EDA PROJECT

MICROSFT MOVIE STUDIO

AUTHORED BY : BENITTO OTIENO

CONTENT

MICROSOFT MOVIE STUDIO EDA PROJECT

BUSINESS PROBLEM

THE MICROSOFT NEW MOVIE STUDIO EDA (EXPLORATORY DATA ANALYSIS) PROJECT AIMS TO ADDRESS THE CHALLENGE OF EFFICIENTLY ANALYZING AND EXTRACTING MEANINGFUL INSIGHTS FROM THE VAST AND DIVERSE DATA ASSOCIATED WITH THE CREATION AND DISTRIBUTION OF MOVIES. THIS INCLUDES DATA RELATED TO SCRIPTWRITING, CASTING, PRODUCTION, MARKETING, AND AUDIENCE ENGAGEMENT. THE GOAL IS TO LEVERAGE EXPLORATORY DATA ANALYSIS TECHNIQUES TO UNCOVER PATTERNS, TRENDS, AND POTENTIAL CORRELATIONS WITHIN THE DATA, ULTIMATELY EMPOWERING MICROSOFT'S MOVIE STUDIO TO MAKE INFORMED DECISIONS, OPTIMIZE PROCESSES, AND ENHANCE THE OVERALL SUCCESS OF THEIR MOVIE PRODUCTIONS.

