**CHAPTER- I**

**INTRODUCTION TO DATA ANALYTICS**

**1.1 Introduction to Data Analysis / Analytics**

In today’s rapidly evolving digital landscape, the importance of data analysis and analytics cannot be overstated. The exponential growth of data generated from various sources—such as social media platforms, Internet of Things (IoT) devices, business transactions, and more—has created an unprecedented opportunity for organizations to harness this vast amount of information for strategic advantage. This abundance of data, often referred to as big data, encompasses a wide variety of data types and formats, including structured data like databases, unstructured data like text and images, and semi-structured data like XML files. The sheer volume and complexity of this data necessitate advanced methods and tools to effectively analyse and interpret it.

Data analysis involves the systematic application of statistical, mathematical, and logical techniques to describe, illustrate, condense, and evaluate data. It is a meticulous process that encompasses several stages, including data collection, inspection, cleaning, transformation, and modelling. The goal of data analysis is to uncover valuable information that can guide decision-making and provide insights into underlying trends and patterns within the data. By transforming raw data into a meaningful and understandable format, data analysis facilitates informed decision-making and problem-solving across various domains.

Data analytics, on the other hand, encompasses a broader scope of activities aimed at extracting meaningful insights and actionable intelligence from raw data. It goes beyond simple data analysis by leveraging advanced tools, algorithms, and technologies to handle complex datasets and perform in-depth analysis. Data analytics involves a variety of methods, including data mining, machine learning, predictive modelling, and statistical analysis, to identify patterns, correlations, and trends within the data. These insights help organizations understand their past performance, predict future outcomes, and make data-driven decisions that enhance operational efficiency, drive innovation, and improve competitive positioning.

**1.2 Data Analytics Approaches**

**Descriptive Analytics:** To summarize and interpret historical data to understand what has happened in the past. Use of statistical measures such as mean, median, mode, and standard deviation. Data visualization tools like charts, graphs, and dashboards are commonly used to present the data in an easily interpretable format. Useful for identifying trends and patterns, such as analysing sales data to understand revenue trends over time.

**Diagnostic Analytics:** To investigate and understand the root causes of past outcomes or events. Use of techniques such as correlation analysis, root cause analysis, and drill-down analysis to explore data and identify underlying factors that contributed to specific outcomes. Commonly used in business to understand why sales dropped in a particular quarter or to determine the factors contributing to product defects.

**Predictive Analytics:** To forecast future trends and behaviours by analysing current and historical data. Implementation of statistical models, machine learning algorithms, and regression analysis to predict future events or trends. Frequently used for demand forecasting, risk assessment, and customer behaviour prediction, such as predicting which customers are likely to churn.

**Prescriptive Analytics:** Prescriptive analytics goes beyond prediction and provides recommendations on the best course of action to achieve desired outcomes. It uses optimization algorithms, simulation models, and decision analysis to generate actionable insights. Prescriptive analytics helps organizations optimize resources, streamline processes, and make data-driven decisions.

**Exploratory Analytics**: To explore data sets to uncover patterns, anomalies, and relationships without having any preconceived notions or hypotheses.

**1.3 Steps of Data Analytics**

**Data Collection: Data Collection** is the foundational step in the data analytics process, where relevant data is gathered from various sources to address specific business questions or research objectives. This stage is critical as the quality and breadth of the data collected directly impact the accuracy and reliability of the subsequent analysis. Data can be collected from a multitude of sources.

**Data Cleaning: Data Cleaning**, also known as data cleansing or data scrubbing, is a critical process in which collected data is prepared for analysis by identifying and rectifying errors, inconsistencies, and inaccuracies. The quality of data can be compromised by various issues, such as missing values, duplicate entries, and incorrect formatting.

**Data Transformation: Data Transformation** is the process of converting data into a suitable format for analysis. This involves modifying the structure and content of the data to meet the requirements of the analysis tools and techniques being used.

**Data Analysis:** Perform exploratory data analysis to understand data distributions, identify patterns, and detect anomalies.Use visualization techniques like histograms, scatter plots, and correlation matrices to gain insights into the data.Identify key features that might influence churn, such as average purchase value, frequency of visits, and customer tenure.

**Data Interpretation:** Create new features from existing data to enhance model performance. For example, compute the average purchase value per month or the time since the last purchase. Select features that are highly correlated with churn and drop irrelevant or redundant features.

**1.4 Applications of Data Analytics**

**Content Strategy:** Data analytics provides a comprehensive understanding of audience preferences, viewing habits, and emerging trends, which can be leveraged to inform content creation. By analysing which genres, themes, and formats resonate most with viewers, content creators can tailor their productions to meet audience demands. This involves assessing metrics such as viewership numbers, completion rates, and user ratings. For instance, if data indicates a high interest in crime dramas, producers can prioritize developing more content in this genre.

