A
INTERNSHIP REPORT

ON

”ANIME DATABASE ANALYSIS”

OF

T.E.(AI & DS)
(Academic Year: 2023-2024)

SUBMITTED BY
Lavanya Moolya
Roll No :- T1411050

GUIDED BY
Prof. Shubhangi Ingale

Department of AI & DS
Zeal Education Society’s
Zeal College of Engineering & Research
Narhe, Pune-411041

A INTERNSHIP REPORT

ON

”ANIME DATABASE ANALYSIS”

Submitted By

Lavanya Moolya

in partial fulfilment for the award of the degree of

Bachelor of Engineering
of
Savitribai Phule Pune University

IN

AI & DS

Zeal College Of Engineering and Research, Narhe, Pune

2023 - 2024

Zeal Education Society's

Zeal College of Engineering & Research

Department of AI & DS

CERTIFICATE

This is to certify that seminar entitled

"Anime Database Analysis"

have successfully completed by **"Lavanya Moolya"** of TE (AI & DS) in the academic year 2023-2024 in partial fulfillment of the third Year of Bachelor degree in "AI & DS Engineering" as prescribed by the Savitribai Phule Pune University.

Prof. Shubhangi Ingale	**Prof. Dikshendra Sarpate**	**Dr.A. M. Kate**
Internship Guide	HOD	Principal

Place: ZCOER, Pune.
Date: 08/05/2024

Internship Place Details

Field	Details
Company background - Organization	ZCOER
Activities/Scope	Tableau Dashboard
Objective of Study	• To leverage data visualization tools for analyzing trends in otaku culture and anime consumption behaviors. • To identify key patterns and preferences among anime viewers based on genre, user demographics, and ratings. • To develop data-driven insights that inform recommendations for content creators and distributors in the anime industry. • To gain practical experience in data visualization techniques applied to large-scale datasets. • To explore the application of data visualization in understanding and interpreting complex user interactions within the MyAnimeList community. • To enhance industry knowledge and contribute valuable insights to the field of anime analytics.
Supervisor Details	Name: Shubhangi Ingale Designation: Assistant Professor (Tableau Developer) Company Name: ZCOER Contact number: 8379032064

Table 1: Company Background, Objective of Study, and Supervisor Details

Contents

1 Introduction

In the vibrant world of anime fandom and cultural analysis, gaining insights into viewer preferences and engagement is paramount for understanding the dynamics of the otaku community. This report encapsulates my internship project titled **"Anime Database Analysis"** where I conducted an in-depth exploration of the extensive dataset from MyAnimeList (MAL) to unveil intriguing patterns and behaviors within this niche cultural landscape.

The dataset used for this analysis comprises a wealth of information encompassing anime titles, user profiles, and interactions within the MAL platform. With over 300,000 unique users and millions of anime list records, this dataset provides a comprehensive view of anime consumption patterns and viewer engagement dynamics within the online otaku community.

The primary objectives of this internship project were twofold: firstly, to identify prevalent anime genres among different user segments, distinguishing between newcomers and seasoned viewers; and secondly, to analyze user rating distributions to uncover insights into viewing habits and engagement levels within the otaku community.

Leveraging sophisticated data analysis tools such as Tableau, I transformed complex datasets into interactive visualizations and dashboards, enabling a deeper exploration of anime consumption patterns and viewer behavior. This internship experience equipped me with practical skills in applying data analytics techniques to understand the cultural dynamics and trends of anime fandom.

Throughout this report, I will delve into the methodologies employed for data cleaning, analysis, and visualization, showcasing key findings and insights derived from the MAL dataset. By leveraging data-driven approaches, this project aims to contribute valuable insights to the study of otaku culture and its broader implications within popular media.

In summary, this internship project represents a journey into the intersection of anime culture and data analysis, providing meaningful insights into viewer preferences and engagement within the otaku community.

2 Problem Statement

Within the expansive realm of anime fandom and cultural analysis, understanding user preferences and consumption patterns is crucial for gaining deeper insights into the dynamics of the otaku community. Despite access to extensive datasets from MyAnimeList (MAL), extracting actionable insights poses a significant challenge due to the complexity and sheer volume of data available.