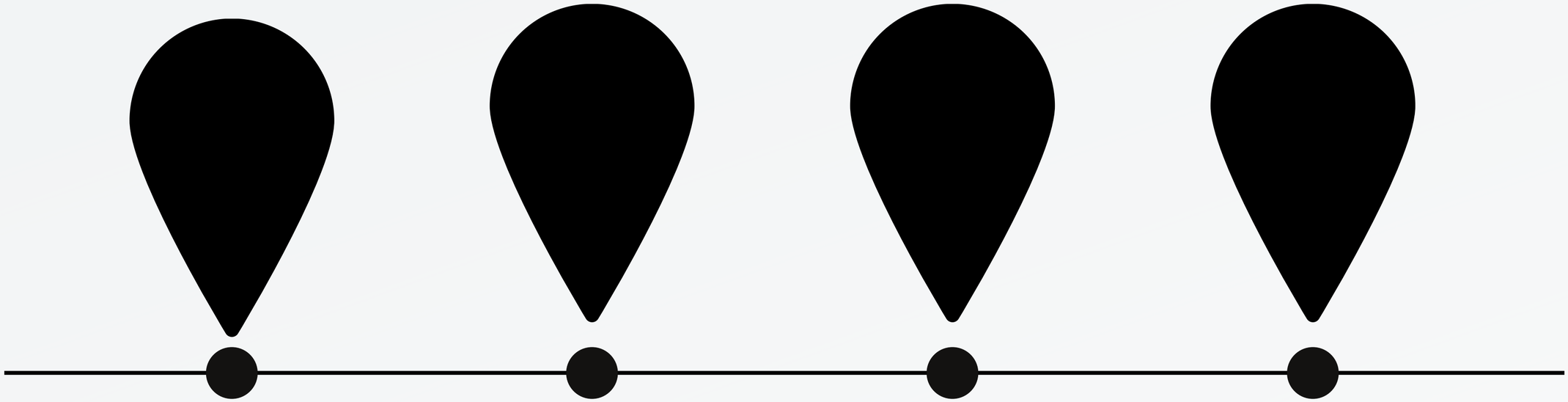
BUSINESS UNDERSTANDING

THE PROBLEM BEING ADDRESSED BY THE MICROSOFT NEW MOVIE STUDIO EDA PROJECT IS THE NEED FOR EFFECTIVE MANAGEMENT AND ANALYSIS OF DIVERSE DATA SETS ASSOCIATED WITH MOVIE PRODUCTION AND DISTRIBUTION. THIS INCLUDES BUT IS NOT LIMITED TO DATA RELATED TO SCRIPT DEVELOPMENT, CASTING DECISIONS, PRODUCTION COSTS, MARKETING STRATEGIES, AND AUDIENCE RESPONSE. THE CHALLENGE LIES IN HARNESSING THIS WEALTH OF INFORMATION TO GAIN ACTIONABLE INSIGHTS THAT CAN INFORM DECISION-MAKING PROCESSES, IMPROVE EFFICIENCY, AND CONTRIBUTE TO THE OVERALL SUCCESS OF MICROSOFT'S MOVIE PRODUCTIONS. THE PROJECT SEEKS TO IMPLEMENT ROBUST EXPLORATORY DATA ANALYSIS TECHNIQUES TO EXTRACT VALUABLE PATTERNS, TRENDS, AND CORRELATIONS FROM THE DATA, ENABLING THE STUDIO TO MAKE INFORMED AND STRATEGIC DECISIONS THROUGHOUT THE ENTIRE MOVIE-MAKING PROCESS.GRAPH TEXT

DATA COLLECTION

1. **Script and Story Development Data:** Information related to script versions, changes, and feedback during the development phase.
2. **Casting Data:** Details about actors, their availability, contracts, and casting decisions.
3. **Production Data:** Data on production schedules, costs, resource allocation, and logistical details.
4. **Financial Data:** Budgets, revenue streams, box office performance, and financial metrics associated with each movie.
5. **Marketing and Promotion Data:** Strategies, expenditures, and the effectiveness of marketing campaigns across different channels.
6. **Audience Engagement Data:** Metrics from social media, reviews, and audience feedback to gauge the reception of the movie.
7. **Distribution Data:** Information on the distribution channels, release dates, and geographic performance.
8. **Industry Trends and Market Data:** External data sources providing insights into broader industry trends, competitor performance, and market dynamics.
9. **Demographic Data:** Information about the target audience, their preferences, and demographics.
10. **Streaming and Digital Platforms Data:** Metrics related to digital distribution, streaming performance, and viewer engagement on online platforms.

