IMDB RATING ANALYSIS

—

PROJECT 5

—

YASHI GUPTA

PROJECT DESCRIPTION

The project aims to analyze imdb movies data to determine the factors that contribute to a movie's success based on its imdb score. The main goal is to examine various elements of movies, including genre, duration, language, director, and budget, to understand how these factors influence the ratings and overall success of a movie on IMDb. By comprehending these connections, individuals involved in the movie industry, such as producers, directors, and investors, can make well-informed choices in their upcoming projects to increase their likelihood of achieving success.

Objectives

1. Determine the most prevalent movie genres and examine their influence on imdb ratings.
2. Analyze the correlation between the length of a movie and its rating on IMDb. Examine the impact of language on the evaluation of movies.
3. Examine the influence of directors in determining the outcome of a film.
4. Investigate the correlation between movie budgets and their financial achievements.

Approach

The approach involved collecting and cleaning the IMDB movies dataset to make it suitable for analysis. Data preprocessing included handling missing values, converting data types where necessary, and feature engineering to ensure accuracy and consistency. After the data was prepared, various analytical techniques were employed using Microsoft Excel to calculate descriptive statistics, correlations, and percentiles to uncover meaningful patterns. The analysis was broken down into distinct tasks based on different factors like genre, duration, language, director, and budget.

The following tasks were executed:

1. **Movie Genre Analysis**: Distribution of movie genres and their relationship with IMDB ratings using descriptive statistics.
2. **Movie Duration Analysis**: Exploring the relationship between movie durations and IMDB ratings through scatter plots and trendlines.
3. **Language Analysis**: Analyzing the impact of movie language on IMDB ratings.
4. **Director Analysis**: Identifying the top directors and their contribution to movie ratings using percentiles.
5. **Budget Analysis**: Analyzing the correlation between movie budgets and gross earnings to identify the movies with the highest profit margins.

Approach

To approach the analysis, I began by cleaning and structuring the data to ensure its consistency. I then used **Microsoft Excel** as the primary tool for performing the analysis due to its powerful functions for data manipulation and statistical analysis. Specifically, I used the following steps:

1. **Data Preprocessing**:
   - Handled missing values where applicable.
   - Removed any duplicates to ensure that the dataset was accurate.
   - Removed useless columns.
   - Converted data types where necessary (e.g., dates, numeric values).

2. **Analysis**:
   - Used **COUNTIF**, **AVERAGE**, **MEDIAN**, **MODE**, **MAX**, **MIN**, **VAR**, and **STDEV** functions for descriptive statistics to analyze each factor (genre, duration, language, director, and budget).
   - Created **scatter plots** and added **trendlines** to visualize the relationship between **duration** and IMDB ratings.
   - Employed **PERCENTILE** function to identify the directors with the highest IMDB scores and compare them to the overall distribution.
   - Used **CORREL** to analyze the relationship between **budget** and **gross earnings**.

3. **Visualization**:
   - Created various charts (e.g., bar charts, scatter plots) to visualize the results and trends.
   - Used trendlines to demonstrate the correlation between variables.

Tech-Stack Used

- **Microsoft Excel 2022**: Excel was used to perform the data analysis and calculations. The software's vast array of built-in functions allowed me to calculate descriptive statistics, correlation coefficients, and percentiles with ease. I also used its charting capabilities to visualize the relationships between different variables.

INSIGHTS

1. **Movie Genre Analysis:**
   - Certain genres like Comedy, Drama,Action and Adventure tend to have higher average IMDB ratings compared to others like Horror or Drama.
   - Genres with more variability in IMDB scores often include those that mix multiple elements (e.g., Fantasy/Action).
   - The analysis suggests that some genres consistently perform well with audiences, leading to higher ratings, possibly due to broad appeal or well-established fan bases.

2. **Movie Duration Analysis:**
   - There is no clear correlation between movie duration and IMDB scores. While some longer movies may receive higher ratings due to rich storytelling, there is a mix of opinions on shorter movies, with some receiving high ratings despite their brief run-time.
   - The average duration of movies with high ratings can vary, indicating that the length of a movie does not always dictate its quality or viewer satisfaction.

3. **Language Analysis:**
   - English is the most common language in the dataset, with a large number of high-rated movies.
   - Non-English languages show a mixed result; while they may not have as many high ratings overall, certain languages like Spanish and Japanese often correlate with niche but high-quality movies (e.g., anime or art films).
   - The analysis shows that movies in English tend to dominate high ratings, but non-English films can often surprise with high-quality offerings and appeal to specific audiences.

4. **Director Analysis:**
   - Directors with a consistent history of producing high IMDB-rated movies (e.g., Christopher Nolan, Steven Spielberg) often have higher average ratings. These directors are in the top percentile of the dataset.
   - The influence of a director's style and reputation plays a significant role in a movie's overall success, showing that a skilled director can create a more immersive and engaging movie, often leading to higher audience ratings.

5. **Budget Analysis:**
   - There is a positive correlation between movie budget and gross earnings, indicating that larger budgets tend to lead to higher earnings. However, the profit margin varies significantly.
   - The movie with the highest profit margin in this dataset is Avatar, which made a massive profit compared to its budget. This confirms that despite its high budget, its financial success was extraordinary, with gross earnings far outpacing production costs.

Conclusion : Avatar stands out as the most profitable movie in this dataset, not only due to its high gross earnings but also because of its massive profit margin, making it a remarkable example of financial success in the film industry.

- The analysis highlights the importance of factors like genre, director reputation, and budget in determining a movie's IMDB rating and its financial success, with certain genres and directors consistently achieving high ratings.