**Audience Engagement:** Data analytics enables the segmentation of audiences based on various criteria such as age, gender, location, and viewing habits. This segmentation helps in understanding the diverse preferences and behaviors of different viewer groups. For instance, younger audiences might prefer action-packed series, while older demographics might lean towards documentaries or classic films. These insights allow Amazon to tailor its content and marketing strategies to meet the specific needs of each segment.

**Marketing:** Data analytics enables the creation of highly targeted marketing campaigns by identifying specific audience segments and their preferences. By analyzing data on user demographics, behavior, and viewing history, marketers can design personalized ads that resonate with potential viewers. For instance, promoting a new thriller series to users who frequently watch crime dramas increases the effectiveness of the campaign. This continuous feedback loop ensures that marketing strategies remain effective and aligned with audience interests.

**Subscription and Pay-Per-View Models:** Data analytics helps identify high-performing content that can drive subscription growth or pay-per-view revenue. By analyzing user engagement and satisfaction metrics, platforms can pinpoint which shows and movies have the greatest potential for monetization. This information can be used to develop pricing strategies and subscription plans that maximize revenue while offering value to users.

**Maximizing Ad Revenue:** For ad-supported content, data analytics helps in optimizing ad placements and targeting. By analyzing user behavior and engagement patterns, platforms can ensure that ads are placed in the most effective positions, enhancing their impact and value. This targeted approach increases the likelihood of viewers interacting with ads, thereby boosting ad revenue.

**Chapter-II**

**OVERVIEW OF THE PROBLEM**

**2.1 Problem Study**

The entertainment landscape has evolved significantly with the rise of streaming platforms, creating an immense pool of data that holds valuable insights into audience preferences and content trends. This project aims to harness this data to analyse Amazon shows and movies, uncovering patterns that can inform strategic decisions for content creators and marketers. By delving into various aspects such as ratings, genres, and viewership, the project seeks to address the following key areas. One of the primary objectives of this project is to gain a deep understanding of audience preferences. Streaming platforms like Amazon Prime Video offer a vast array of content spanning different genres, formats, and themes. However, not all content resonates equally with viewers. By analysing viewer ratings and engagement metrics, the project aims to identify which genres and types of content are most favoured by audiences. This insight is crucial for content creators who must align their production strategies with audience tastes to maximize viewership and satisfaction.

The digital entertainment industry is dynamic, with trends continuously evolving in response to changing viewer behaviours and preferences. This project seeks to identify these trends by analysing historical data on Amazon shows and movies. Understanding which genres are gaining popularity, how viewer preferences shift over time, and what content characteristics lead to higher engagement can provide valuable foresight. This foresight enables content creators and platform managers to anticipate future demands and adapt their strategies accordingly.

Predicting the future performance of new content is a significant challenge for streaming platforms. This project aims to address this challenge by developing predictive models that can forecast how new shows and movies might perform based on historical data. By analysing patterns in viewer ratings, genre preferences, and other relevant factors, the models can estimate potential viewership and engagement levels for upcoming content. Such predictive insights are invaluable for making informed decisions about content acquisition, production investments, and marketing strategies.

The insights gained from this analysis have far-reaching implications for both content creation and marketing. For content creators, understanding what resonates with audiences can guide the development of new shows and movies that are more likely to succeed. For marketers, insights into viewer preferences and trends can inform targeted promotional strategies, ensuring that content reaches the right audience at the right time.

**2.2 Existing & Proposed System**

**Existing System:** Current methodologies in analysing streaming content often involve basic metrics such as view counts, ratings, and simple trend analysis. These approaches may lack depth and fail to uncover deeper insights due to

* **Limited Data Integration**: Not combining data from various sources for a comprehensive view.
* **Basic Analytical Techniques**: Relying on simple statistical methods rather than advanced analytics.
* **Manual Processes**: Involvement of manual data handling, which is time-consuming and prone to errors.

**Proposed System:** The proposed system aims to utilize advanced data analytics techniques to provide a more comprehensive and in-depth analysis. This includes:

The proposed system aims to overcome the limitations of existing systems by:

* **Automated Data Collection**: Using APIs and web scraping to gather data efficiently.
* **Advanced Analytics**: Employing machine learning models, predictive analytics, and sentiment analysis.
* **Data Visualization**: Creating interactive dashboards and visualizations for better understanding and presentation of insights.
* **Integration of Multiple Data Sources**: Combining structured and unstructured data from various sources to get a holistic view.