EDA PROCESS



LOADING DATA

OVERVIEW OF DATA/DATA CLEANING

SUMMARY STAISTICS/ DATA MODELLING

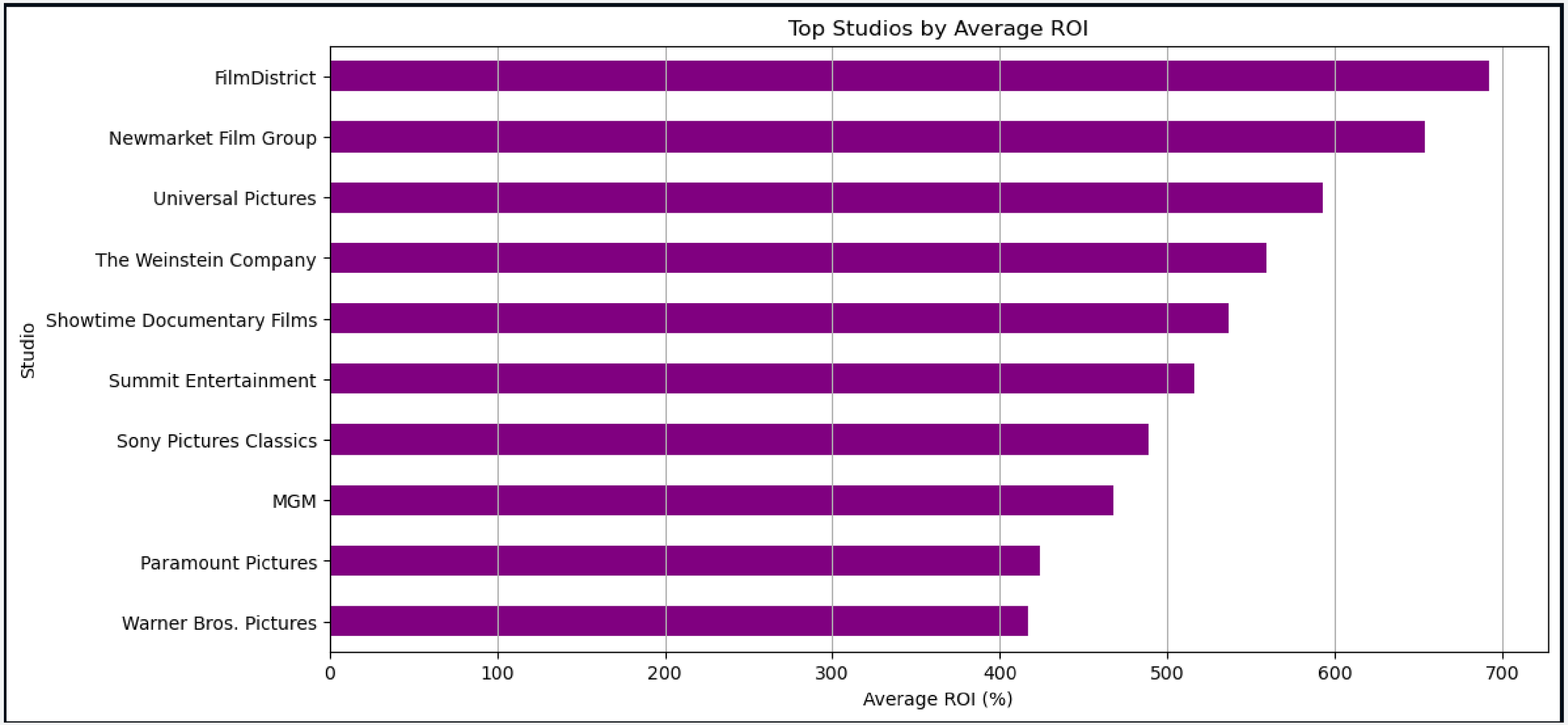
VISUALISATION

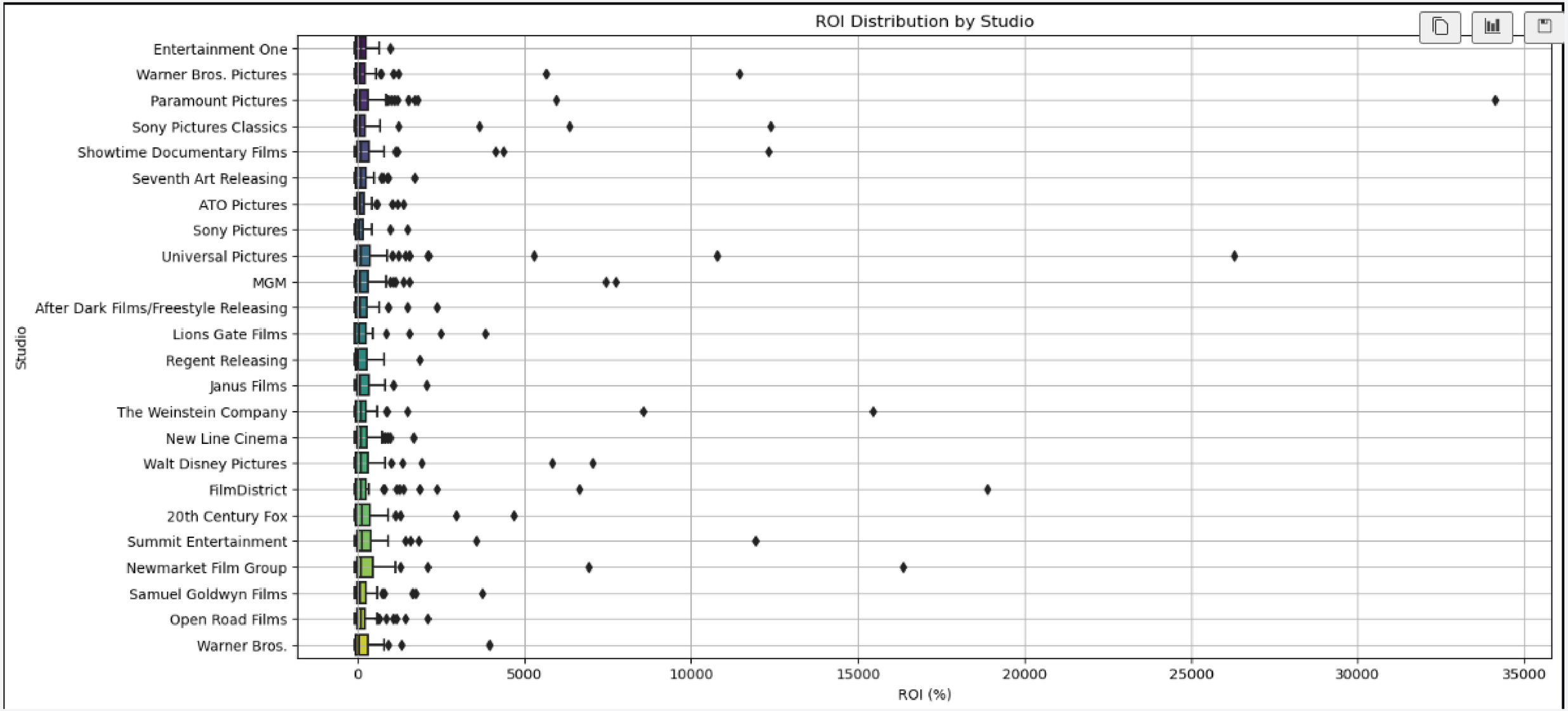
After loading the data I imported popular data analysis libraries such as Pandas, Matplotlib, and Seaborn.