Tasks

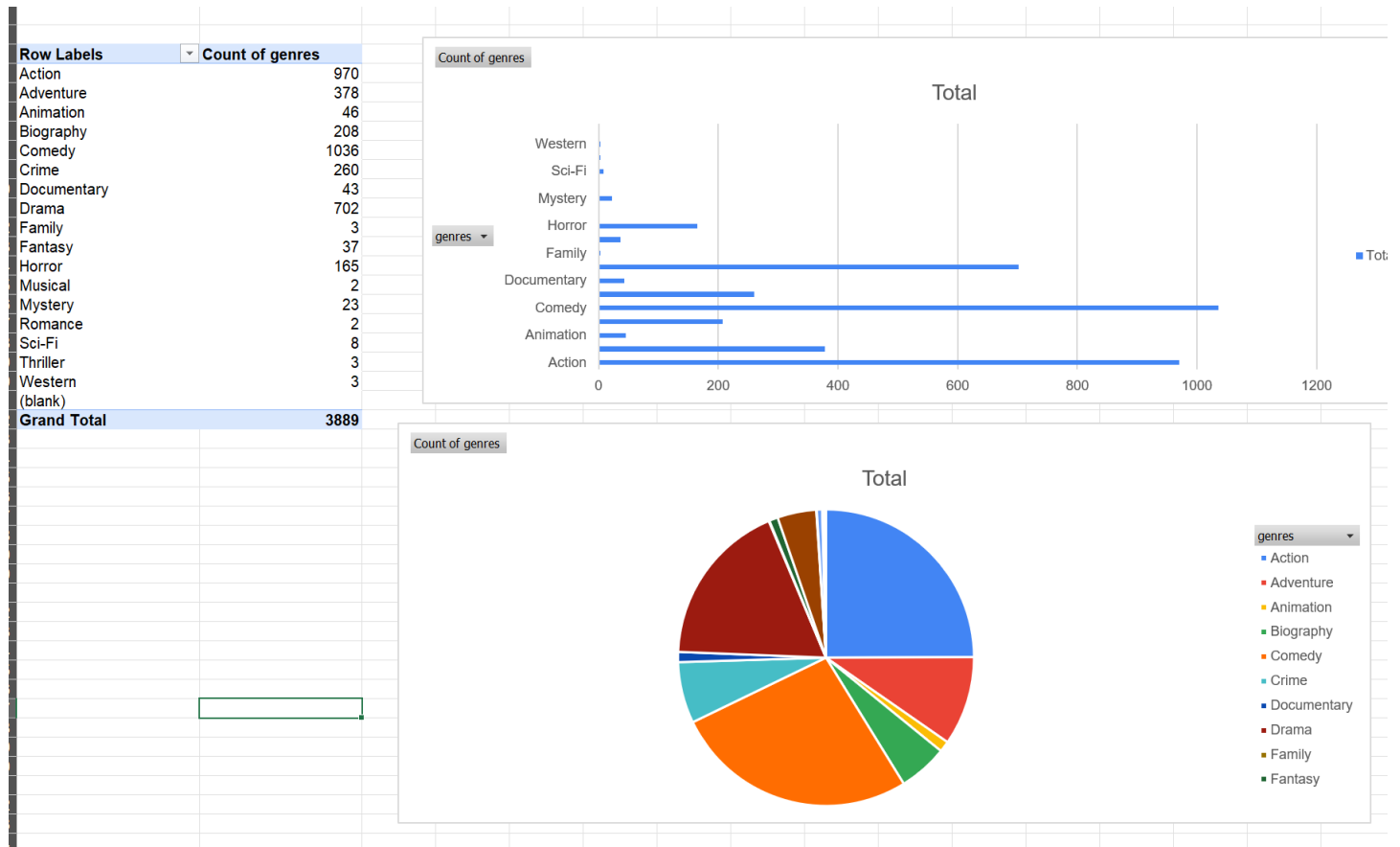**A. Movie Genre Analysis:** Analyze the distribution of movie genres and their impact on the IMDB score.

Step 1

Separated the genre column based on different genre.

