

# IMDB Movie Analysis

## Project Description

This project is about the data analysis using excel sheet and statistics.

IMDB movie analysis is the fundamental and the most important to get insight related to movies. This analysis is very useful to identify profit or loss of movie, favorite actor, likes and views.

Statistics is collection, organization, analysis, implementation and presentation of data. It is a review, gather, analyze and raw conclusion from data. Here we also use pie chart, bar chart, histogram, frequency table and many more to understanding data in graphical format which is very easy to understanding. Also, here we can use 5 Whys technique to determine its root cause by repeatedly asking the question “Why”. It will very helpful to understand data and tasks.

## Approach

IMDB movie analysis is also an important part of data analyst. Here we use excel sheet and statistics to get output from insights. Here they want some data related to highest profitable movie, top 250 movie, best director, popular genera's, critics and audience's favorite actor and some graph from data set. Also they want clean dataset on existion dataset

With the help of Microsoft office Excel & Google Sheet, resolve query with the help of video lecture and database which already provided in our course. They provide relevant data like movie name, year, actors name, director name, movie bugget, gross, movie rating, user votes, facebook likes, duration and many more. Using these data and their requirement we did some operation on the data.

## Tech-Stack Used

### 1. Microsoft Excel

Analyze Data in Excel empowers you to understand your data through natural language queries that allow you to ask questions about your data without having to write complicated formulas. In addition, Analyze Data provides high-level visual summaries, trends, and patterns.

### 2. Google Sheet

Get summaries and charts of your data with the click of a button in Sheets. It's kind of like having an expert in data analysis sitting next to you. In a spreadsheet, select a range of cells, columns, or rows. Otherwise, you'll get insights based on where your cursor is.

### 3. Google Colab

Use to clean dataset and perform some python groupby operation to get data.

## Result

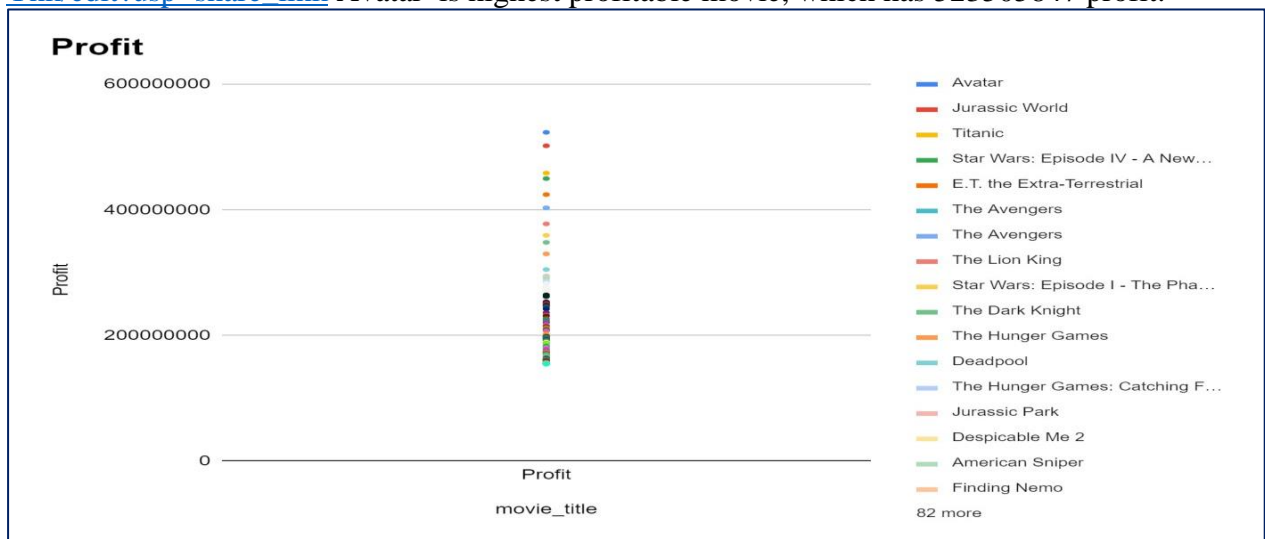
### A : Cleaning the data Clean the data

[https://docs.google.com/spreadsheets/d/1yRDBRy-h9qcXDyCaCgUqJRhGbJaS0T\\_xCWIFNYSk9Y/edit?usp=share\\_link](https://docs.google.com/spreadsheets/d/1yRDBRy-h9qcXDyCaCgUqJRhGbJaS0T_xCWIFNYSk9Y/edit?usp=share_link)

### B : Movies with highest profit

Find the movies with the highest profit?

[https://docs.google.com/spreadsheets/d/1oB\\_WZ6o1Wh16JCd7Va9mff0vb1Y2bmp7aFTmA7BYnk/edit?usp=share\\_link](https://docs.google.com/spreadsheets/d/1oB_WZ6o1Wh16JCd7Va9mff0vb1Y2bmp7aFTmA7BYnk/edit?usp=share_link) Avatar is highest profitable movie, which has 523505847 profit.