**2.3 Hardware / System Requirements**

For fast data processing and running complex analytical models. To store large datasets securely and efficiently. Essential for training machine learning models and handling intensive computations. The system used in this project is Windows 11. It has 11th Gen Intel(R) Core (TM) i3-1115G4 @ 3.00GHz processor with 8.00GB RAM and 64-bit operating system, x64-based processor system type.GPU is Optional for large-scale data processing.

**2.4 Software, Tools and Libraries Requirements**

* **Operating System**: Windows 11
* **Programming Language**: Python 3.11.4
  + **Libraries**:
    - pandas: For data manipulation and analysis.
    - NumPy: For numerical computations.
    - **Plotly**: An interactive graphing library for Python.
    - Seaborn: For data visualization.
    - Matplotlib: For graphical plotting.
  + **Development Environment**: Jupyter Notebook or any Python IDE.
  + **Data Source:** Excel, CSV files.

**PYTHON:**

Python is a high-level programming language that is widely used for various purposes, such as web development, data analysis, machine learning, scientific computing, and more.

Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations, and narrative text. It is commonly used for data analysis, scientific computing, machine learning, and more.

**CHAPTER- III**

**DATA PREPARATION**

**3.1 Data Collection Approaches**

**Amazon Prime Video API**: An API provided by Amazon to access data related to its streaming content. This can include metadata about shows and movies, user ratings, reviews, and more. APIs allow for programmatic access to data, making it easier to collect large amounts of information quickly and consistently.

**Web Scraping:** The process of extracting data from websites by parsing the HTML content. This can be done using tools like Beautiful Soup, Scrappy, or Selenium in Python. Web scraping can be used to gather data from Amazon Prime Video's website, such as user reviews, ratings, and other content-related information that might not be available through the API.

**Third-Party Datasets:** Pre-collected datasets that are available from third-party sources, such as data repositories, academic publications, or commercial data providers. These datasets can provide additional context or supplementary data that might not be directly available from Amazon, such as demographic information, industry reports, or competitor data.

**User Surveys and Feedback**: Collecting data directly from users through surveys, feedback forms, and interviews. Gathering qualitative data on viewer preferences, satisfaction, and opinions about specific shows and movies. This can provide insights that quantitative data alone might miss.

**Social Media Monitoring**: Analysing social media platforms such as Twitter, Facebook, and Instagram for mentions, reviews, and discussions about Amazon shows and movies. Social media analysis can reveal trends in public opinion, emerging popular shows, and real-time feedback from viewers.

**3.2 Data Method**

**Exploratory Data Analysis:**

Exploratory Data Analysis is a method of evaluating or comprehending data in order to derive insights or key characteristics. EDA can be divided into two categories: graphical analysis and non-graphical analysis. EDA is a critical component of any data science or machine learning process.

**Data Frame NaN Summary:**

Data Frame NaN Summary provides an overview of missing or Nan (Not a Number) values in a pandas Data Frame, typically presented as counts per column.



(Fig 3.2)

**Data Cleaning:**

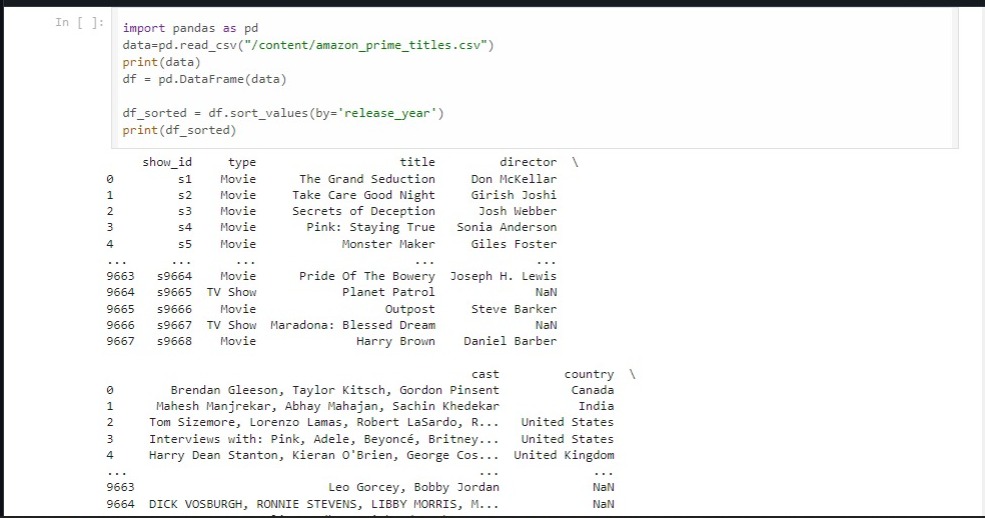
Missing or irrelevant information was removed to ensure data quality.