movie_title	genres	imdb_scor	list of genres	Action	Adventure	Drama	Animation	Comedy	Mystery	Crime	Biography	Fantasy	Documentary	Sci-Fi	Horror	Romance
Avatar	Action|Adventure|Fantasy|Sci-Fi	7.9	Action	7.9	7.9	6.7	7.8	7.8	7.5	7.2	5.5	7.9	7.8	7.9	7	6.2
Pirates of the Caribbean:	Action|Adventure|Fantasy	7.1	Adventure	7.1	7.1	7.2	7.3	6.8	7.5	6.2	7.5	7.1	7.3	6.6	5.2	7.8
Spectre	Action|Adventure|Thriller	6.8	Drama	6.8	6.8	7.7	6.3	7.3	6.5	9	6.8	7.8	8	7.5	7.2	7.2
The Dark Knight Rises	Action|Thriller	8.5	Animation	8.5	6.6	7.3	8.3	6.3	7.5	6.7	7	7.5	7.1	6.9	6.2	7.7
John Carter	Action|Adventure|Sci-Fi	6.6	Comedy	6.6	6.2	6.8	7.2	8.3	7.6	6.1	7.8	7.3	1.6	6.1	5.8	7.3
Spider-Man 3	Action|Adventure|Romance	6.2	Mystery	6.2	7.8	6.6	8.4	7.2	6.7	6.6	8.2	7.2	5.9	7.2	5.7	7.3
Tangled	Adventure|Animation|Comedy|Family|Fantasy|Musical|Rom	7.8	Crime	7.5	7.5	9	6.8	6.2	6.1	6.6	7	6.6	4.1	8.1	6.6	6.6
Avengers: Age of Ultron	Action|Adventure|Sci-Fi	7.5	Biography	6.9	7.5	7.5	8.3	8.3	7.7	6.4	8	6.7	8	6.8	4.9	6.6
Harry Potter and the Half-	Adventure|Family|Fantasy|Mystery	7.5	Fantasy	6.1	6.9	8.3	6.5	6.5	8.1	7.3	8.2	6.8	5.4	8.2	5.7	7
Batman v Superman: Daw	Action|Adventure|Sci-Fi	6.9	Documentary	6.7	6.1	7.8	8.3	4.8	6.4	7.5	7.1	7.5	7.5	5.9	7	7.8
Superman Returns	Action|Adventure|Sci-Fi	6.1	Sci-Fi	7.3	6.7	6.1	6.4	6.9	7.4	4.8	8.4	7	6.5	7	4.9	6.1
Quantum of Solace	Action|Adventure	6.7	Horror	6.5	7.3	7.6	7.9	5.4	7.8	6.2	7	7.9	6.6	7.2	4.2	5.5
Pirates of the Caribbean:	Action|Adventure|Fantasy	7.3	Romance	7.2	6.5	6.3	7.8	8.3	6.6	6.4	7.9	6.1	6.9	6.8	5.1	6.6
The Lone Ranger	Action|Adventure|Western	6.5	Family	6.6	7.2	7.9	6.6	6.4	7.5	6.5	7.4	7.3	2.7	6	5.6	7.3
Man of Steel	Action|Adventure|Fantasy|Sci-Fi	7.2	Western	8.1	6.6	8.6	8.2	7.9	7	5.8	7.2	6.5	8.5	5.7	6.2	6.4
The Chronicles of Narnia:	Action|Adventure|Family|Fantasy	6.6	Musical	6.7	8.1	7.8	6.1	7.8	7	3.7	7.1	6.8	7	6.7	6.3	5.5
The Avengers	Action|Adventure|Sci-Fi	8.1	Thriller	6.8	6.7	5.6	8	6.1	5.8	6.5	6.6	7.3	7.6	6.8	5.2	5.5
Pirates of the Caribbean:	Action|Adventure|Fantasy	6.7	History	7	6.8	6.1	6.7	6.7	7.1	7	7	6.4	6.6	5.6	6.4	6.6
Men in Black 3	Action|Adventure|Comedy|Family|Fantasy|Sci-Fi	6.8	Music	6.7	7.5	5.5	7.6	8	7.7	7.8	6.6	6.7	8	6.6	5.9	5.3
The Hobbit: The Battle of	Adventure|Fantasy	7.5	War	7.2	7	6.4	6.9	6.7	7.4	6.4	7.4	8.3	5.4	7	5.5	6.3
The Amazing Spider-Man	Action|Adventure|Fantasy	7	Sport	8.2	6.7	7.2	5.1	5.9	6.1	6.7	6.7	8	7.2	8	5.9	6.5
Robin Hood	Action|Adventure|Drama|History	6.7	Short	5.9	7.9	6.9	6.2	7.6	6.9	8.2	7.3	6.3	6.6	7.8	5.4	3.7
The Hobbit: The Desolatic	Adventure|Fantasy	7.9	Film-Noir	7	6.1	5.8	7.3	6.9	5.7	3.3	7.9	6.6	6.4	7	6.8	7.1
The Golden Compass	Adventure|Family|Fantasy	6.1		7.8	7.2	6.6	5.4	5.1	7.3	5.3	8	6.2	5.1	6.3	6.3	6.9
King Kong	Action|Adventure|Drama|Romance	7.2		7.3	8.2	6.4	6.7	6.2	6.3	7.5	6.3	7.2	6.7	7.5	5.7	4.4
Titanic	Drama|Romance	7.7		7.2	5.9	7.7	6.9	6.2	7.3	6.6	6.9	6.8	7.7	8.4	5.7	8.5
Captain America: Civil Wa	Action|Adventure|Sci-Fi	8.2		6.8	7	6.1	6.9	7.3	6.6	5.5	7	6.9	7.6	5.8	6.7	7
Battleship	Action|Adventure|Sci-Fi|Thriller	5.9		6	7.8	7.7	6.3	5.4	7.8	8.5	8	5.2	7.1	5.4	4.8	5.5
Jurassic World	Action|Adventure|Sci-Fi|Thriller	7		5.7	7.3	7.4	7.2	6.7	7.6	7.8	7.4	5.4	8	7.3	6	6.9
Skyfall	Action|Adventure|Thriller	7.8		6.7	7.2	5.5	7.3	6.9	4.9	7.6	8.2	8.3	8.3	6.5	5.6	6
Spider-Man 2	Action|Adventure|Fantasy|Romance	7.3		6.8	6.5	8.1	7.9	6.9	7.4	6.4	6.7	7.8	8.1	7.9	6.4	3.3
Iron Man 3	Action|Adventure|Sci-Fi	7.2		5.6	6.8	5.5	6.4	5.5	6.4	6.9	7.3	6.1	6.8	4.8	5.9	7.6
Alice in Wonderland	Adventure|Family|Fantasy	6.5		6.6	7.3	5.8	5.9	6.1	6.9	7	7.6	7	7.8	6.9	6.4	4.3
X-Men: The Last Stand	Action|Adventure|Fantasy|Sci-Fi|Thriller	6.8		7.2	6	6.6	6.5	7.2	7.8	6.1	6.7	6.3	6.6	7.9	6.4	6.9
Monsters University	Adventure|Animation|Comedy|Family|Fantasy	7.3		7	5.7	6.4	7.2	6.1	6.8	6.7	6.4	6.4	7.4	6.1	6.8	7
Transformers: Revenge o	Action|Adventure|Sci-Fi	6		8	6.4	8	6.7	7.9	7.3	7.8	6.8	6.6	6.2	5.8	6.7	6.3
Transformers: Age of Exti	Action|Adventure|Sci-Fi	5.7		7.8	6.7	6.6	7.6	6	6.7	6.8	6	8.2	7.4	7	7.6	7.3
Oz the Great and Powerfi	Adventure|Family|Fantasy	6.4		6.6	6.8	5.3	6.7	5.9	4.9	6.2	6.4	7.9	7.7	7.6	5.9	6.2
The Amazing Spider-Man	Action|Adventure|Fantasy|Sci-Fi	6.7		7	6.3	6	6.7	6.5	7.7	7.3	7.3	7.8	8.2	7.8	7.6	7.7
TRON: Legacy	Action|Adventure|Sci-Fi	6.8		6.3	5.6	6.1	6.5	7.2	5.9	6.9	7.1	6.1	7.7	6.5	6.1	6.5
Cars 2	Adventure|Animation|Comedy|Family|Sport	6.3		6.2	8.3	7.8	7.3	6.7	8.1	6	7.7	6.4	7.3	7.9	6.6	5.8
Green Lantern	Action|Adventure|Sci-Fi	5.6		7.5	6.6	6.1	6.4	7.6	6.6	7.7	7	6.1	7.6	7.8	6.3	7.1
Toy Story 3	Adventure|Animation|Comedy|Family|Fantasy	8.3		6.2	7	7.1	5.4	7.4	4.9	6.7	7.6	7.3	6.8	5.5	6.2	6.9