### C : Top 250 Find IMDB Top 250

[https://docs.google.com/spreadsheets/d/1oQb4y4YZ-6BOZiIdLqkmUpRGEZuPaFOab2nTzwUWufM/edit?usp=share\\_link](https://docs.google.com/spreadsheets/d/1oQb4y4YZ-6BOZiIdLqkmUpRGEZuPaFOab2nTzwUWufM/edit?usp=share_link)

1	2	3	4	5
1	2	3	4	5
1	1	The Shawshank Redem	1689764	9.3
2	2	The Godfather	1155770	9.2
3	3	The Dark Knight	1676169	9
4	4	The Godfather: Part II	790926	9
5	5	Pulp Fiction	1324680	8.9
6	6	The Lord of the Rings: The	1215718	8.9
7	7	Schindler's List	865020	8.9
8	8	The Good, the Bad and	503509	8.9
9	9	Inception	1468200	8.8
10	10	Fight Club	1347461	8.8
11	11	Forrest Gump	1251222	8.8
12	12	The Lord of the Rings: The	1238746	8.8
13	13	Star Wars: Episode V -	837759	8.8
14	14	The Matrix	1217752	8.7
15	15	The Lord of the Rings: The	1100446	8.7
16	16	Star Wars: Episode IV -	911097	8.7
17	17	The Godfather: Part I	790926	8.7
18	18	The Godfather: Part II	790926	8.7

## **D : Best Directors**

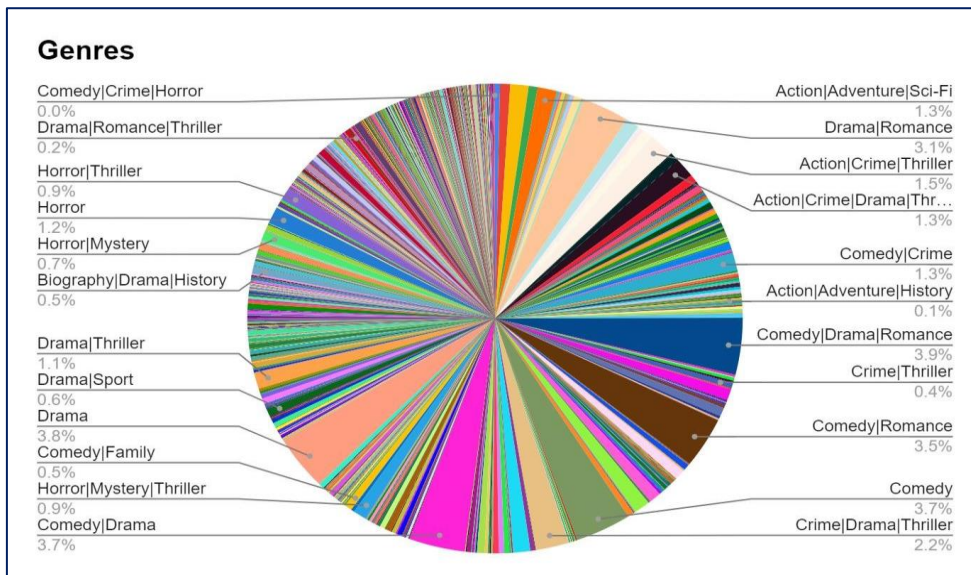
Find the best directors

[https://docs.google.com/spreadsheets/d/1P2zo2AZmsoT105laGIUWCppMId6yKG4JxC8P4WD0cy8/edit?usp=share\\_link](https://docs.google.com/spreadsheets/d/1P2zo2AZmsoT105laGIUWCppMId6yKG4JxC8P4WD0cy8/edit?usp=share_link)

	top10director	movie_title	imdb_score
1	Frank Darabont	The Shawshank	9.3
2	Francis Ford Coppola	The Godfather	9.2
3	Christopher Nolan	The Dark Knight	9
4	Peter Jackson	The Lord of the Rings	8.9
5	Quentin Tarantino	Pulp Fiction	8.9
6	Sergio Leone	The Good, the Bad and the Ugly	8.9
7	Steven Spielberg	Schindler's List	8.9
8	David Fincher	Fight Club	8.8
9	Irvin Kershner	Star Wars: Episode V - The Empire Strikes Back	8.8
10	Robert Zemeckis	Forrest Gump	8.8

## **E: Popular Genres** Find popular genres

[https://docs.google.com/spreadsheets/d/1qGZY6OHU4n0Fc8AzZS6Ww-Dd39A80KW03jzFBVZXqI/edit?usp=share\\_link](https://docs.google.com/spreadsheets/d/1qGZY6OHU4n0Fc8AzZS6Ww-Dd39A80KW03jzFBVZXqI/edit?usp=share_link)



## E : Charts

Find the critic-favorite and audience-favorite actors

[https://docs.google.com/spreadsheets/d/1YgADqEbrv2nXic9vwdN72MUKXUr9tSjn/edit?usp=share\\_link&ouid=114387490144908412553&rtpof=true&sd=true](https://docs.google.com/spreadsheets/d/1YgADqEbrv2nXic9vwdN72MUKXUr9tSjn/edit?usp=share_link&ouid=114387490144908412553&rtpof=true&sd=true)

Leonardo DiCaprio is critic-favorite and audience-favorite actors based on critics review and user review.

num_critic_for_reviews		num_user_for_reviews	
Brad Pitt	245	Brad Pitt	742.3529412
Leonardo DiCaprio	330.1905	Leonardo DiCaprio	914.4761905
Meryl Streep	181.4545	Meryl Streep	297.1818182

[https://docs.google.com/spreadsheets/d/1IGhhCCz1f00IK9PCqhVxtvCPdHOIEWWtSuWOCJYnUrY/edit?usp=share\\_link](https://docs.google.com/spreadsheets/d/1IGhhCCz1f00IK9PCqhVxtvCPdHOIEWWtSuWOCJYnUrY/edit?usp=share_link)

