

Excel Exercise: IMDB Movie Analysis (Solutions)

Step 1: Import Data from the Web

Since your data is coming from the web, follow these steps to import it:

1. Open **Excel**.
 2. Rename **Sheet1** as "**Top 250 Movies**".
 3. Go to **Data** → **Get & Transform Data** → **From Web**.
 4. Enter the following URL for the **Top 250 Movies** list 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 it contains the following columns:
Rank, Title, Year, Rating
 7. Click **Load** to insert the data into the "**Top 250 Movies**" sheet.
 8. Repeat the same process for the **Top 250 English Movies** list using the URL:
`https://mikhail-ict.github.io/imdb-english`
 9. Rename the new sheet as "**Top 250 English Movies**".
-

Step 2: Sorting Data

Task 1: Sort by Rating (Highest to Lowest)

1. Select the dataset (Columns A to D).
2. Go to **Data** → Click **Sort**.
3. In the **Sort** window:
 - **Sort by:** "Rating"

- **Order:** Largest to Smallest

4. Click **OK**.

Question: Which movie has the highest rating? (*Check the first row after sorting.*)

Task 2: Sort by Year (Newest to Oldest)

1. Select the dataset.
2. Go to **Data** → Click **Sort**.
3. In the **Sort** window:
 - **Sort by:** "Year"
 - **Order:** Newest to Oldest
4. Click **OK**.

Question: What is the most recent movie in the dataset? (*Check the first row after sorting.*)



Step 3: Apply Formulas & Functions

Task 3: Identify English Movies (**VLOOKUP** + **IFERROR**)

1. Insert a new column **E** in the "**Top 250 Movies**" sheet and name it "**Language**".
2. In cell **E2**, enter the following formula:

```
=IFERROR(IF(VLOOKUP([@Title],EnglishMovies[Title],1,FALSE) <> "",  
"English"), "Non-English")
```

3. Press **Enter**.
4. Drag the formula down to apply it to all rows.

Question: How many **Non-English** movies are in the Top 250? (*Use **COUNTIF** on Column E to count "Non-English."*)

Task 4: Categorise Movies by Rating (**IF** Function)

1. Insert a new column **F** and name it "**Rating Category**".

2. In **cell F2**, enter:

```
=IF(D2>8.5, "Highly Rated", "Average")
```

3. Drag the formula down.

Question: How many movies are categorized as "Highly Rated"? (Use **COUNTIF** on Column F.)

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

1. In an empty cell (e.g., **H2**), enter:

```
=COUNTIF(D:D, ">9")
```

2. Press **Enter**.

Question: How many movies have a rating above 9?

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

1. **Total of all ratings:**

```
=SUM(D:D)
```

2. **Total of ratings for movies released after 2000:**

```
=SUMIF(C:C, ">2000", D:D)
```

3. **Average rating of all movies:**

```
=AVERAGE(D:D)
```

Question: What is the average rating of all movies?

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

1. Highest Rating:

```
=MAX(D:D)
```

2. Lowest Rating:

```
=MIN(D:D)
```

Question: What is the highest-rated movie on the list? (Find using MAX.)

Question: What is the lowest-rated movie? (Find using MIN.)

Task 8: Identify Modern Classics (AND Function)

1. Insert a new column **G** and name it "Modern Classic?".
2. In cell **G2**, enter:

```
=IF(AND(D2>8.5, C2>2000), "Yes", "No")
```

3. Drag the formula down.

Question: How many movies qualify as a "Modern Classic"? (Use COUNTIF on Column G.)

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

1. Insert a new column **H** and name it "Popular or Highly Rated?".
2. In cell **H2**, enter:

```
=IF(OR(D2>8, C2>2010), "Yes", "No")
```

3. Drag the formula down.

Question: How many movies meet at least one of these criteria? (Use **COUNTIF** on Column H.)

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

1. Click on an empty cell (e.g., **J2**).
2. Enter:

```
=FILTER(A:D, D:D>8.8, "No movies found")
```

3. Press **Enter**.

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 on **Padlet**.

#datainsight/day2