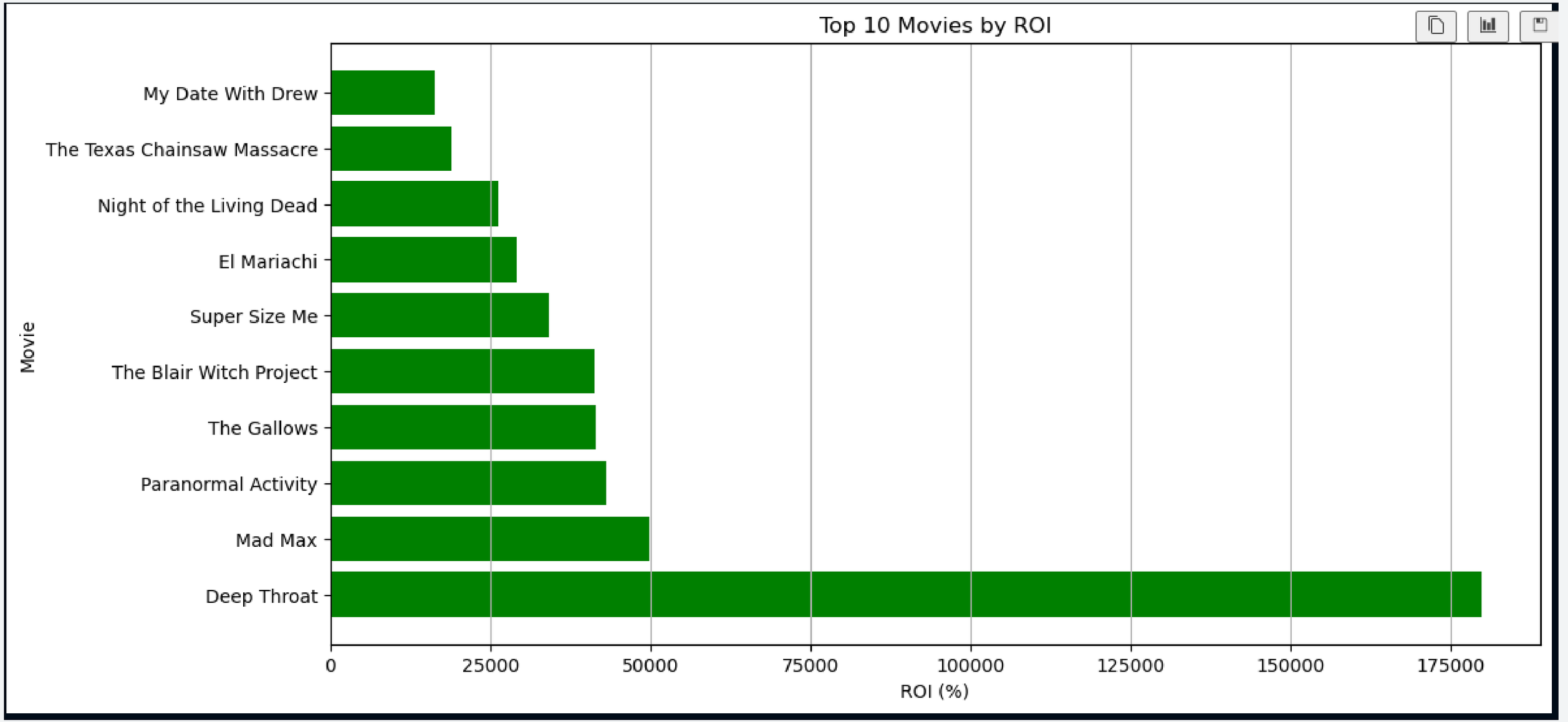
Data cleaning and Modelling involved :