Step 2

Used pivot table function for further calculations

Row Labels	Count of genres
Action	970
Adventure	378
Animation	46
Biography	208
Comedy	1036
Crime	260
Documentary	43
Drama	702
Family	3
Fantasy	37
Horror	165
Musical	2
Mystery	23
Romance	2
Sci-Fi	8
Thriller	3
Western	3
(blank)	
Grand Total	3889

Calculated descriptive statistics for imdb score for each genre using excel formulas:

DES STAT (IMDB RATING)	Action	Adventure	Drama	Animation	Comedy	Mystery	Crime	Biography	Fantasy	Documentary	Sci-Fi	Horror	Romance	Family	Western	Musical	Thriller	History	Music	War	Sport	Short	Film-Noir
AVG	6.290619	6.455723	6.781693	6.700503	6.184514	6.468638	6.545556	7.141803	6.281431	7.01194	6.323952	5.92796	6.430552	6.209934	6.740323	6.559223	6.378388	7.125641	6.4636	7.048148	6.607237	6.8	7.7
MEDIAN	6.35	6.6	6.9	6.8	6.3	6.5	6.6	7.2	6.4	7.2	6.4	6	6.5	6.3	6.75	6.7	6.4	7.2	6.7	7.1	6.8	6.8	7.7
MODE	6.6	6.7	6.7	6.7	6.7	6.6	6.6	7	6.7	6.6	6.7	5.9	6.5	5.4	6	7.1	6.5	7.7	6.2	7.1	7.2	#N/A	#N/A
STD DEV	1.037485	1.108754	0.8973	0.986109	1.040426	1.017403	0.982673	0.706102	1.134906	1.190951	1.15787	0.997124	0.964045	1.159151	0.971938	1.135434	0.970135	0.681333	1.181725	0.804787	1.03759	0.3	0
MAX	9	8.9	9.3	8.6	8.8	8.6	9.3	8.9	8.9	8.5	8.8	8.6	8.5	8.6	8.9	8.5	9	8.9	8.5	8.6	8.4	7.1	7.7
MIN	2.1	2.3	2.1	2.8	1.9	3.1	2.4	4.5	2.2	1.6	1.9	2.3	2.1	1.9	4.1	2.1	2.7	5.5	1.6	4.3	2	6.5	7.7
VAR	1.076376	1.229335	0.805147	0.972412	1.082487	1.035109	0.965647	0.49858	1.288011	1.418365	1.340664	0.994256	0.929382	1.343632	0.944664	1.289211	0.941163	0.464214	1.396475	0.647682	1.076592	0.09	0

Chart Title

Legend: AVG, MEDIAN, MODE, STD DEV, MAX, MIN, VAR

Categories: Action, Adventure, Drama, Animation, Comedy, Mystery, Crime, Biography, Fantasy, Documentary, Sci-Fi, Horror, Romance, Family, Western, Musical, Thriller, History, Music, War, Sport, Short, Film-Noir
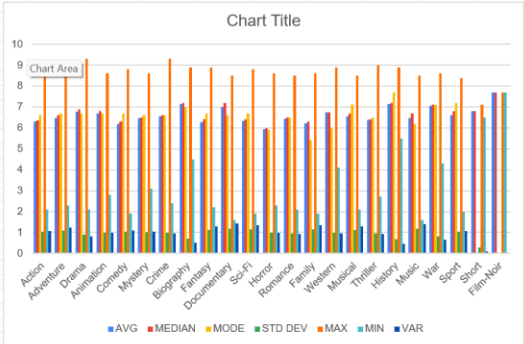
B. Movie Duration Analysis: Analyze the distribution of movie durations and its impact on the IMDB score.

imdb_score	duration			
7.9	178			
7.1	169	avg	109.8982	
6.8	148	median	106	
8.5	164	std dev	22.70651	
6.6	132	mode	101	
6.2	156			
7.8	100			
7.5	141			
7.5	153			
6.9	183			
6.1	169			
6.7	106			
7.3	151			
6.5	150			
7.2	143			
6.6	150			
8.1	173			
6.7	136			
6.8	106			
7.5	164			
7	153			
6.7	156			
7.9	186			
6.1	113			
7.2	201			
7.7	194			
8.2	147			
5.9	131			
7	124			
7.8	143			
7.3	135			
7.2	195			
6.5	108			
6.8	104			
7.3	104			
6	150			
5.7	165			
6.4	130			
6.7	142			
6.8	125			
6.3	106			

RATING VS duration

Insights:

The trendline indicates a slight positive correlation between movie duration and IMDB score. This suggests that, generally, longer movies tend to have higher ratings. However, the correlation appears weak because the data points are widely scattered around the trendline.