(Fig 3.3)

**Data Transformation:**

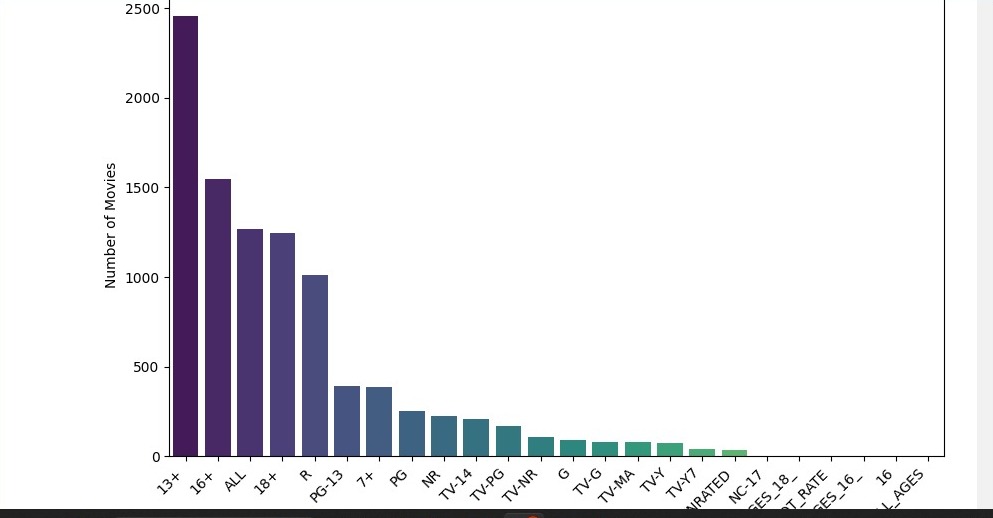
Data transformation refers to the process of converting raw data into a more useful format, often involving cleaning, normalization, aggregation, or other operations to prepare it for analysis or machine learning tasks.



(Fig 3.4)

**Data Visualization:**

Data visualization is the graphical representation of information and data. It involves creating visual elements like charts, graphs, and maps to make complex data more accessible, understandable, and usable.



(fig 3.5)

**Feature engineering:** Feature engineering involves creating new features or transforming existing ones in a dataset to improve the performance of machine learning models.



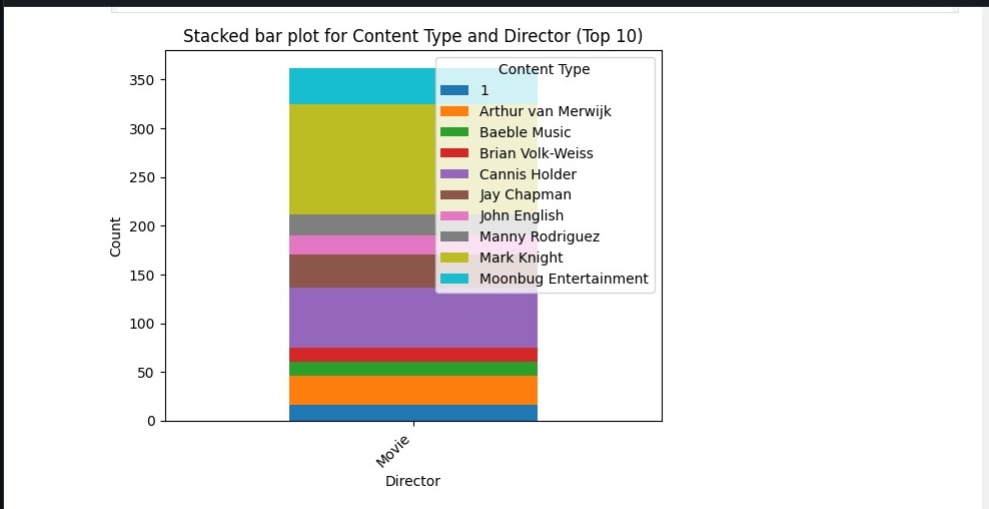
(Fig 3.6)

**Distribution of ratings:** The "distribution of ratings" refers to how ratings or scores are spread across different categories or levels, showing the frequency or proportion of each rating value within a dataset.



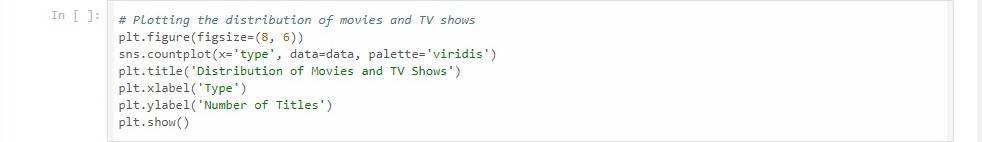
(fig 3.7)

**Stacked bar plot:** A "stacked bar plot" is a type of chart that displays multiple categories of data stacked on top of each other. It visualizes the composition of each category by showing the cumulative contribution of different variables or subcategories.