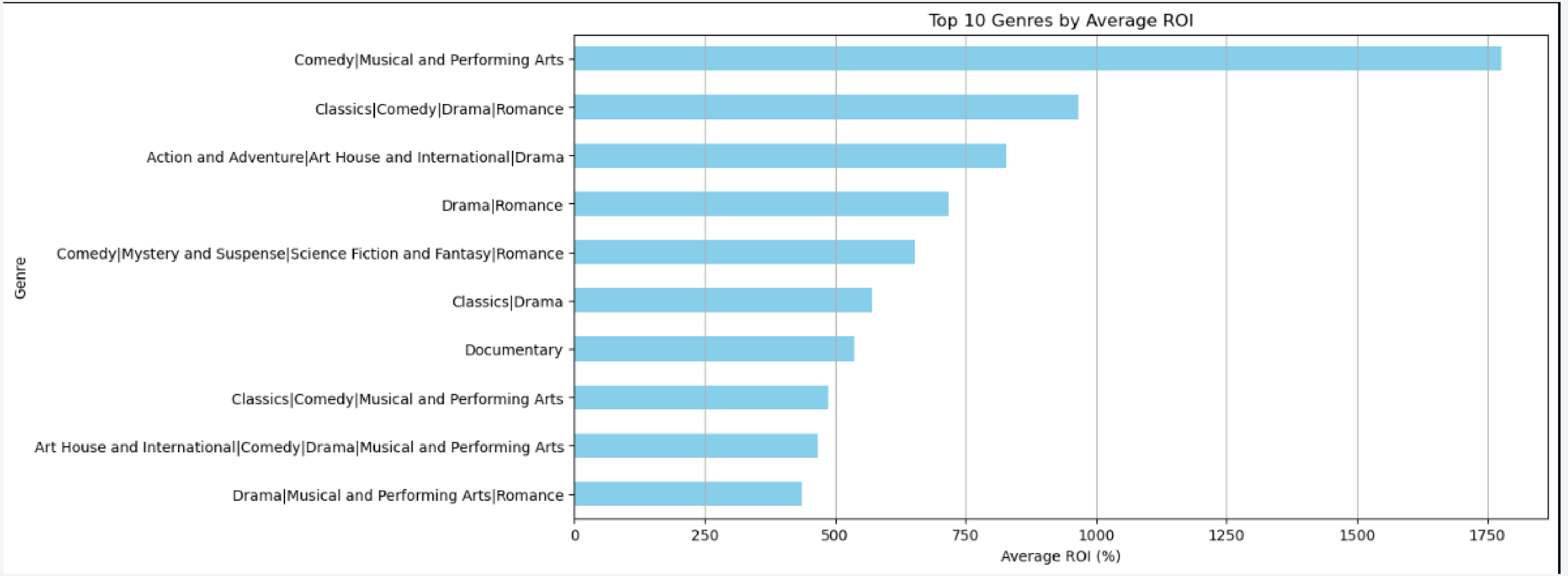
- **Handling Missing Values**
- **Removing Duplicates**
- Outlier Detection and Handling
- Merging of Data frames
- joining different tables