There is significant variability in ratings for movies of similar durations. For instance: Some shorter movies (e.g., under 100 minutes) have both low and high ratings. Longer movies (e.g., 150+ minutes) also show a wide range of ratings.
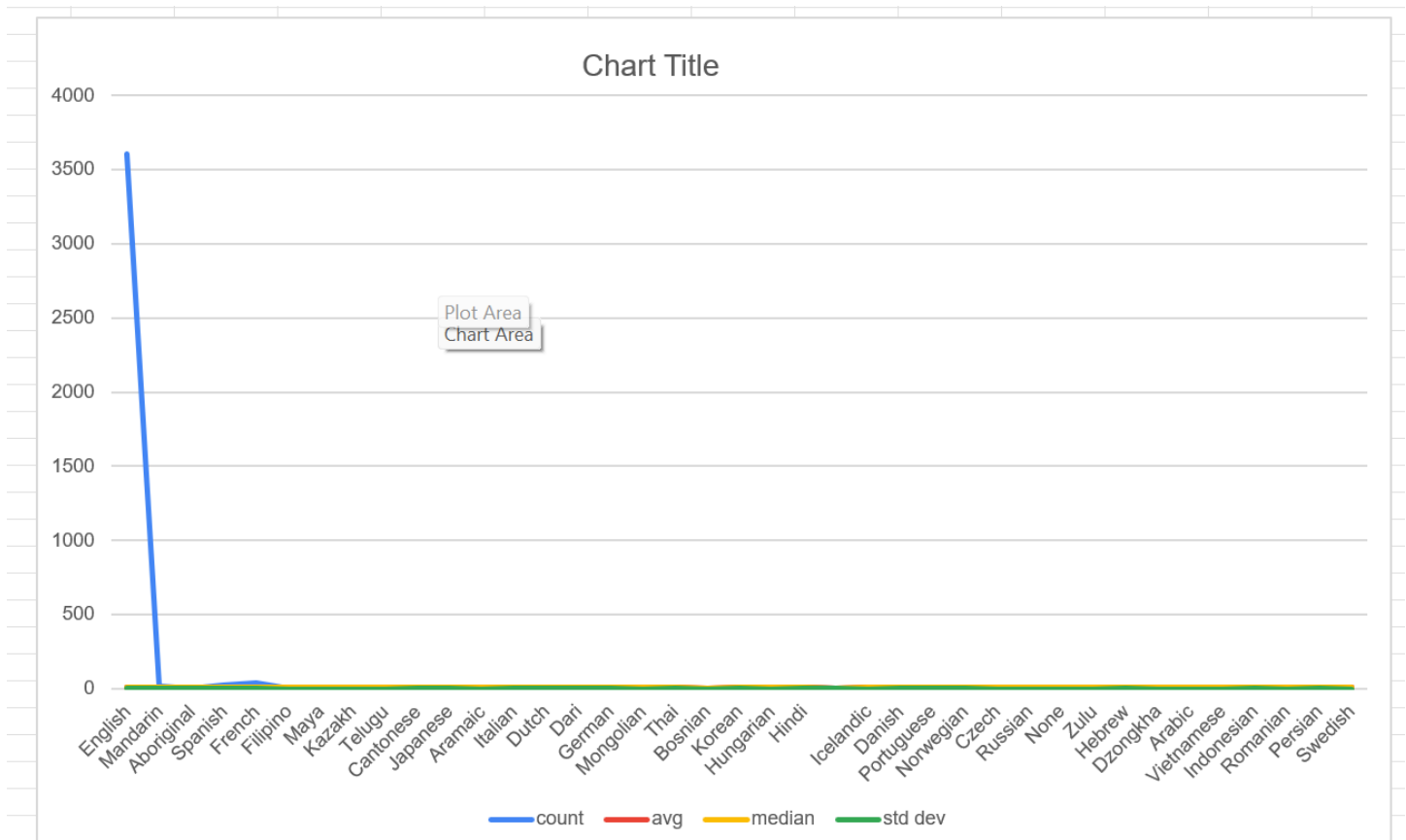
There are a few data points that deviate significantly from the trend (e.g., movies with high durations and low ratings or vice versa). These could represent unique cases like critically panned long movies or highly-rated short ones.

While longer movies might generally perform better in ratings, duration alone does not strongly determine the IMDB score. Other factors, such as genre, storyline, direction, and acting, likely play a larger role.

C. Language Analysis: Situation: Examine the distribution of movies based on their language.

language	imdb_score
English	7.9
English	7.1
English	6.8
English	8.5
English	6.6
English	6.2
English	7.8
English	7.5
English	7.5
English	6.9
English	6.1
English	6.7
English	7.3
English	6.5
English	7.2
English	6.6
English	8.1
English	6.7
English	6.8
English	7.5
English	7
English	6.7
English	7.9
English	6.1
English	7.2
English	7.7
English	8.2
English	5.9
English	7
English	7.8
English	7.3
English	7.2
English	6.5
English	6.8
English	7.3
English	6
English	5.7
English	6.4
English	6.7
English	6.8
English	6.3
English	5.6

language	count	avg	median	std dev
English	3608	6.421258	6.5	1.052352
Mandarin	14	7.021429	7.25	0.73793
Aboriginal	2	6.95	6.95	0.55
Spanish	26	7.05	7.15	0.810152
French	37	7.286486	7.2	0.553691
Filipino	1	6.7	6.7	0
Maya	1	7.8	7.8	0
Kazakh	1	6	6	0
Telugu	1	8.4	8.4	0
Cantonese	8	7.2375	7.3	0.412121
Japanese	12	7.625	7.8	0.861322
Aramaic	1	7.1	7.1	0
Italian	7	7.185714	7	1.069618
Dutch	3	7.566667	7.8	0.329983
Dari	2	7.5	7.5	0.1
German	13	7.692308	7.7	0.615769
Mongolian	1	7.3	7.3	0
Thai	3	6.633333	6.6	0.368179
Bosnian	1	4.3	4.3	0
Korean	5	7.7	7.7	0.509902
Hungarian	1	7.1	7.1	0
Hindi	10	6.76	7.05	1.054704
	0	5.933333	5.8	0.984322
Icelandic	1	6.9	6.9	0
Danish	3	7.9	8.1	0.432049
Portuguese	5	7.76	8	0.875443
Norwegian	4	7.15	7.3	0.497494
Czech	1	7.4	7.4	0
Russian	1	6.5	6.5	0
None	1	8.5	8.5	0
Zulu	1	7.3	7.3	0
Hebrew	3	7.5	7.3	0.355903
Dzongkha	1	7.5	7.5	0
Arabic	1	7.2	7.2	0
Vietnames	1	7.4	7.4	0
Indonesian	2	7.9	7.9	0.3
Romanian	1	7.9	7.9	0
Persian	3	8.133333	8.4	0.449691
Swedish	1	7.6	7.6	0

Chart Title

Derived unique languages using data< removes duplicates then used countif function

=COUNTIF(language_range, "English")

Descriptive Statistics for IMDB Ratings by Language:

For each language, calculate the mean, median, and standard deviation of the IMDB scores using :

    Mean: =AVERAGE(range_of_IMDB_scores)
    Median: =MEDIAN(range_of_IMDB_scores)
    Standard Deviation: =STDEV.P(range_of_IMDB_scores)

Insights:

English Dominates: 3608 movies with a moderate average IMDb score of 6.42. High variability (std dev: 1.05) reflects diverse quality.

Top Performers: Telugu: Highest average (8.4), but only 1 movie. Persian: Strong average (8.13) and consistent quality (std dev: 0.45). Danish, Portuguese, and Indonesian: High scores (7.76–7.9) with low variability.