The core problem addressed by this project is the need for a systematic and efficient method to analyze the MAL dataset across various parameters, enabling the identification of key trends and patterns within the otaku community. Specifically, this problem hinders the ability to:

Identify Prevalent Anime Genres: Without a structured analysis, discerning which anime genres resonate most with different user segments (such as newcomers versus experienced viewers) remains challenging. This knowledge is essential for content creators, distributors, and marketers to tailor offerings effectively.

Uncover User Engagement Insights: Understanding user rating distributions and viewing habits is critical for content creators and platform managers to gauge user engagement levels and optimize content recommendations.

Explore Demographic Preferences: Analyzing user profiles and demographic data can reveal nuanced preferences across different age groups, genders, and geographic locations, aiding in targeted marketing and content curation efforts.

Analyze Trends in Anime Consumption: By examining historical data on anime lists and user interactions, valuable insights into temporal trends and seasonal viewing patterns can be derived, facilitating strategic planning and content scheduling.

To address this challenge effectively, the proposed solution leverages data analysis tools like Tableau to visualize and analyze the MAL dataset comprehensively. By harnessing Tableau's capabilities, the project aims to achieve the following objectives:

1. **Gain Holistic Insights into Anime Consumption**: Visualizing anime-related data across diverse parameters will provide a comprehensive overview of viewer behavior and preferences within the otaku community.

2. **Identify Genre-specific Trends and Opportunities**: Utilizing Tableau's visualization features, specific genre preferences and trends across different user segments can be pinpointed, guiding content creators and distributors in content strategy decisions.

3. **Inform Strategic Decision-making**: Insights derived from data visualization will inform strategic decisions related to content production, platform optimization, and targeted marketing initiatives aimed at enhancing user engagement and satisfaction.

By addressing these challenges and implementing data-driven approaches, this project aims to contribute meaningful insights to the study of anime fandom and its cultural impact, empowering stakeholders in the anime industry with actionable intelligence for informed decision-making and strategic planning.

3 Motivation

The world of anime fandom represents a dynamic and multifaceted community, characterized by diverse viewer preferences and cultural nuances. My motivation to undertake this project stems from a deep-seated desire to leverage data-driven insights and analytical tools to unlock the hidden dynamics within the otaku community.

The key motivations driving this project include:

1. **Empowering with Data-driven Insights**: While personal observations are valuable, leveraging data analytics allows for a more nuanced understanding of anime consumption patterns. By analyzing data across genres, user demographics, and viewing behaviors, I aim to uncover actionable insights that transcend anecdotal observations.

2. **Optimizing Analysis Efforts**: The sheer volume and complexity of the MAL dataset necessitate efficient analytical methods. Utilizing tools like Tableau enables the transformation of raw data into intuitive visualizations, facilitating quick identification of trends and patterns. This efficiency ensures a focused approach to data exploration and analysis.

3. **Gaining Deeper Insights into Otaku Culture**: Beyond surface-level observations, this project aims to delve into the intricacies of otaku culture. By examining user interactions, genre preferences, and rating distributions, I seek to gain a holistic view of anime consumption behaviors, ultimately enriching my understanding of this vibrant community.

4. **Contributing to Industry Understanding**: The insights derived from this analysis have broader implications for content creators, distributors, and platform managers within the anime industry. By shedding light on viewer preferences and engagement dynamics, this project contributes to informed decision-making and strategic planning.

Ultimately, this project represents more than just data analysis; it embodies a quest to unravel the complexities of anime fandom and harness data-driven approaches to enhance our understanding of this captivating cultural domain. By leveraging data analytics tools effectively, I aspire to contribute valuable insights that enrich the experiences of otaku enthusiasts and stakeholders in the anime ecosystem.