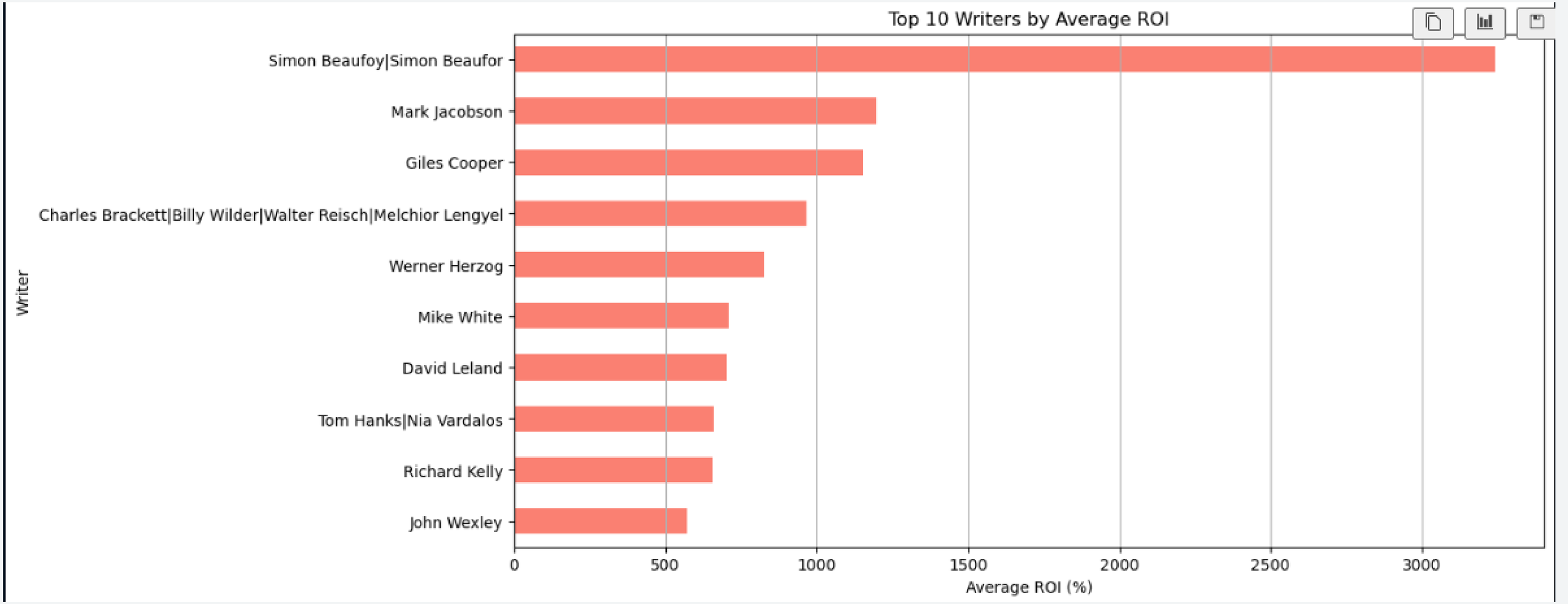
USED visuals like graphs & charts to SHOWCASE RESULTS & FINDINGS.