Low Performers: Bosnian: Lowest rating (4.3). Thai, Kazakh, and Russian: Scores below 6.7.

Consistent Languages: French, Japanese, and German movies have high scores with low variability, indicating steady quality.

 Conclusion: English leads in volume, but smaller languages like Persian and Telugu excel in quality.

D. Director Analysis: Influence of directors on movie ratings.

Average IMDB Score by Director

director_name	imdb_score		director_name	count	avg
James Cameron	7.9		James Cameron	7	7.914286
Gore Verbinski	7.1		Gore Verbinski	7	6.985714
Sam Mendes	6.8		Sam Mendes	8	7.5
Christopher Nolan	8.5		Christopher Nolan	8	8.425
Andrew Stanton	6.6		Andrew Stanton	3	7.733333
Sam Raimi	6.2		Sam Raimi	12	6.85
Nathan Greno	7.8		Nathan Greno	1	7.8
Joss Whedon	7.5		Joss Whedon	4	7.925
David Yates	7.5		David Yates	4	7.05
Zack Snyder	6.9		Zack Snyder	8	7.175
Bryan Singer	6.1		Bryan Singer	8	7.2875
Marc Forster	6.7		Marc Forster	7	7.228571
Gore Verbinski	7.3		Andrew Adamson	4	7.15
Gore Verbinski	6.5		Rob Marshall	5	6.6
Zack Snyder	7.2		Barry Sonnenfeld	7	6.457143
Andrew Adamson	6.6		Peter Jackson	12	7.675
Joss Whedon	8.1		Marc Webb	3	7.133333
Rob Marshall	6.7		Ridley Scott	17	7.070588
Barry Sonnenfeld	6.8		Chris Weitz	5	6.08
Peter Jackson	7.5		Anthony Russo	4	7
Marc Webb	7		Peter Berg	6	6.666667
Ridley Scott	6.7		Colin Trevorrow	2	7
Peter Jackson	7.9		Shane Black	2	7.4
Chris Weitz	6.1		Tim Burton	16	6.93125
Peter Jackson	7.2		Brett Ratner	10	6.41
James Cameron	7.7		Dan Scanlon	1	7.3
Anthony Russo	8.2		Michael Bay	13	6.638462
Peter Berg	5.9		Joseph Kosinski	3	6.866667
Colin Trevorrow	7		John Lasseter	5	7.38
Sam Mendes	7.8		Martin Campbell	9	6.711111
Sam Raimi	7.3		Lee Unkrich	1	8.3
Shane Black	7.2		McG	6	6.083333
Tim Burton	6.5		James Wan	7	7.2
Brett Ratner	6.8		J.J. Abrams	4	7.45
Dan Scanlon	7.3		Baz Luhrmann	5	7.12
Michael Bay	6		Mike Newell	7	6.842857
Michael Bay	5.7		Guillermo del Toro	5	7.14
Sam Raimi	6.4		Steven Spielberg	25	7.544
Marc Webb	6.7		Mark Andrews	1	7.2
Joseph Kosinski	6.8		Justin Lin	5	6.84
John Lasseter	6.3		Roland Emmerich	8	6.1875
Martin Campbell	5.6		Robert Zemeckis	13	7.307692
Lee Unkrich	8.3		Lana Wachowski	6	6.916667

Identify Top Directors:

- Use the PERCENTILE function to calculate the percentile ranks of the directors' average IMDB scores:

=PERCENTILE(H:H, 0.9) [7.5]

	director_name	avg
1		
5	Christopher Nolan	8.425
561	S.S. Rajamouli	8.4
627	Richard Marquand	8.4
694	Sergio Leone	8.433333
859	Alfred Hitchcock	8.5
1191	Tony Kaye	8.6
1441	Ron Fricke	8.5
1474	Damien Chazelle	8.5
1586	Charles Chaplin	8.6
1656	Asghar Farhadi	8.4
1682	Marius A. Markevicius	8.4
1727	Majid Majidi	8.5

TOP 10

- **Christopher Nolan (8.425 AVG)** and **Sergio Leone (8.433 AVG)** both have multiple movies and a consistently high rating.
- Other directors such as **Damien Chazelle** and **Richard Marquand** also exhibit high scores with fewer movies but strong averages.

