TMDB Movies Q&A

Introduction:

I have chosen a TMDB Movie dataset to analyze, and through this analysis we will try to answer some questions which are Which year has the highest and Lowest release of movies, Which genres are most popular from year to year, What is the lowest and highest movie Profit, What is the High and Low Movie runtime, What is the best movie for each year, who is the most director their movies get the highest and lowest rate, Which Month Released Highest Number of Movies, Which production companies Released Highest Number of Movies , and I will use some Python libraries like Pandas, NumPy and matplotlib to help me to analysis and answer these questions.

Data Wrangling:

In this part of the report, I will explain how the data was dealt with to clean it. At first, the columns that were not important in the analysis were removed, namely imdb_id, homepage, tagline, keywords, overview, budget_adj, revenue_adj, and then the missing rows that would affect the analysis process were removed from a column director and then the missing data in the other columns was dealt with by placing 0 in some columns like revenue and budget and Nan in others to ensure the completeness of the data to start the analysis, In addition to, I add a new column named Profit to hold the profit for each Movie because it will help to answer some question in analysis process.

Exploration Phase:

In this section, we will clarify the questions that we want to answer, as well as the answer for each question, and the penalty for the code that shows how this answer was obtained will be found in the code file.

- ## *Question 1:*

  - ➢ **Which year has the highest and Lowest release of movies?**
    - o **2014 the highest with 695 Movie.**
    - o **1961 and 1969 the lowest movie in year with 31 Movie.**

- ## *Question 2:*

  - ➢ **Which genres are most popular from year to year?**
    - o **1960   Drama**
    - o **1961   Drama**
    - o **1962   Drama**
    - o **1963   Thriller**
    - o **1964   Comedy**
    - o **1965   Drama**
    - o **1966   Drama**
    - o **1967   Drama**
    - o **1968   Drama**
    - o **1969   Drama**
    - o **1970   Drama**
    - o **1971   Drama**
    - o **1972   Drama**
    - o **1973   Drama**
    - o **1974   Drama**
    - o **1975   Drama**
    - o **1976   Drama**
    - o **1977   Adventure**
    - o **1978   Drama**
    - o **1979   Drama**
    - o **1980   Action**
    - o **1981   Adventure**
    - o **1982   Drama**
    - o **1983   Action**
    - o **1984   Action**
    - o **1985   Comedy**
    - o **1986   Drama**
    - o **1987   Comedy**
    - o **1988   Comedy**
    - o **1989   Comedy**
    - o **1990   Drama**
    - o **1991   Comedy**
    - o **1992   Comedy**
    - o **1993   Drama**

- 1994　　Drama
- 1995　　Drama
- 1996　　Drama
- 1997　　Thriller
- 1998　　Drama
- 1999　　Drama
- 2000　　Comedy
- 2001　　Drama
- 2002　　Drama
- 2003　　Action
- 2004　　Comedy
- 2005　　Drama
- 2006　　Drama
- 2007　　Drama
- 2008　　Drama
- 2009　　Drama
- 2010　　Drama
- 2011　　Drama
- 2012　　Drama
- 2013　　Drama
- 2014　　Drama
- 2015　　Drama

Question 3:

> **What is the lowest and highest movie Profit?**
> - **Avatar director by James Cameron release in 2009 become the highest movie Profit with (2544505847$).**
> - **The Warrior's Way director by Sngmoo Lee release in 2010 become the lowest movie Profit with (-413912431$).**

Question 4:

> **What is the High and Low Movie runtime?**
> - **The Story of Film: An Odyssey director by Mark Cousins release in 2011 become the highest movie runtime with 900 Minute.**
> - **Fresh Guacamole - Cousin Ben Troop Screening - Luxo Jr. - Bambi Meets Godzilla become the lowest movies runtime with 2 Minute.**

Question 5:

> **What is the best movie for each year?**

- 1960　8.0
- 1961　7.5
- 1962　7.7
- 1963　7.5
- 1964　7.7
- 1965　7.6
- 1966　7.9
- 1967　7.6
- 1968　7.7
- 1969　7.3
- 1970　7.6
- 1971　7.7
- 1972　8.3
- 1973　7.7
- 1974　8.1
- 1975　8.0
- 1976　7.8
- 1977　7.9
- 1978　8.0
- 1979　7.8
- 1980　8.0
- 1981　8.5
- 1982　7.9
- 1983　8.1
- 1984　8.4
- 1985　8.3
- 1986　8.0
- 1987　7.7
- 1988　8.0
- 1989　8.4
- 1990　8.0
- 1991　7.9
- 1992　8.4
- 1993　8.1
- 1994　8.4
- 1995　8.5
- 1996　7.9
- 1997　7.8
- 1998　8.0
- 1999　8.1
- 2000　8.5
- 2001　8.0
- 2002　8.3