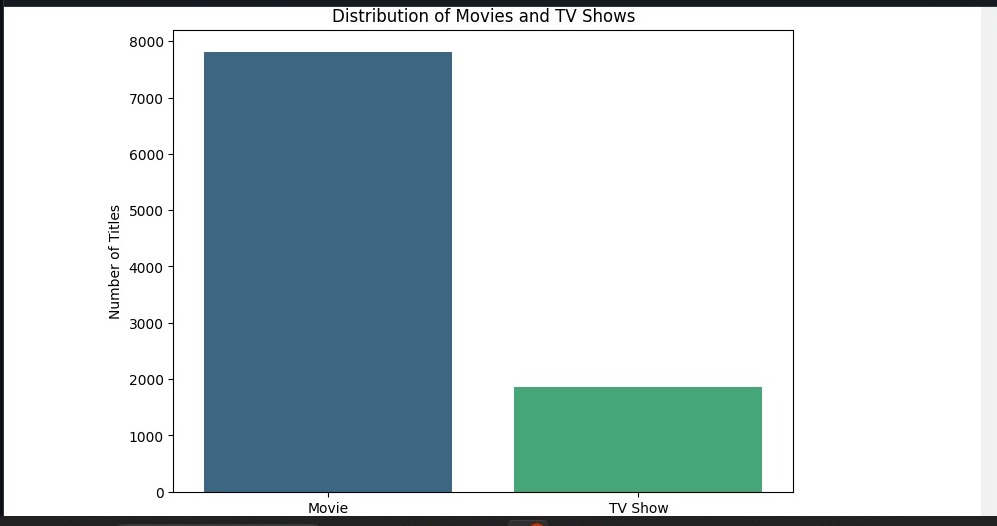


(Fig 3.8)

**Distribution of movies and tv shows:** The "distribution of movies and TV shows" refers to how these entertainment categories are divided or spread across a dataset, showing the relative frequency or proportion of each category within a collection or platform.



(fig 3.9)



(fig 3.10)

**3.3 Purpose of Data**

The purpose of collecting data on Amazon shows and movies is multifaceted. Primarily, it aims to analyse trends within the content available on the platform. This analysis helps to identify popular genres, themes, and formats that resonate with audiences. By understanding these preferences, content creators and platform managers can make informed decisions about what types of shows and movies to produce or acquire.

Additionally, the data serves a predictive role by enabling the development of models that forecast the performance of new content. These predictive models can estimate viewer engagement, ratings, and potential viewership based on historical data and the characteristics of past successful productions. Such insights are crucial for optimizing content strategies, allocating resources effectively, and maximizing the overall appeal and success of Amazon's entertainment offerings.

**CHAPTER- IV**

**METHODOLOGY**

The methodology for analysing Amazon shows and movies encompasses several key stages to extract meaningful insights and predictions. Here's an elaboration on each step along with additional methodologies:

**Data Pre-processing:** This initial stage involves cleaning the collected data to handle missing values, remove duplicates, and format data into a consistent structure.

**Feature Engineering:** Feature engineering enhances the predictive power of models by creating new features or transforming existing ones.

**Model Training:** This step involves splitting the data into training and validation sets, selecting appropriate algorithms, and tuning model parameters to optimize performance.

**Exploratory Data Analysis (EDA):** EDA is crucial for gaining a deeper understanding of the dataset.

**Evaluation:** Models are evaluated using appropriate metrics (e.g., accuracy, precision, recall, or RMSE for regression) to assess how well they generalize to new data.

**Deployment:** Successful models are deployed to provide actionable insights and predictions for decision-making.

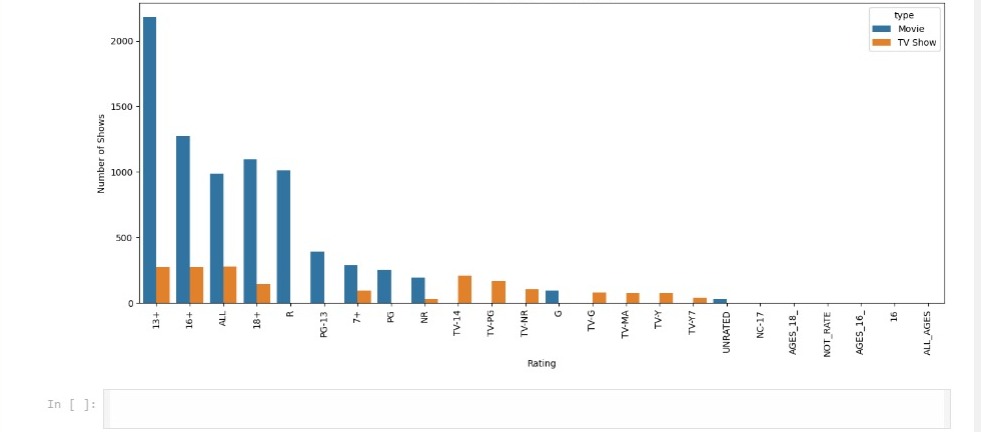
**Time Series Analysis:** Applied to understand seasonal viewership patterns, detect trends in content popularity over time, and forecast future viewership or engagement metrics.