E. Budget Analysis: Explore the relationship between movie budgets and their financial success.
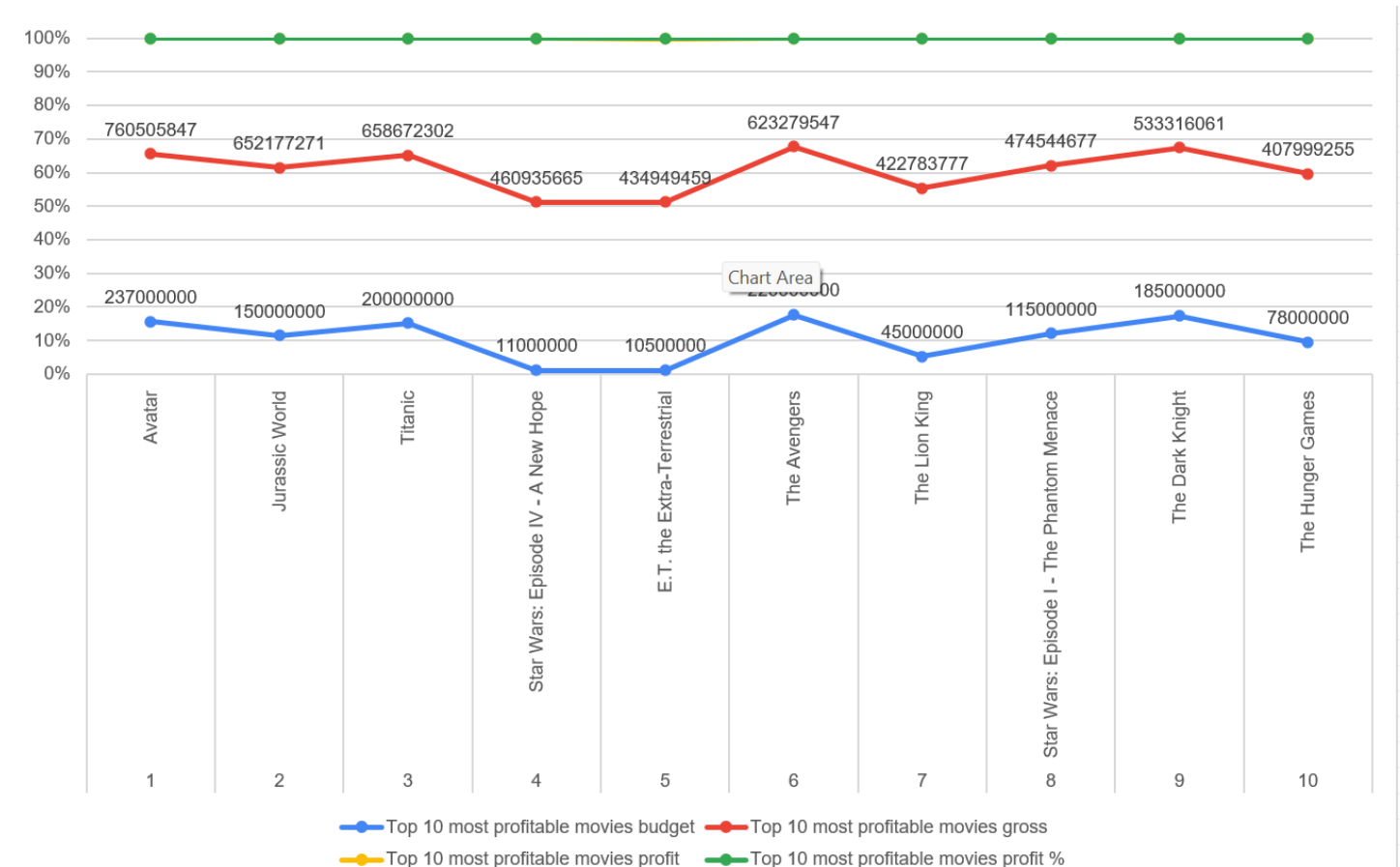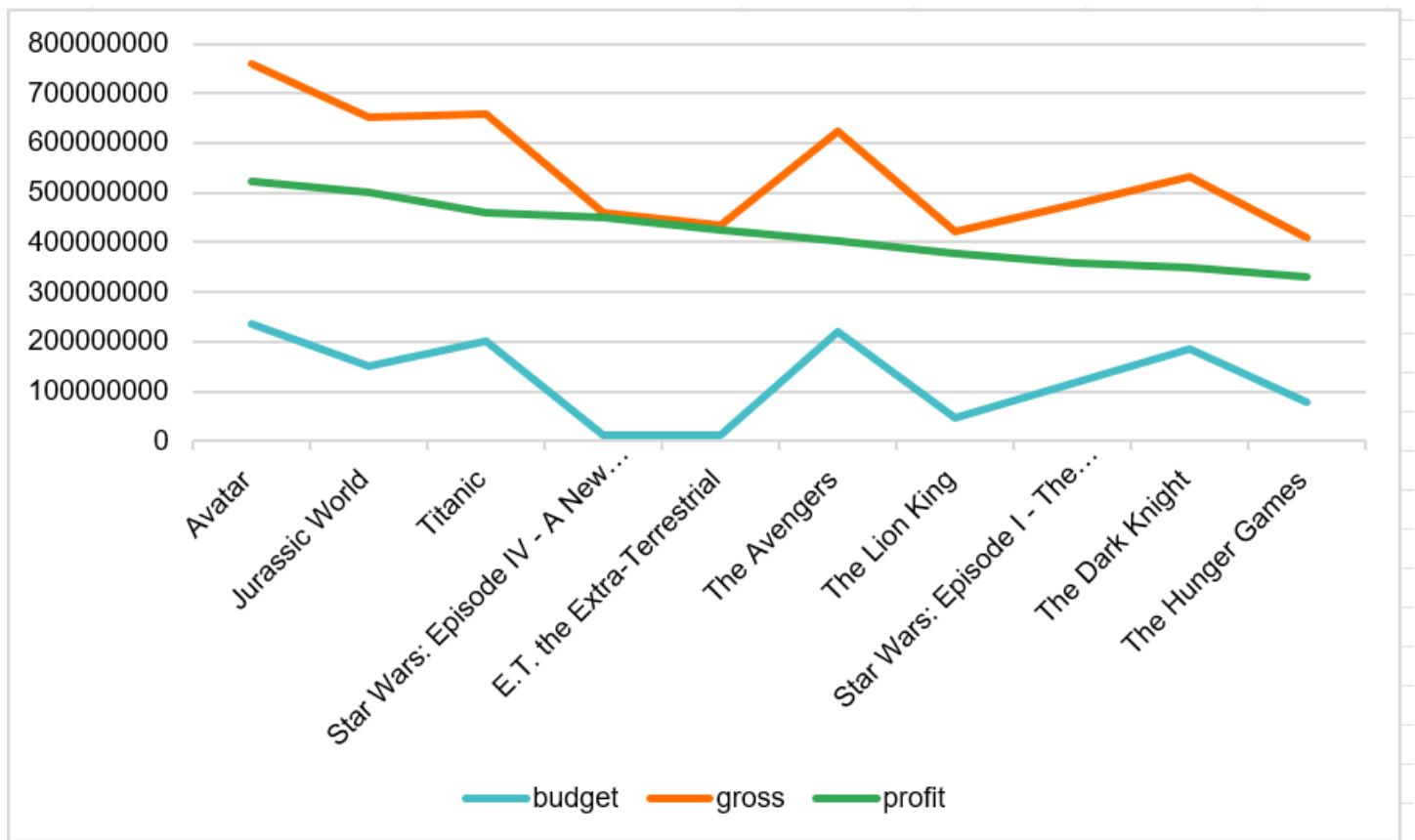
Created a profit column