- o **2003   8.0**
- o **2004   7.8**
- o **2005   8.3**
- o **2006   8.7**
- o **2007   7.9**
- o **2008   8.3**
- o **2009   8.3**
- o **2010   8.8**
- o **2011   9.2**
- o **2012   8.3**
- o **2013   8.5**
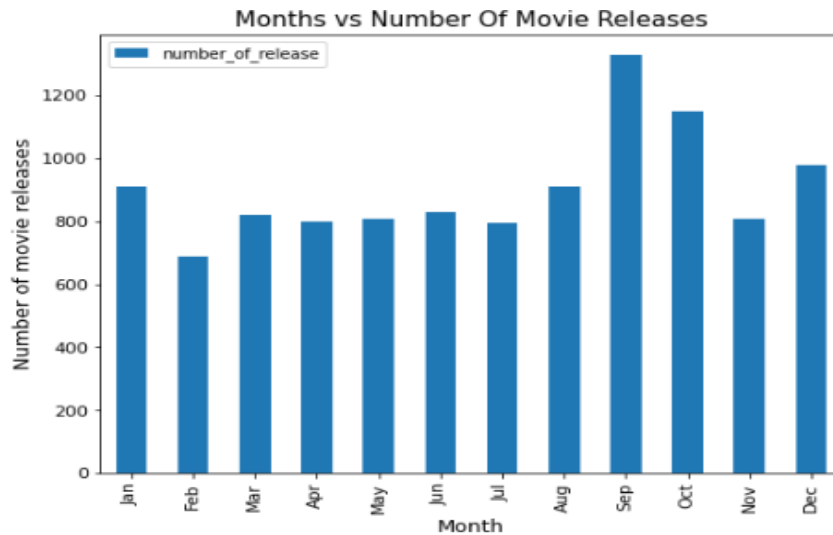- o **2014   8.8**
- o **2015   8.9**

- *Question 6:*

  - ➤ **who is the most director their movies get the highest and lowest rate?**
    - o **Mark Cousins director of The Story of Film: An Odyssey become the highest rate with 9.2.**
    - o **Leigh Scott director of Transmorphers and Harold P. Warren director of Manos: The Hands of Fate become the lowest rate with 1.5.**

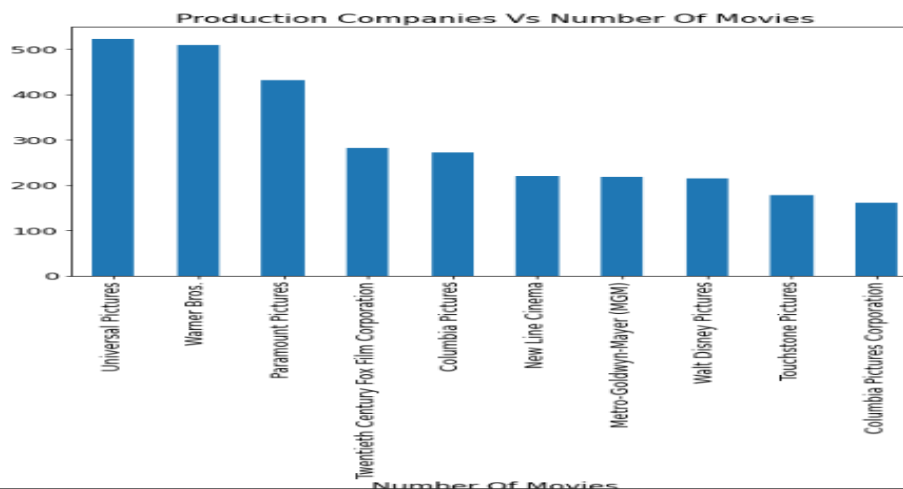## The project's visualizations:

- *Question 1:*

  - ➤ **Which Month Released Highest Number of Movies?**

Months vs Number Of Movie Releases

o **According to the chart, higher number of releases in September and October month.**

- ## *Question 2:*

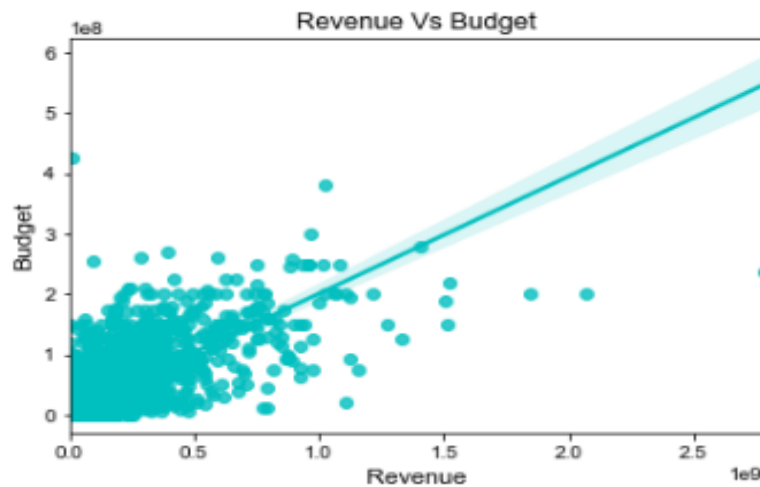 ➤ **Which production companies Released Highest Number of Movies?**

Production Companies Vs Number Of Movies

o **According to the chart, higher number of releases by Universal Studios and Warner Bros.**

- ## *Question 3:*

  ➤ **How Does the Revenue differ Budget?**

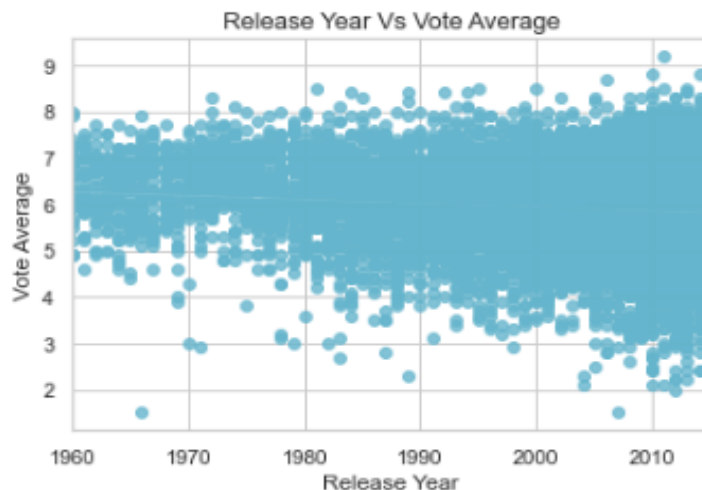Correlation Between Revenue And Budget : 0.6885179716225603

Revenue Vs Budget

○ **Budget vs Revenue: Budget and revenue both have positive correlation (0.68) between them. Means there is a good possibility that movies with higher investments result in better revenues.**

● *Question 4:*

➢ **How Does the Release Year differ vote average?**

Correlation Between Release Year And Vote Average : -0.12056518481297653

Release Year Vs Vote Average

○ **Release Year Vs Vote Average: Release year and vote Average have negative correlation (-0.11). Means that movie ratings (vote average) do not depends on the release year.**

Conclusions:
- **in this phase I will try to describe all project.**

- firstly, I have answer all question I describe in above phases by analyzing data, so I find 2014 become the most year have movies released by 695 and the 1961 and 1969 become the lowest year have movies released by 31 movie, and the drama movies is the most movies released along most years, also I have the lowest and highest profit movies, in addition to, the strangest things during analysis I have the movie with 900 min runtime and it is the highest movie rating between all movies with 9.2 rate for director Mark Cousins
- secondly, by using some vis charts I have seen the most months have released movies is September and October this maybe led to because it is the summer and off time in most country, also the Universal Studios and Warner Bros is the most production companies released movies, in addition to Budget and revenue both have positive correlation (0.68) between them. Means there is a good possibility that movies with higher investments result in better revenues, Finally, Release year and vote Average have negative correlation (-0.11). Means that movie ratings (vote average) do not depends on the release year.