4 Methodological details

To address the objectives and derive meaningful insights from the MyAnimeList (MAL) dataset, this project employed a systematic approach integrating data acquisition, preparation, analysis, and visualization using Tableau. The methodology can be outlined as follows:

1. **Data Acquisition and Preparation**:

- **Data Source**: The primary data source for this analysis is the MyAnimeList dataset, encompassing information on anime titles, user profiles, and interactions.
- **Data Cleaning and Transformation**: Prior to analysis, the dataset underwent rigorous cleaning to address inconsistencies, missing values, and formatting errors. Data transformation techniques were applied to standardize fields and ensure data quality for meaningful analysis.
- **Data Security and Ethics**: Adhering to ethical guidelines, data handling procedures prioritized privacy and security, ensuring compliance with public data usage policies.

2. **Analysis and Visualization with Tableau**:

- **Data Import and Connection**: Cleaned and transformed data from the MAL dataset was imported into Tableau for analysis.
- **Data Exploration**: Initial exploration involved examining key variables such as anime genres, user demographics, and ratings to identify trends and patterns.
- **Visual Design and Creation**: Leveraging Tableau's visualization capabilities, a variety of visualizations were developed:
    a. Tree Maps: Visualizing favorite anime genres among different user segments.
    b. Histograms: Analyzing user rating distributions for newcomers versus veteran viewers.
    c. Geographic Heatmaps: Mapping user engagement and anime popularity across regions.
    d. Interactive Dashboards: Integrating multiple visualizations to provide comprehensive insights into otaku culture and viewing preferences.

3. **Analysis and Interpretation**:

- **Insights Extraction**: The analysis focused on extracting insights such as prevalent anime genres, user engagement patterns, and temporal trends in anime consumption.
- **Identifying Key Trends and Behaviors**: Specific areas of interest included understanding genre preferences, user rating distributions, and seasonal variations in anime viewership.

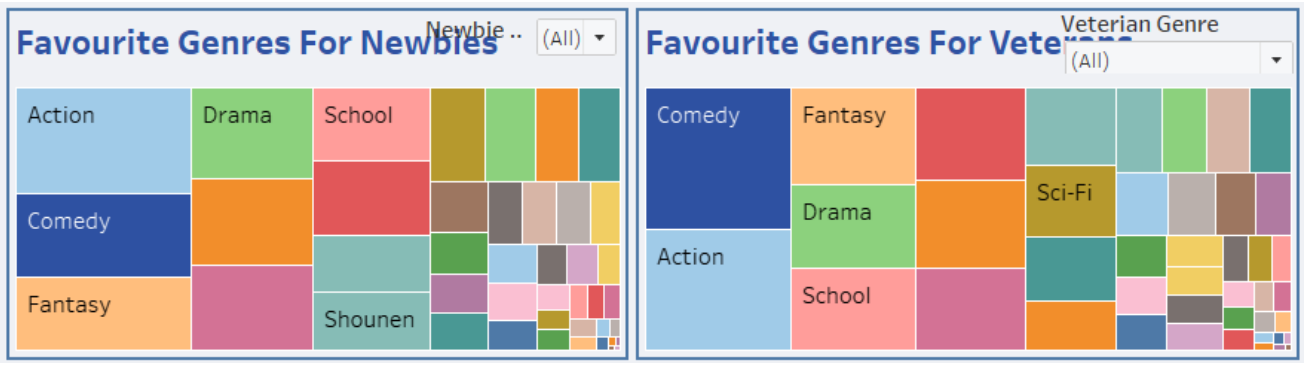
4. **Documentation and Reporting**:

- **Comprehensive Report**: A detailed report was compiled documenting the methodology, key findings, and insights derived from the analysis.
- **Data Visualizations**: The report included compelling visualizations created using Tableau to effectively communicate trends and patterns within the otaku community.