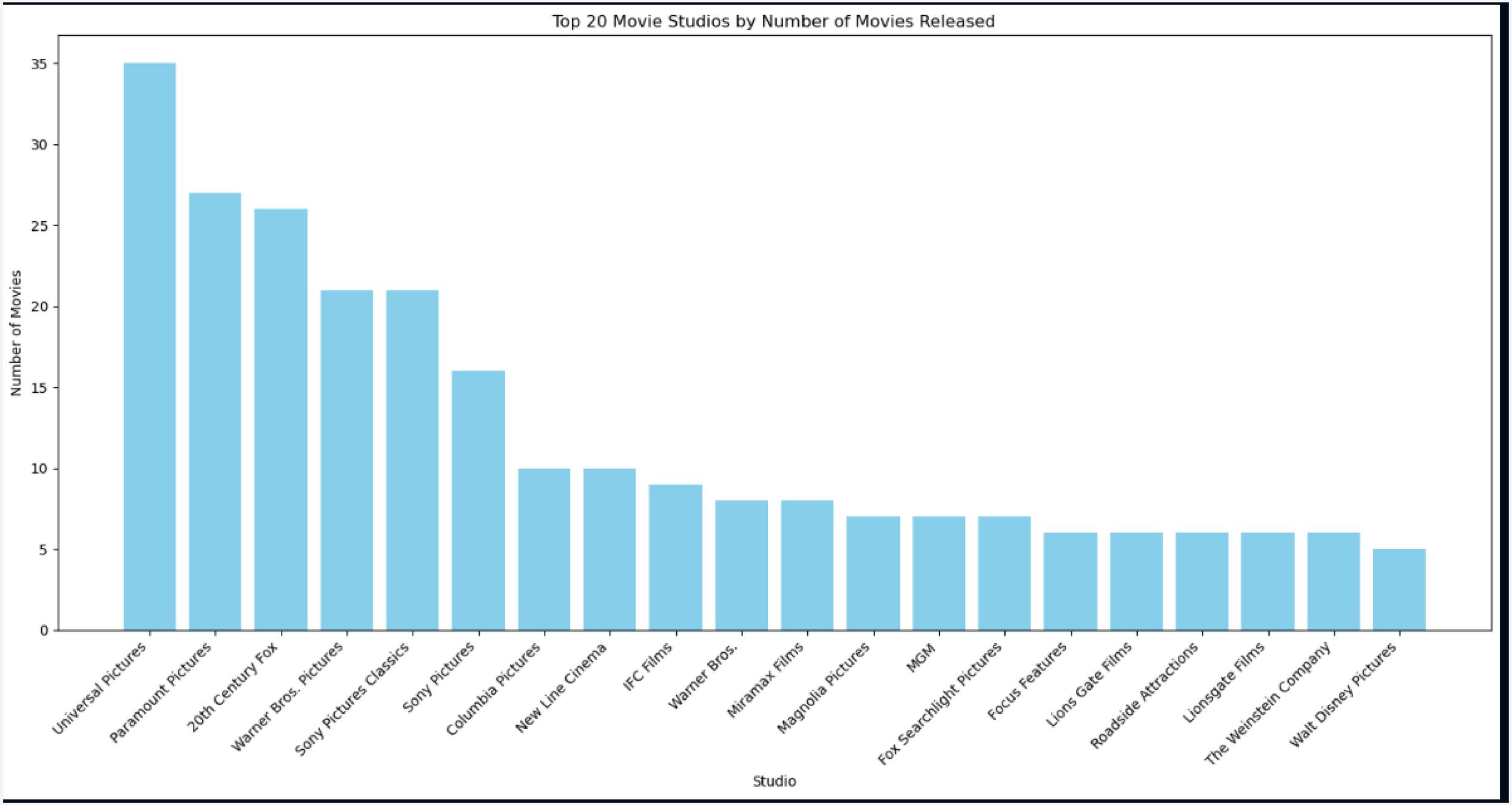

Top Studios by Average ROI
FilmDistrict
Newmarket Film Group
Universal Pictures
The Weinstein Company
Showtime Documentary Films
Summit Entertainment
Sony Pictures Classics
MGM
Paramount Pictures
Warner Bros. Pictures
Studio
0
100
200
300
400
500
600
700
Average ROI (%)


ROI Distribution by Studio
Entertainment One
Warner Bros. Pictures
Paramount Pictures
Sony Pictures Classics
Showtime Documentary Films
Seventh Art Releasing
ATO Pictures
Sony Pictures
Universal Pictures
MGM
After Dark Films/Freestyle Releasing
Lions Gate Films
Regent Releasing
Janus Films
The Weinstein Company
New Line Cinema
Walt Disney Pictures
FilmDistrict
20th Century Fox
Summit Entertainment
Newmarket Film Group
Samuel Goldwyn Films
Open Road Films
Warner Bros.
Studio
0
5000
10000
15000
20000
25000
30000
35000
ROI (%)


Top 10 Movies by ROI
My Date With Drew
The Texas Chainsaw Massacre
Night of the Living Dead
El Mariachi
Super Size Me
The Blair Witch Project
The Gallows
Paranormal Activity
Mad Max
Deep Throat
Movie
0
25000
50000
75000
100000
125000
150000
175000
ROI (%)


