1. Project Objectives:

- Analyze a dataset containing various features of movies such as box office revenue, budget, director, genre, release date, cast, etc.
- Perform exploratory data analysis (EDA) to identify trends and patterns in movie performance.
- Build a predictive model to estimate a movie's success based on key attributes (such as budget, genre, or cast).
- Visualize data insights and findings in an intuitive manner using Python libraries.

2. Data Collection and Preprocessing:

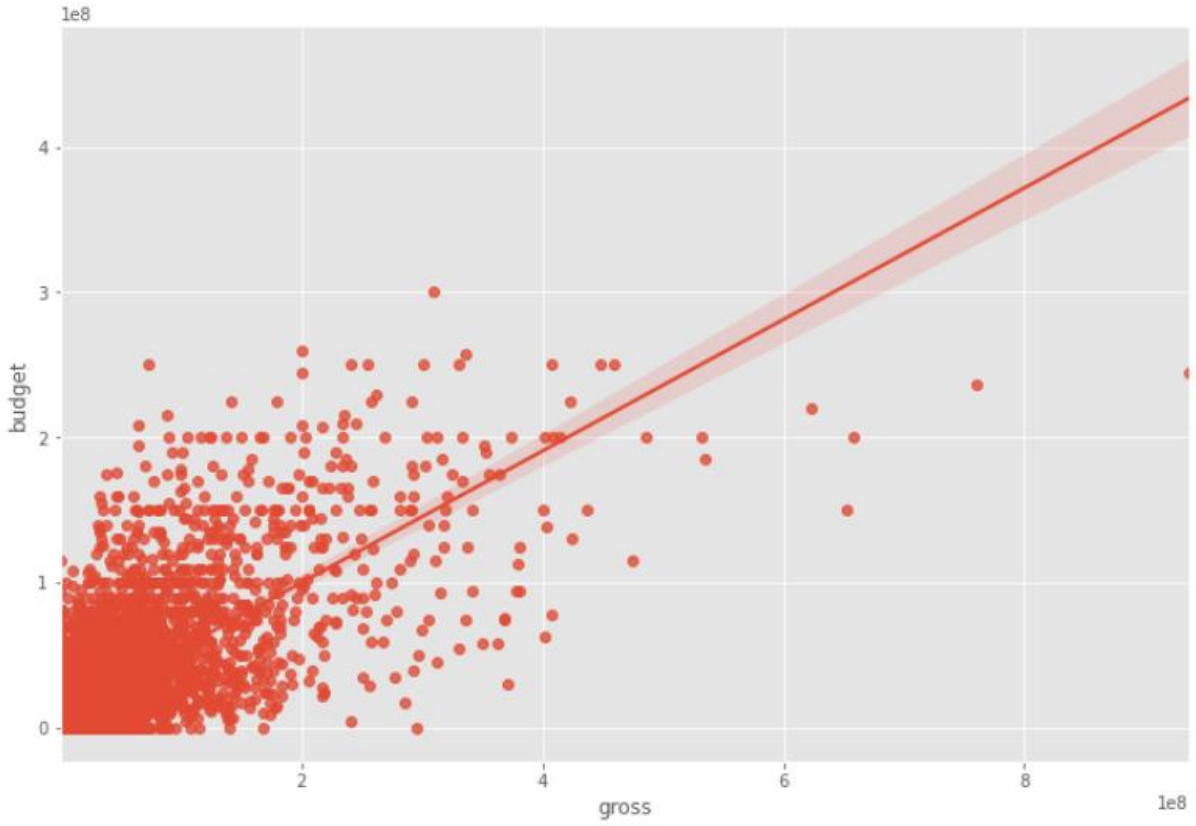
- Data sources: Working with existing datasets (CSV files).
- Data cleaning: Handling missing values, duplicates, or inconsistencies in data.
- Feature engineering: Creating new variables based on existing ones, such as calculating the ROI (Return on Investment) of movies, or categorizing movies by era (e.g., pre-2000 vs. post-2000).

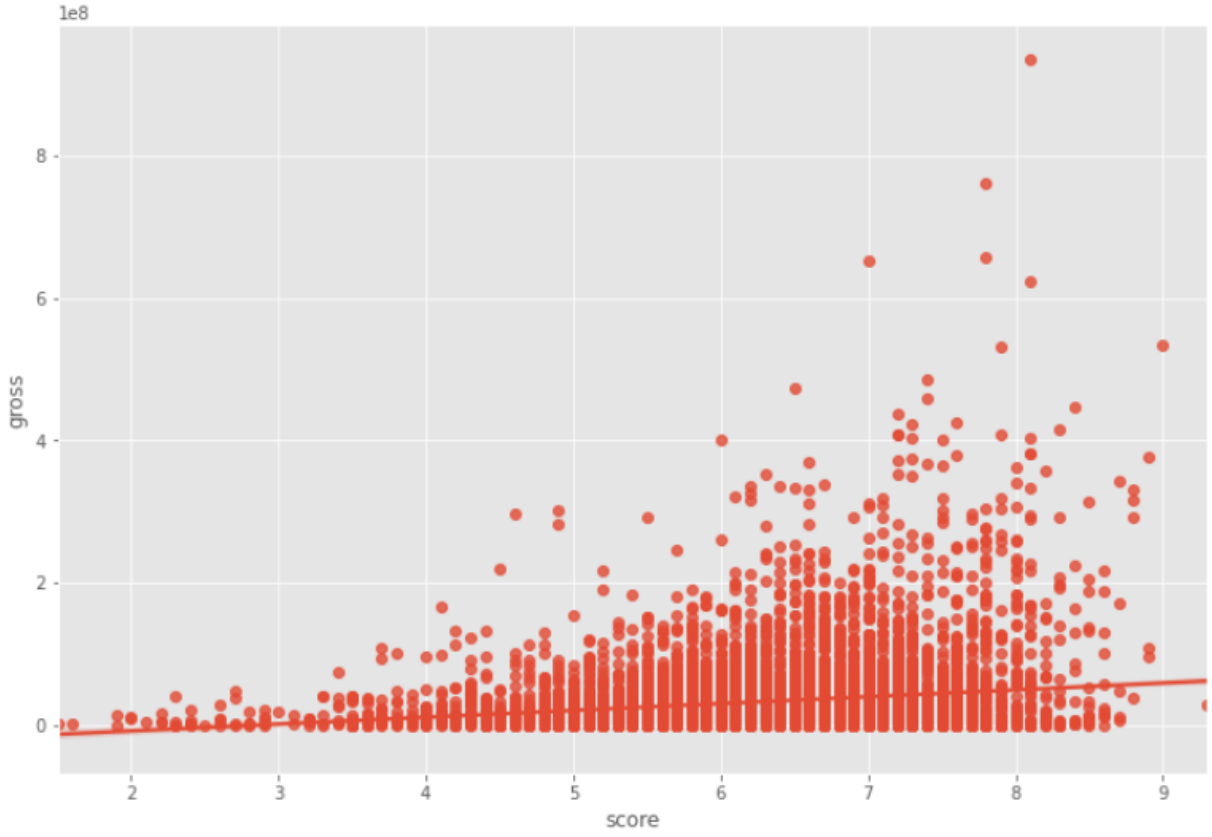
3. Exploratory Data Analysis (EDA):