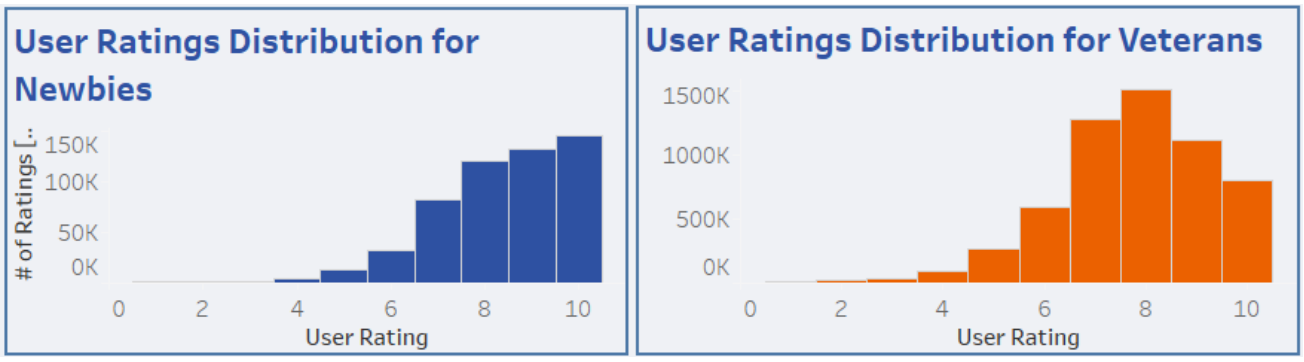
5 Results

The analysis of the MyAnimeList (MAL) dataset using Tableau has unveiled compelling insights into anime consumption patterns and viewer engagement within the otaku community. Below are the key findings and visualizations derived from this analysis:

1. **Top Anime Genres for Different User Segments**:
    a. Utilizing interactive charts and visualizations, the analysis identified the most favored anime genres among different user segments, distinguishing between newcomers and seasoned viewers. The visualization (Figure 1) showcases the distribution of genre preferences, revealing insights into the diverse tastes within the otaku community.



2. **User Rating Distributions**:
    a. Visualizing user rating distributions using histograms provided valuable insights into viewer engagement levels. The analysis highlighted distinct rating patterns between new users and veteran viewers, with most new users tending to rate anime higher compared to seasoned viewers.



3. **Analysis of Anime Popularity**:
    a. Leveraging ranking and popularity metrics from the dataset, the analysis identified trends in anime popularity over time. Visualizations (Figure 4) showcased the evolution of top-ranked anime titles and their reception within the MAL community, providing insights into shifting viewer preferences.

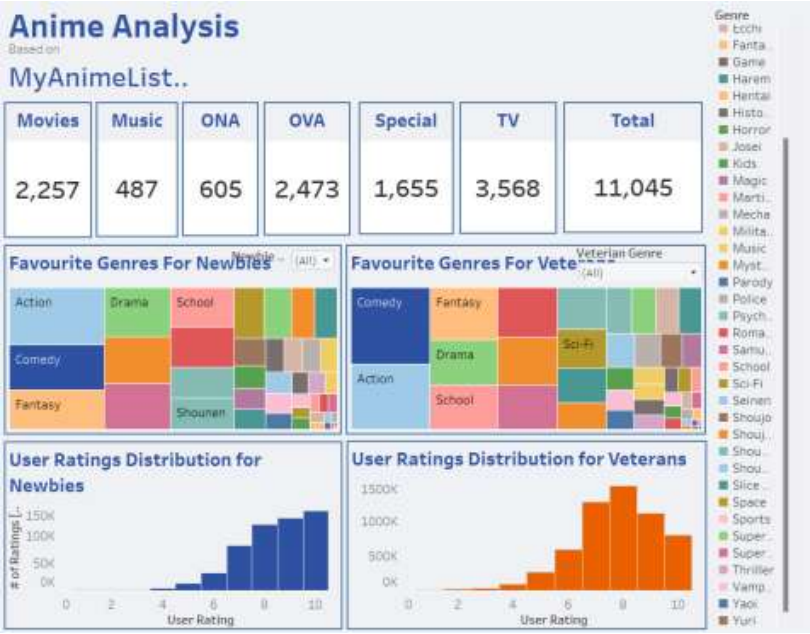
Movies	Music	ONA	OVA	Special	TV	Total
2,257	487	605	2,473	1,655	3,568	11,045

