

Excel Exercise: IMDB Movie Analysis

Objective:

In this exercise, you will use Excel to analyse **IMDB Top 250 Movies** and **IMDB Top 250 English Movies**. You will learn how to:

- **Import data from the web**
 - **Sort data**
 - **Apply formulas and functions**, including:
 - VLOOKUP
 - IF, SUMIF, COUNTIF
 - SUM, MAX, MIN
 - AND, OR
 - FILTER
 - **Use Conditional Formatting** to highlight key information.
-

Step 1: Import Data from the Web

1. Open **Excel**.
2. Rename **Sheet1** as "**Top 250 Movies**".
3. Go to **Data** → **Get & Transform Data** → **From Web**.
4. Enter the **IMDB Top 250 Movies** URL and click **OK**:
`https://mikhail-ict.github.io/imdb`
5. Use the **Navigator Pane** to select the correct table.
6. Cleanse the dataset to make sure you get the following columns: **Rank**, **Title**, **Year**, **Rating**
7. Click **Load** to insert the data into "**Top 250 Movies**".
8. Repeat the steps for the **IMDB Top 250 English Movies** and load them into a new sheet named

"Top 250 English Movies":

```
https://mikhail-ict.github.io/imdb-english
```

Step 2: Sorting Data

Task 1: Sort by Rating (Highest to Lowest)

- Sort the "Top 250 Movies" sheet by "Rating" from highest to lowest.

Question: Which movie has the highest rating?

Task 2: Sort by Year (Newest to Oldest)


- Sort the movies by "Year" from newest to oldest.

Question: What is the most recent movie in the dataset?

Step 3: Apply Formulas & Functions

Task 3: Identify English Movies (VLOOKUP + IFERROR)

- Create a new column in the "Top 250 Movies" sheet named "Language".
- Use VLOOKUP to check if the movie title exists in the "Top 250 English Movies" sheet.
- If the movie is found, mark it as "English", otherwise, mark it as "Non-English".

 *Hint:* Use VLOOKUP combined with an IFERROR or IF & ISNA functions.

Question: How many Non-English movies are in the Top 250?

Task 4: Categorise Movies by Rating (IF Function)

- **Create a new column** named **"Rating Category"**.
- **Use IF** to classify movies:
 - If **Rating > 8.5**, label it **"Highly Rated"**.
 - Otherwise, label it **"Average"**.

Question: How many movies are categorised as "Highly Rated"?

Task 5: Count Movies Above a Certain Rating (COUNTIF)

- **Use COUNTIF** to count how many movies have a **rating above 9**.

Question: How many movies have a rating above 9?

Task 6: Calculate the Total and Average Ratings (SUM , SUMIF)

- **Use SUM** to calculate the total of all ratings.
- **Use SUMIF** to calculate the total ratings **only for movies released after 2000**.
- **Use AVERAGE** to find the **average rating** for all movies.

Question: What is the average rating of all movies?

Task 7: Find the Highest & Lowest Ratings (MAX & MIN)

- **Use MAX** to find the highest rating.
- **Use MIN** to find the lowest rating.

Question: What is the highest-rated movie on the list?

Question: What is the lowest-rated movie?

Task 8: Identify Modern Classics (AND Function)

- **Create a new column** named **"Modern Classic?"**.
- **Use** **AND** in an **IF** statement to classify a movie as **"Yes"** if:
 - **Rating > 8.5 AND Year > 2000.**

Question: How many movies qualify as a "Modern Classic"?

Task 9: Identify Popular or Highly Rated Movies (**OR** Function)

- **Create a new column** named **"Popular or Highly Rated?"**.
- **Use** **OR** in an **IF** statement to classify a movie as **"Yes"** if:
 - **Rating > 8 OR Year > 2010.**

Question: How many movies meet at least one of these criteria?

Task 10: Use **FILTER** to Display Highly Rated Movies

- **Use the** **FILTER** **function** to display only movies with a **rating above 8.8.**

Question: How many movies have a rating above 8.8?

Step 4: Apply Conditional Formatting

Task 11: Highlight High-Rated Movies

- **Use Conditional Formatting** to highlight movies with a **rating greater than 9.**
 - **Question:** What colour did you use for formatting?
-

Task 12: Highlight Old Movies

- **Use Conditional Formatting** to highlight movies released **before 1970.**

Question: How many movies were released before 1970?

Task 13: Highlight English Movies

- **Use Conditional Formatting** to highlight English movies in the **"Language"** column.

Question: What percentage of the Top 250 Movies are in English?



Final Review Questions

1. Which movie has the highest rating?
 2. What is the most recent movie in the dataset?
 3. How many **Non-English** movies are in the Top 250?
 4. How many movies are categorised as "Highly Rated"?
 5. How many movies have a rating above 9?
 6. What is the average rating of all movies?
 7. What is the highest-rated movie on the list?
 8. What is the lowest-rated movie?
 9. How many movies qualify as a "Modern Classic"?
 10. How many movies meet at least one of these criteria?
 11. How many movies have a rating above 8.8?
 12. What colour did you use for formatting high-rated movies?
 13. How many movies were released before 1970?
 14. What percentage of the Top 250 Movies are in English?
-



Submission Instructions

- Save your Excel file as **IMDB_Analysis_(Your_Name).xlsx**.
- Submit it via **Padlet**.
- Put in the answers for each of the above questions right in a card on **Padlet** where you

submitted the file.

#datainsight/day2