By integrating these methodologies, Amazon can effectively leverage data-driven insights to enhance content curation, improve viewer engagement, and optimize business strategies in the competitive landscape of digital entertainment.

**CHAPTER- V**

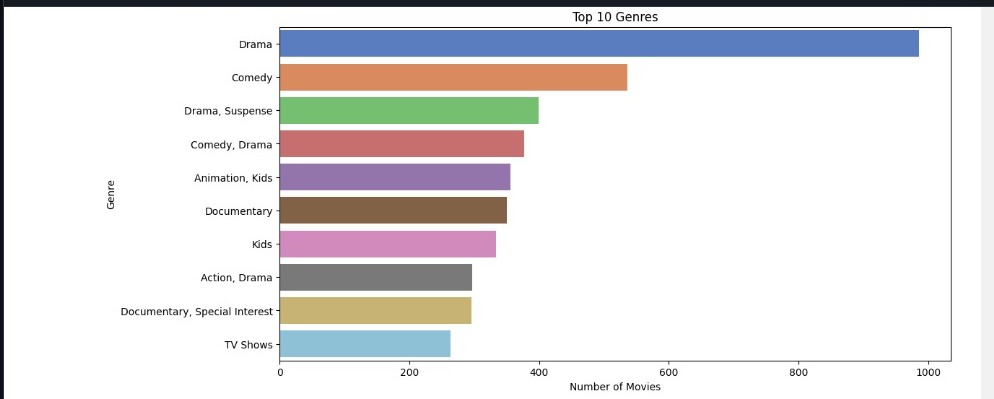
**RESULTS & FINDING INSIGHTS**

**Show and Movie Ratings:** Analysis of viewer ratings revealed that certain genres consistently receive higher ratings than others. This includes genres such as drama, thriller, and comedy, indicating strong audience preferences for these types of content.

****

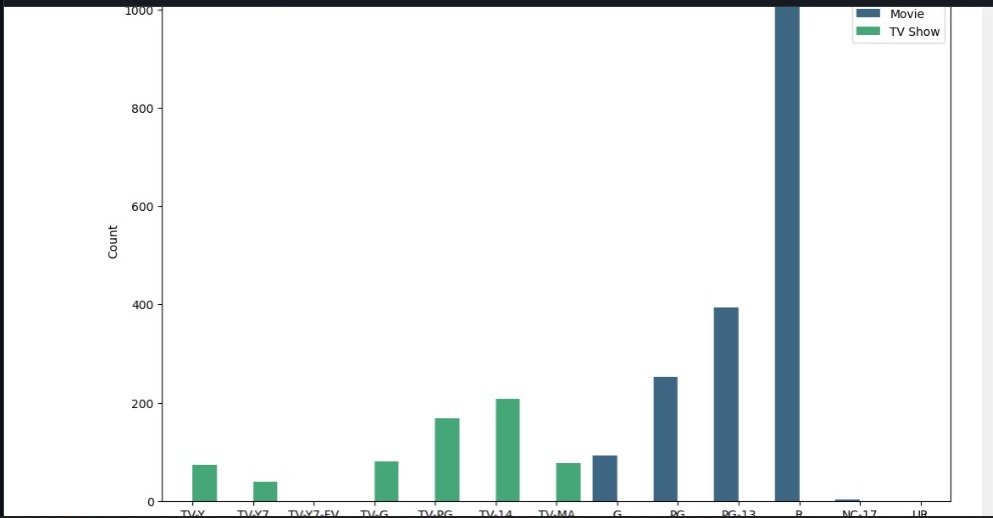
**(fig 5.1)**

**Genre Preferences:** Genre emerged as a critical predictor of content success. Specific genres, such as sci-fi or romance, demonstrated varying levels of popularity and viewer engagement. Understanding these preferences helps in tailoring content offerings to better meet audience expectations.

****

**(fig 5.2)**

**Viewership Patterns:** Patterns in viewership across different genres and formats (e.g., series vs. movies) were observed. Insights into when viewers engage most actively with content, such as during weekends or evenings, provide strategic timing for content releases and marketing campaigns.