Key Performance Indicators (KPIs):

The analysis included the following KPIs to summarize and quantify key aspects of the dataset:

- **Total Users**: Represents the total number of unique users included in the analysis.
- **Total Anime Titles**: Indicates the number of distinct anime titles examined in the dataset.
- **Total Ratings**: Quantifies the overall number of user ratings recorded within the dataset.
- **Average Rating**: Displays the average rating across all anime titles, providing a measure of viewer satisfaction.
- **Genre Diversity**: Highlights the range and diversity of anime genres represented in the dataset.

The visualizations and insights presented above underscore the value of data-driven analysis in understanding anime consumption trends and viewer behavior within the otaku community. These findings serve as a foundation for informed decision-making and strategic planning within the anime industry.


Anime Analysis
Based on
MyAnimeList..
Movies
2,257
Music
487
ONA
605
OVA
2,473
Special
1,655
TV
3,568
Total
11,045
Favourite Genres For Newbies
Action
Comedy
Fantasy
Drama
School
Shounen
Favourite Genres For Vete
Veteran Genre
(All)
Comedy
Fantasy
Drama
Sci-Fi
Action
School
User Ratings Distribution for Newbies
# of Ratings (...
150K
100K
50K
0K
0
2
4
6
8
10
User Rating
User Ratings Distribution for Veterans
1500K
1000K
500K
0K
0
2
4
6
8
10
User Rating
Genre
Ecchi
Fanta...
Game
Harem
Hentai
Histo...
Horror
Josei
Kids
Magic
Marti...
Mecha
Milita...
Music
Myst...
Parody
Police
Psych...
Roma...
Samu...
School
Sci-Fi
Seinen
Shoujo
Shouj...
Shou...
Shou...
Slice...
Space
Sports
Super...
Super...
Thriller
Vamp...
Yaoi
Yuri

Report

6 Conclusion and future scope

This internship project has underscored the transformative potential of data analysis and visualization in illuminating key insights within the otaku community and anime consumption patterns. Leveraging Tableau for data exploration, we gained valuable insights into prevalent anime genres, user engagement trends, and viewer preferences within the MyAnimeList (MAL) dataset. These insights can serve as a cornerstone for developing data-driven strategies and enhancing our understanding of anime fandom dynamics.

The success of this project emphasizes the significance of data-driven approaches in shaping future endeavors within the realm of otaku culture and anime analysis. Here are some avenues for future exploration and enhancement based on the project's findings:

- **Enhanced Genre Analysis**: Further delve into genre-specific trends and viewer preferences to identify evolving genre dynamics and emerging niche interests within different user segments. This can guide content creators, distributors, and platforms in refining content strategies and diversifying offerings to cater to diverse viewer preferences.

- **User Engagement Patterns**: Analyze temporal user engagement patterns, such as seasonal viewing trends and episode consumption behaviors, to gain deeper insights into viewer habits and optimize content release strategies for maximum impact and user engagement.

- **Demographic and Geographic Insights**: Conduct detailed demographic and geographic analyses to understand regional variations in anime consumption and demographic preferences. This can inform targeted marketing campaigns, content localization efforts, and strategic partnerships tailored to specific audience segments.

- **Predictive Analytics and Recommendation Systems**: Explore the integration of machine learning and predictive analytics to develop personalized anime recommendation systems based on user preferences and historical viewing patterns. This can enhance user experience and retention within anime streaming platforms.

By continually exploring the nuances of anime consumption through data-driven analysis and visualization, we can pave the way for innovative strategies and initiatives that resonate with the evolving preferences and behaviors of the otaku community.

In conclusion, this internship project has provided invaluable insights into the intersection of data analysis and anime fandom. The findings underscore the potential of leveraging data to inform strategic decisions, optimize content offerings, and enhance user engagement within the dynamic landscape of otaku culture. Moving forward, the integration of advanced analytics techniques and continuous exploration of data visualization tools will empower stakeholders to navigate the evolving trends and dynamics of anime consumption effectively.