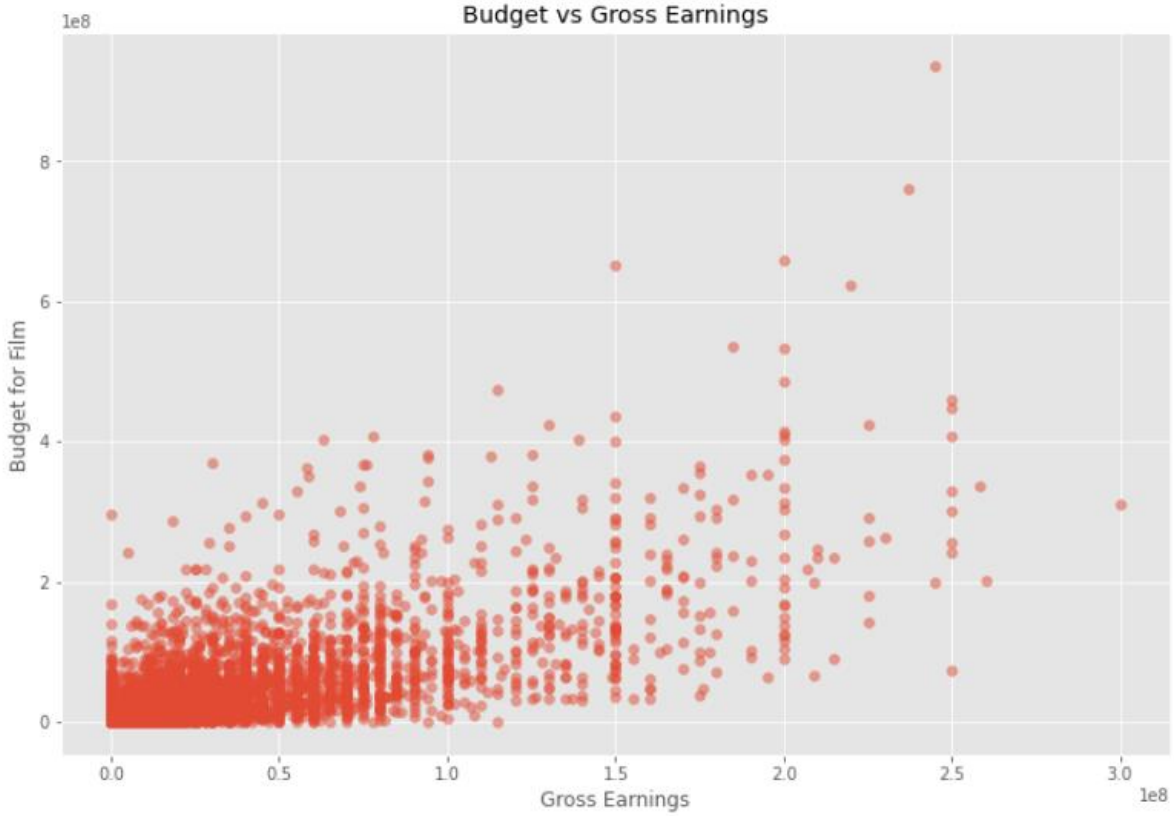
- Univariate Analysis: Analyze individual variables such as budget, revenue, and ratings to understand their distributions.
- Bivariate/Multivariate Analysis: Correlation analysis between variables, such as budget vs. box office revenue, or genre vs. average rating.
- Trend analysis: Identify trends over time, such as how movie budgets or box office revenues have changed by year, or the popularity of genres over time.
- Visualization: Use Python libraries like matplotlib, seaborn, or plotly to generate graphs such as histograms, box plots, and scatter plots to represent data trends and insights.

4. Results and Conclusions:

- Present findings and insights from the data analysis.
- Discuss any predictive models built, including their performance (accuracy, precision, recall, etc.) and possible improvements.




1e8
8
6
4
2
0
gross
2
3
4
5
6
7
8
9
score


Budget vs Gross Earnings
1e8
8
6
4
2
0
Budget for Film
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Gross Earnings
1e8