1. Introduction and Objective

The project aimed to explore the potential relationships between the ratings, budget, and box office earnings of Marvel movies and understand whether these factors can serve as predictive indicators of a film's success.

2. Data Gathering

The first step in this analysis involved extracting data from the TMDB API. The dataset was chosen for its extensive coverage of Marvel films, including relevant details such as release year, rating, director, main actors, budget, and box office earnings.

3. Data Cleaning and Preprocessing

Once the data was gathered, the next step was to clean and preprocess the data to make it suitable for analysis. This involved removing any irrelevant columns, handling missing values, and converting data types. For instance, the 'Budget' and 'Box Office Earnings' columns were initially in string format, and they were converted into numeric values.

**4. Exploratory Data Analysis (EDA)**

After data cleaning, the exploratory data analysis was performed using Python and its libraries like Pandas, Matplotlib, and Seaborn. I used descriptive statistics to understand the central tendency, dispersion, and distribution of the data.

Budget vs Box Office Earnings
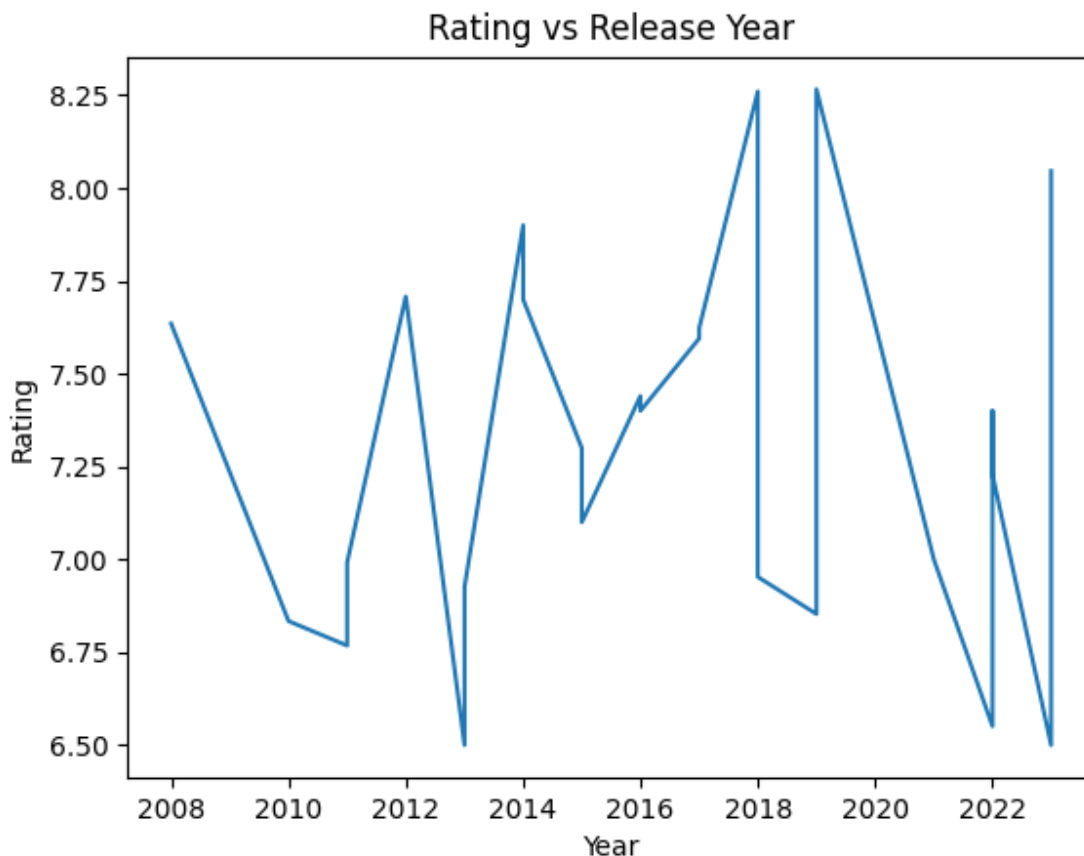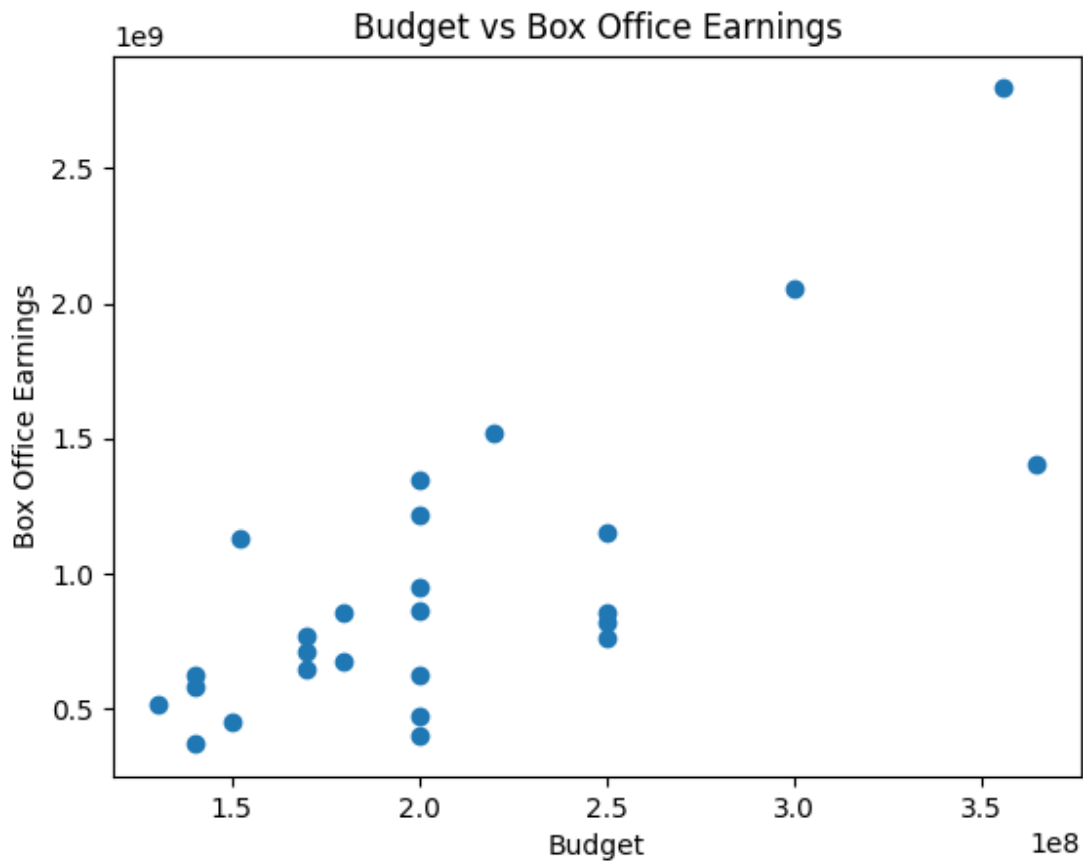
5. Correlation Analysis
The correlation between the variables was examined using a correlation matrix.
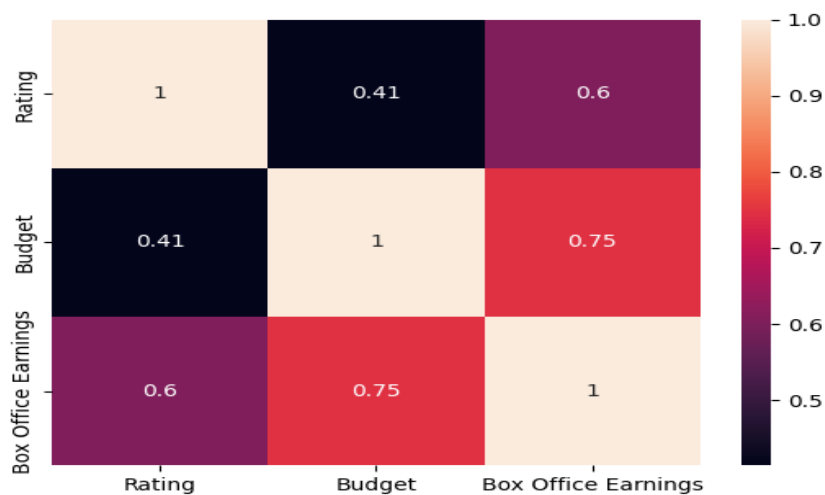
6. Regression Analysis
A multiple regression model was fitted with 'Rating' as the dependent variable and 'Budget' and 'Box Office Earnings' as the independent variables. The results were interpreted accordingly.