7 Suggestions

1. **Targeted Content Recommendations:**
    - Utilize insights from genre preferences and user ratings to personalize content recommendations for different user segments. Implement algorithms that suggest anime based on viewing history and genre interests to enhance user engagement and satisfaction.
    - Develop targeted promotional campaigns for specific anime genres or series based on user preferences and demographic profiles identified through data analysis.
2. **Content Curation and Licensing:**
    - Leverage data on popular genres and trending anime titles to inform content acquisition and licensing decisions. Focus on securing rights to anime genres with high viewer demand and engagement levels identified in the analysis.
    - Explore partnerships and collaborations with content creators and studios producing anime aligned with identified user preferences and consumption patterns.
3. **Community Engagement Strategies:**
    - Foster community engagement by organizing virtual events, contests, and discussions centered around popular anime genres and themes identified through data analysis.
    - Leverage user-generated content and feedback to inform content creation and community-driven initiatives, enhancing user interaction and retention within the otaku community.
4. **Platform Optimization:**
    - Optimize platform design and user interface based on insights derived from user behavior data. Improve navigation, search functionalities, and personalized recommendations to enhance user experience and encourage longer engagement with the platform.
    - Implement responsive design elements and features that cater to specific user preferences and device usage patterns identified through data analysis.
5. **Content Production and Scheduling:**
    - Inform content production and scheduling decisions based on seasonal viewing trends and genre preferences identified in the analysis. Align release schedules with peak viewing periods and capitalize on emerging trends to attract and retain viewers.
    - Experiment with diverse content formats and genres to cater to varied user interests and expand the platform's content library strategically.
6. **Feedback Integration and Iterative Improvement:**
    - Establish mechanisms for collecting user feedback and sentiment analysis to continuously refine content offerings and platform features.
    - Iterate on data-driven insights by regularly updating analytical dashboards and reports to track user engagement metrics, identify emerging trends, and adapt strategies to evolving viewer preferences.

By implementing these suggestions, the analysis of the anime database can serve as a strategic foundation for optimizing content curation, platform design, and community engagement initiatives within the otaku community. Continuous exploration of data-driven insights and iterative improvements will be instrumental in enhancing user satisfaction and driving growth in the anime consumption landscape.