Profit = Gross Earnings - Budget

movie_title	budget	gross	profit	Profit %
Avatar	237000000	760505847	523505847	220.88854
Jurassic World	150000000	652177271	502177271	334.78485
Titanic	200000000	658672302	458672302	229.33615
Star Wars: Episode IV - A New Hope	11000000	460935665	449935665	4090.3242
E.T. the Extra-Terrestrial	10500000	434949459	424449459	4042.3758
The Avengers	220000000	623279547	403279547	183.30889
The Lion King	45000000	422783777	377783777	839.5195
Star Wars: Episode I - The Phantom Menace	115000000	474544677	359544677	312.64755
The Dark Knight	185000000	533316061	348316061	188.27895
The Hunger Games	78000000	407999255	329999255	423.07597
Deadpool	58000000	363024263	305024263	525.9039
The Hunger Games: Catching Fire	130000000	424645577	294645577	226.65044
Jurassic Park	63000000	356784000	293784000	466.32381
Despicable Me 2	76000000	368049635	292049635	384.27584
American Sniper	58800000	350123553	291323553	495.44822
Finding Nemo	94000000	380838870	286838870	305.14773
Shrek 2	150000000	436471036	286471036	190.98069
The Lord of the Rings: The Return of the King	94000000	377019252	283019252	301.08431
Star Wars: Episode VI - Return of the Jedi	32500000	309125409	276625409	851.1551
Forrest Gump	55000000	329691196	274691196	499.43854
Star Wars: Episode V - The Empire Strikes Back	18000000	290158751	272158751	1511.9931
Home Alone	18000000	285761243	267761243	1487.5625
Star Wars: Episode III - Revenge of the Sith	113000000	380262555	267262555	236.51554
Spider-Man	139000000	403706375	264706375	190.43624
Minions	74000000	336029560	262029560	354.094
The Sixth Sense	40000000	293501675	253501675	633.75419
Jaws	8000000	260000000	252000000	3150
Frozen	150000000	400736600	250736600	167.15773
The Secret Life of Pets	75000000	323505540	248505540	331.34072
The Twilight Saga: New Moon	50000000	296623634	246623634	493.24727
The Lord of the Rings: The Two Towers	94000000	340478898	246478898	262.21159
The Hangover	35000000	277313371	242313371	692.32392
My Big Fat Greek Wedding	5000000	241437427	236437427	4728.7485
The Twilight Saga: Eclipse	68000000	300523113	232523113	341.94575
Independence Day	75000000	306124059	231124059	308.16541
The Blind Side	29000000	255950375	226950375	782.5875
Raiders of the Lost Ark	18000000	242374454	224374454	1246.5247
The Lord of the Rings: The Fellowship of the Ring	93000000	313837577	220837577	237.45976
Beverly Hills Cop	14000000	234760500	220760500	1576.8607
Batman	35000000	251188924	216188924	617.68264
Toy Story 3	200000000	414984497	214984497	107.49225
The Hunger Games: Mockingjay - Part 1	125000000	337103873	212103873	169.6831

	Top 10 highest grossing movies collection amount wise				
Rank	movie_title	budget	gross	profit	profit %
1	Avatar	237000000	760505847	523505847	220.89
2	Titanic	200000000	658672302	458672302	229.34
3	Jurassic World	150000000	652177271	502177271	334.78
4	The Avengers	220000000	623279547	403279547	183.31
5	The Dark Knight	185000000	533316061	348316061	188.28
6	Star Wars Chart Area - The Phantom Menace	115000000	474544677	359544677	312.65
7	Star Wars: Episode IV - A New Hope	11000000	460935665	449935665	4090.32
8	Avengers: Age of Ultron	250000000	458991599	208991599	83.60
9	The Dark Knight Rises	250000000	448130642	198130642	79.25
10	Shrek 2	150000000	436471036	286471036	190.98

	Top 10 most profitable movies				
Rank	movie_title	budget	gross	profit	profit %
1	Avatar	237000000	760505847	523505847	220.89
2	Jurassic World	150000000	652177271	502177271	334.78
3	Titanic	200000000	658672302	458672302	229.34
4	Star Wars: Episode IV - A New Hope	11000000	460935665	449935665	4090.3
5	E.T. the Extra-Terrestrial	10500000	434949459	424449459	4042.4
6	The Avengers	220000000	623279547	403279547	183.31
7	The Lion King	45000000	422783777	377783777	839.52
8	Star Wars: Episode I - The Phantom Menace	115000000	474544677	359544677	312.65
9	The Dark Knight	185000000	533316061	348316061	188.28
10	The Hunger Games	78000000	407999255	329999255	423.08

While Avatar is the top most profitable movie in the list, other high-budget films like Jurassic World and Titanic also performed exceedingly well, proving that strategic investments in popular genres, established franchises, and strong global appeal can yield immense profits. Additionally, the significant profits from lower-budget films like E.T. the Extra-Terrestrial highlight that profitability can also be achieved through effective storytelling and audience engagement.

Chart 1: Line chart showing budget, gross, and profit for top 10 movies.

X-axis categories: Avatar, Jurassic World, Titanic, Star Wars: Episode IV - A New..., E.T. the Extra-Terrestrial, The Avengers, The Lion King, Star Wars: Episode I - The..., The Dark Knight, The Hunger Games

Y-axis: 0 to 800000000

Legend: budget, gross, profit

Chart 2: Line chart with data labels.

	Avatar	Jurassic World	Titanic	Star Wars: Episode IV - A New Hope	E.T. the Extra-Terrestrial	The Avengers	The Lion King	Star Wars: Episode I - The Phantom Menace	The Dark Knight	The Hunger Games
gross	760505847	652177271	658672302	460935665	434949459	623279547	422783777	474544677	533316061	407999255
budget	237000000	150000000	200000000	11000000	10500000	220000000	45000000	115000000	185000000	78000000

X-axis index: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Legend: Top 10 most profitable movies budget, Top 10 most profitable movies gross, Top 10 most profitable movies profit, Top 10 most profitable movies profit %

LINKS

EXCEL WORKBOOK-
https://docs.google.com/spreadsheets/d/13fOFz_yR8QMEOOtAWILRf3Axpr3L2LFp/edit?usp=drive_link&ouid=108286913145936487778&rtpof=true&sd=true

LOOM VIDEO-

https://www.loom.com/share/b04074b8550b463088d1c7e407c174cb?sid=0ee4d9ea-c4a8-4e92-93a4-121e47f12db3