal patterns.
   -Select the Pivot Table
   -Go to the Insert Menu and choose the appropriate chart as mentioned in the **6.DATA VISUALIZATION**

3. IMDb Score Analysis:
   - Insert a pivot table.
   - Drag the "IMDb Score" field to the Values area and set it to count to see the IMDb scores of movies.
   - You can also analyze IMDb score trends over time by dragging the "Year" field to the Rows area.
   -Select the Pivot Table
   -Go to the Insert Menu and choose the appropriate chart as mentioned in the **6.DATA VISUALIZATION**

4. Language Analysis:
   - Insert a pivot table.
   - Drag the "Language" field to the Rows area.
   - Drag any numerical field (e.g., "Title") to the Values area to see counts by language.
   - Optionally, you can add filters for other fields like "Genre" to analyze language trends within genres.
   -Select the Pivot Table
   -Go to the Insert Menu and choose the appropriate chart as mentioned in the **6.DATA VISUALIZATION**

5. Top Movies Analysis:
   - Insert a pivot table.
   - Drag the "Title" field to the Rows area.
   - Drag the "IMDb Score" field to the Values area and set it to max to see the highest IMDb score for each movie.
   - You can also include other fields like "Genre" or "Language" to analyze characteristics of top-rated movies.
   -Select the Pivot Table
   -Go to the Insert Menu and choose the appropriate chart as mentioned in the **6.DATA VISUALIZATION**

These steps should help you create pivot tables to analyze various aspects of the Netflix movie dataset. Adjust the fields and settings based on your specific analysis needs.

6.DATA VISUALIZATION

1. Genre Analysis:
   - Bar chart for visualizing the distribution of movies across different genres.

2. Year Analysis:
   - Bar chart or bar chart to visualize the number of movies premiered each year.

3. IMDb Score Analysis:
   - Line chart to show trends in IMDb scores over the years.

5. Language Analysis:
   - Bar chart or pie chart to visualize the distribution of movies by language.

6. Top Movies Analysis:
   - Bar chart to display the top-rated movies based on IMDb score.

These chart types should help you effectively visualize and analyze the various aspects of the Netflix movie dataset. Adjustments can be made based on the specific insights you want to highlight or the audience you're presenting to.