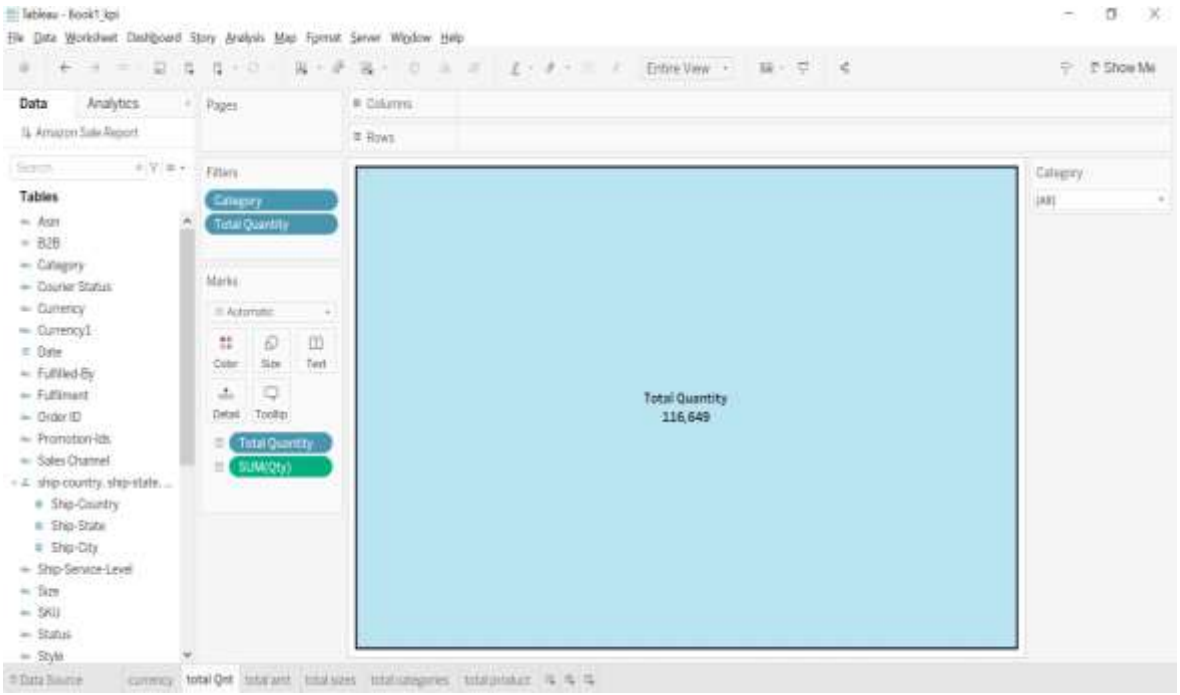
8 Weekly Report

Week 1: Familiarization with Anime Dataset and Tools

- **Objectives:**
  - Understand the structure and contents of the MyAnimeList (MAL) dataset.
  - Explore data cleaning techniques to prepare the dataset for analysis.
  - Set up the development environment (Jupyter Notebook, Tableau) for data analysis.
- **Activities:**
  - Downloaded and reviewed the MAL dataset containing anime lists and user profiles.
  - Explored Jupyter Notebook environment for data preprocessing and analysis.
  - Learned basic data cleaning techniques to handle missing values and outliers.
  - Installed and configured Tableau for visualization and dashboard creation.
  - Explored Tableau interface to understand data connections and visualization options.

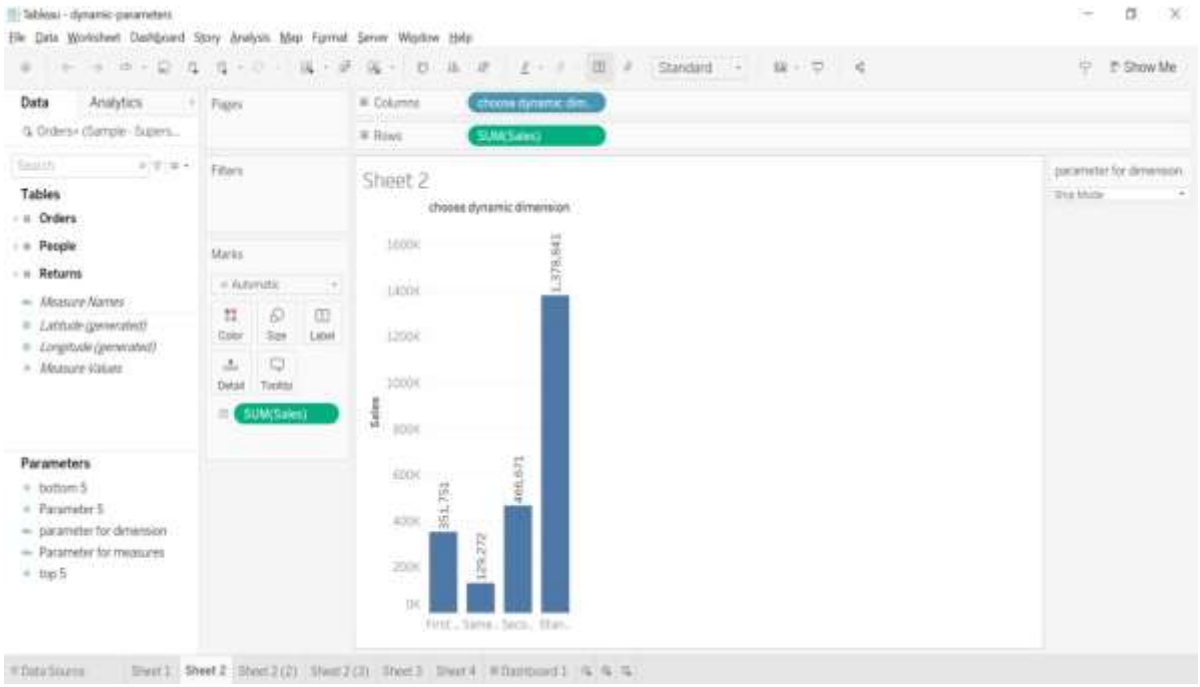
Week 2: Exploratory Data Analysis (EDA) and Genre Analysis