Top 10 Genres by Average ROI
Comedy|Musical and Performing Arts
Classics|Comedy|Drama|Romance
Action and Adventure|Art House and International|Drama
Drama|Romance
Comedy|Mystery and Suspense|Science Fiction and Fantasy|Romance
Classics|Drama
Documentary
Classics|Comedy|Musical and Performing Arts
Art House and International|Comedy|Drama|Musical and Performing Arts
Drama|Musical and Performing Arts|Romance
Genre
0
250
500
750
1000
1250
1500
1750
Average ROI (%)


Top 10 Writers by Average ROI
Simon Beaufoy|Simon Beaufor
Mark Jacobson
Giles Cooper
Charles Brackett|Billy Wilder|Walter Reisch|Melchior Lengyel
Werner Herzog
Mike White
David Leland
Tom Hanks|Nia Vardalos
Richard Kelly
John Wexley
Writer
0
500
1000
1500
2000
2500
3000
Average ROI (%)


Top 20 Movie Studios by Number of Movies Released
Number of Movies
0
5
10
15
20
25
30
35
Universal Pictures
Paramount Pictures
20th Century Fox
Warner Bros. Pictures
Sony Pictures Classics
Sony Pictures
Columbia Pictures
New Line Cinema
IFC Films
Warner Bros.
Miramax Films
Magnolia Pictures
MGM
Fox Searchlight Pictures
Focus Features
Lions Gate Films
Roadside Attractions
Lionsgate Films
The Weinstein Company
Walt Disney Pictures
Studio

Recommendations for the Business:

- **Genre Strategy**: Focus on genres that have demonstrated higher average ROIs, as identified in the "Top 10 Genres by Average ROI" analysis. Allocate resources and investments strategically based on genre preferences.

- **Writer Collaborations**: Consider collaborating with writers who have a track record of contributing to movies with higher ROIs, as indicated by the "Top 10 Writers by Average ROI" analysis. Building strong partnerships with successful writers can enhance the creative quality of the content.

- **Budget Allocation:** Analyze the ROI vs. Production Budget plot to guide budget allocation. Identify budget ranges that show a positive correlation with higher ROIs. Carefully plan and allocate resources for movie production based on this analysis.

- **Strategic Partnerships**: Explore partnerships or collaborations with top-performing studios, as highlighted in the "Top Studios by Average ROI" plot. Consider joint ventures or acquisitions to leverage the expertise of successful studios.

- **Diversification**: Recognize the importance of a diversified portfolio. While focusing on high-ROI genres and writers, diversification can help mitigate risks associated with the unpredictability of the movie industry.

Limitations and Considerations:

- Past Performance Doesn't Guarantee Future Success: The analysis is based on historical data, and the movie industry is dynamic. Trends and audience preferences can change, and past success does not guarantee future performance.

- External Factors: Movie success is influenced by various external factors, including market trends, competition, and unforeseen events. These factors may not be fully captured in the analysis.

- Incomplete Information: The dataset may not include all relevant variables influencing movie success. Additional data on marketing strategies, audience demographics, and critical reviews could enhance the analysis.

Future Improvements:

- Incorporate Additional Data: Include more comprehensive data, such as marketing budgets, social media engagement, and critical reviews, to provide a more holistic view of movie success factors.

- Machine Learning Models: Implement predictive modeling using machine learning algorithms to forecast potential movie success. This could involve building models to predict ROI based on various features.

- Real-time Monitoring: Establish a system for real-time monitoring of industry trends and audience preferences. Regularly update the analysis to stay ahead of evolving market dynamics.

- Feedback Mechanism: Implement a feedback mechanism that collects audience feedback and incorporates it into future projects. Continuous improvement based on audience responses can enhance success rates.

- Scenario Analysis: Conduct scenario analyses to assess the impact of different market scenarios on the success of movie projects. This can help in developing robust strategies for various contingencies.

- In conclusion, while the analysis provides valuable insights, it is essential for Microsoft to adapt its strategies based on the dynamic nature of the entertainment industry. Continuous monitoring, flexibility, and a data-driven approach will contribute to the success of Microsoft's venture into the movie industry.