****

**(fig 5.3)**

**Key Predictors of Success:** Factors such as content duration and age appropriateness were identified as predictors of viewer engagement and satisfaction. Content that aligns well with viewer expectations in these areas tends to perform better in terms of ratings and viewership.

**Key Findings:**

 Genre**-driven Preferences:** Viewers exhibit clear preferences for specific genres, influencing their ratings and viewership behaviours. This underscores the importance of genre diversity in content libraries to cater to varied audience tastes

 Content **Duration Impact:** Shorter or longer durations can affect viewer satisfaction differently across genres. Understanding optimal content length for different genres can enhance viewer retention and engagement

 Emerging **Trends:** Analysis highlighted emerging trends in content consumption, such as increasing popularity in niche genres or a shift towards binge-watching series over standalone movies. These trends provide strategic opportunities for content creation and marketing strategies.

* **Visualizations and Metrics:** Visual representations, such as stacked bar charts or heatmaps, effectively illustrated genre popularity and viewership trends. Metrics like average ratings per genre or viewership peaks provided actionable insights for content strategy and decision-making.

these results and findings equip Amazon with actionable insights to optimize content offerings, enhance viewer satisfaction, and drive strategic decisions in content acquisition, production, and marketing initiatives.

**CHAPTER-VI**

**DICUSSION & FUTURE WORK**

**Discussion**

The analysis of Amazon shows and movies has revealed insightful trends that are crucial for understanding audience preferences and content performance. One of the most significant findings is the strong correlation between content genre and viewer ratings. This relationship suggests that viewers have distinct preferences for certain genres, which in turn influence their overall satisfaction and engagement with the content. For instance, genres such as drama, thriller, and comedy consistently attract higher ratings, indicating their broad appeal. This preference underscores the importance of offering a diverse range of genres to cater to varied audience tastes. The findings also highlight emerging trends in content consumption, such as the growing preference for binge-watching series over standalone movies. This trend suggests a shift in viewing habits where audiences increasingly favor continuous, long-form storytelling. Understanding these trends allows for better planning of content releases and marketing strategies to capture and retain viewer interest.

The visualizations used in the analysis, such as heatmaps and bar charts, have effectively illustrated the distribution of genre popularity and viewership patterns. These tools provide a clear and actionable view of the data, enabling stakeholders to make informed decisions. Moreover, the identification of key predictors of content success, including genre, ratings, and duration, provides a foundation for developing targeted strategies to enhance content offerings and improve viewer satisfaction.

**Future Work**

* **Deeper Audience Engagement Analysis:** A more granular analysis of audience engagement could be conducted, examining not just the ratings but also viewer retention rates, watch time, and interaction metrics such as likes, comments, and shares. This would provide a more comprehensive view of how audiences interact with content and identify factors that contribute to long-term engagement and loyalty.
* **Cross-Platform Comparative Analysis:** Expanding the analysis to include other streaming platforms like Netflix, Hulu, or Disney+ would provide valuable comparative insights. This would help identify platform-specific trends and preferences, as well as uncover potential areas for Amazon to differentiate itself and enhance its competitive positioning.
* **Refinement of Predictive Models:** The predictive models could be refined to enhance their accuracy and predictive power. This could involve incorporating additional variables such as marketing spend, release timing, and social media buzz, which might influence content performance.
* **Incorporating Viewer Demographics and Psychographics:** Future research could integrate demographic and psychographic data to understand how different audience segments respond to various types of content. This includes age, gender, location, and lifestyle preferences.
* **Exploring Interactive and New Media Content:** With the rise of interactive content and new media formats such as virtual reality (VR) and augmented reality (AR), future work could explore how these formats impact viewer engagement and satisfaction.

**CHAPTER- VII**

**SUMMARY & CONCLUSION**

**Summary:**

This project underscores the pivotal role of data analytics in comprehending and forecasting trends within the realm of Amazon shows and movies. By leveraging a robust analytics framework, the study has provided a comprehensive understanding of the factors that drive content success and audience engagement. The primary focus has been on analysing a vast dataset to uncover patterns related to viewer ratings, genre preferences, and viewership behaviours. Key findings from the analysis include a strong correlation between content genre and viewer ratings, indicating clear genre preferences among audiences. The data also revealed that content duration significantly impacts viewer satisfaction, with varying optimal lengths for different genres. Additionally, the analysis identified emerging trends such as the increasing popularity of binge-watching series, which reflects a shift in viewer habits towards continuous, long-form storytelling.

The methodologies employed in this project, including data pre-processing, exploratory data analysis (EDA), feature engineering, and predictive modelling, have been instrumental in extracting meaningful insights from the data. These insights are visualized through effective tools such as heatmaps and bar charts, which illustrate genre popularity and viewership patterns in a clear and actionable manner.