7. Results and Interpretation

The regression model revealed that the 'Box Office Earnings' variable was a significant predictor for the movie ratings, with a p-value of 0.014. However, 'Budget' was not a significant predictor.

```
                          OLS Regression Results
==============================================================================
Dep. Variable:                 Rating   R-squared:                       0.367
Model:                            OLS   Adj. R-squared:                  0.312
Method:                 Least Squares   F-statistic:                     6.675
Date:                Wed, 05 Jul 2023   Prob (F-statistic):            0.00518
Time:                        15:24:18   Log-Likelihood:                -12.880
No. Observations:                  26   AIC:                             31.76
Df Residuals:                      23   BIC:                             35.54
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                        coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const                 6.8541      0.310     22.111      0.000       6.213       7.495
Budget            -6.698e-10   2.06e-09     -0.325      0.748   -4.93e-09    3.59e-09
Box Office Earnings 6.203e-10   2.33e-10      2.666      0.014    1.39e-10     1.1e-09
==============================================================================
Omnibus:                        0.413   Durbin-Watson:                   1.724
Prob(Omnibus):                  0.814   Jarque-Bera (JB):                0.558
Skew:                           0.169   Prob(JB):                        0.757
Kurtosis:                       2.367   Cond. No.                     4.14e+09
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 4.14e+09. This might indicate that there are
strong multicollinearity or other numerical problems.

Process finished with exit code 0
```

8. Conclusion and Communication of Results

The analysis was summarized in a detailed report, with an emphasis on the findings' interpretation in the context of the project objectives. It was concluded that a movie's box office earnings appear to significantly impact its rating, while the production budget does not necessarily influence the film's success as much.

9. Documentation

Every step of the project, from data collection to conclusions, was meticulously documented to ensure reproducibility. Screenshots of all the graphs created during the analysis were taken and included in the final report for comprehensive understanding and transparency.
This project demonstrates the power of data analysis in revealing interesting insights and trends, serving as a foundation for future research exploring other influencing factors of Marvel movie's success.

Tools Used
* Data Gathering: TMDB API
* Data Analysis: Python (libraries - Pandas, NumPy, Scikit-learn)
* Data Visualization: Matplotlib, Seaborn