Furthermore, the project has highlighted the importance of understanding audience engagement at a deeper level, which can inform content creation and marketing strategies. The identification of key predictors of content success, such as genre, rating, and duration, provides a foundation for developing targeted approaches to enhance content offerings and improve viewer satisfaction.

Overall, the project demonstrates how data analytics can be harnessed to gain valuable insights into content trends and audience preferences, ultimately leading to more informed and effective decision-making in the digital entertainment industry.

**Conclusion:**

In conclusion, the project has successfully demonstrated the significant impact of data analytics on understanding and predicting trends in Amazon shows and movies. By implementing a comprehensive analytics framework, the study has provided valuable insights that can be leveraged to enhance content strategy, improve audience engagement, and inform marketing decisions.

The strong correlation between content genre and viewer ratings underscores the necessity for a diverse and well-curated content library that caters to various audience preferences. The findings also emphasize the importance of tailoring content duration to match genre-specific viewer expectations, which can lead to increased viewer satisfaction and retention. The predictive models developed in this project, while already providing valuable insights, can be further refined and expanded to improve their accuracy and applicability. Integrating additional variables such as demographic data, social media engagement, and cross-platform comparisons can enhance the predictive power of these models, leading to more precise forecasts and better decision-making.

Future research should focus on deepening the analysis of audience engagement, exploring comparative studies with other streaming platforms, and incorporating innovative content formats like interactive media and virtual reality. By continuing to build on the insights gained from this project, Amazon can further enhance its content offerings and ensure a more engaging and satisfying experience for its viewers.

**CHAPTER- VIII**

**REFERENCES**

1. **Data from Amazon API**

<https://rapidapi.com/collection/amazon-products>

1. **IMDb Dataset**

<https://www.kaggle.com/datasets/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews>   
**3. Academic Papers on Data Analytics**

Provost, F., & Fawcett, T. (2013). Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking. O'Reilly Media. Link

1. **Relevant Articles on Content Trends**

 “Streaming Wars: How Platforms Are Competing for Viewer Attention" by Variety. Link

 “Trends in Digital Content Consumption in 2024" by TechCrunch. Link

 “The Future of Streaming: Predictions and Trends" by Forbes. Link

1. **Exploratory data analytics**

<https://www.geeksforgeeks.org/what-is-exploratory-data-analysis/>

1. **Jupyter notebook tutorial**

<https://www.dataquest.io/blog/jupyter-notebook-tutorial/>

**Books:**

**"Python for Data Analysis"** by Wes McKinney

**CHAPTER- IX**

**APPENDIX**

**SOURCE CODE**

**Show and Movie Ratings:**

import matplotlib.pyplot as plt

import seaborn as sns

# Create a DataFrame with the count of each rating

data\_count1 = data['rating'].value\_counts().reset\_index()

# Initialize a figure with a specific size

plt.figure(figsize=(16, 6))

# Create a count plot for 'rating', distinguishing by 'type'

sns.countplot(x='rating', data=data, hue='type', order=data['rating'].value\_counts().index)

# Rotate the x-axis labels for better readability

plt.xticks(rotation=90)

# Add title and axis labels to the plot

plt.title('Distribution of Show Rating')

plt.xlabel('Rating')

plt.ylabel('Number of Shows')

# Display the plot

plt.show()

**Genre Preferences:**

**import matplotlib.pyplot as plt**

**import seaborn as sns**

**# Extract top 10 genres based on the number of movies listed in each**

**top\_genres = data['listed\_in'].value\_counts().head(10)**

**# Initialize a figure with a specific size**

**plt.figure(figsize=(12, 6))**

**# Create a bar plot for the top 10 genres**

**sns.barplot(x=top\_genres.values, y=top\_genres.index, palette='muted')**

**# Add title and axis labels to the plot**

**plt.title('Top 10 Genres')**

**plt.xlabel('Number of Movies')**

**plt.ylabel('Genre')**

**# Display the plot**

**plt.show()**

**Viewership Patterns:**

import matplotlib.pyplot as plt

import seaborn as sns

# Define the specific order of ratings

rating\_order = ['TV-Y', 'TV-Y7', 'TV-Y7-FV', 'TV-G', 'TV-PG', 'TV-14', 'TV-MA', 'G', 'PG', 'PG-13', 'R', 'NC-17', 'UR']

# Initialize a figure with a specific size

plt.figure(figsize=(12, 8))

# Create a stacked bar plot for content type by rating

sns.countplot(x='rating', hue='type', data=df, order=rating\_order, palette='viridis')

# Add title and axis labels to the plot

plt.title('Stacked Bar Plot for Content Type by Rating')

plt.xlabel('Rating')

plt.ylabel('Count')

# Customize and position the legend

plt.legend(title='Type', loc='upper right')

# Display the plot

plt.show()