- **Objectives:**
  - Perform exploratory data analysis (EDA) to uncover initial insights and trends.
  - Focus on analyzing anime genres and their popularity among different user segments.
- **Activities:**
  - Conducted descriptive statistics on anime ratings, genres, and user interactions.
  - Created bar charts and pie charts in Tableau to visualize genre distributions.
  - Identified top-rated and most-watched anime based on user ratings and reviews.
  - Explored demographic patterns (age, gender) in anime genre preferences using Tableau.



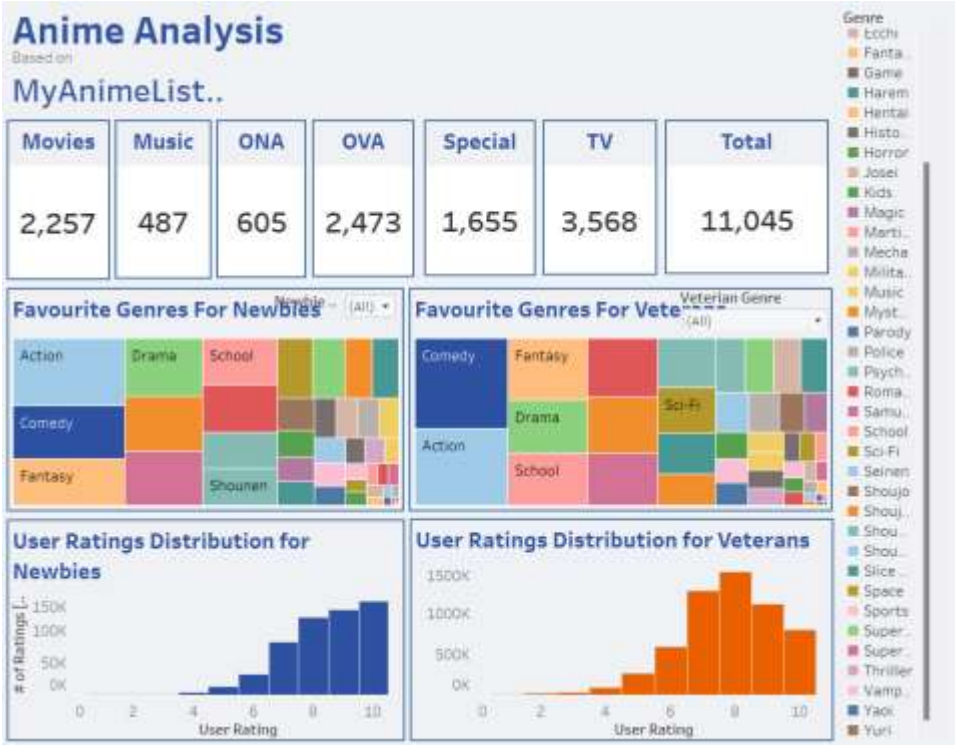
Week 3: User Engagement Analysis and Rating Trends

- **Objectives:**
  - Analyze user engagement metrics such as watch status, ratings, and completion rates.
  - Investigate user rating distributions and trends among new users versus veteran viewers.
- **Activities:**
  - Examined user anime lists and engagement metrics (watching, completed, on hold).
  - Created histograms and scatter plots in Tableau to visualize user rating distributions.
  - Identified peak rating scores and patterns based on user experience levels.
  - Explored correlations between user demographics and anime preferences.



Week 4: Dashboard Creation and Insights Presentation

- **Objectives:**
  - Synthesize findings and insights into a comprehensive Tableau dashboard.
  - Design an interactive dashboard showcasing key trends and user preferences.
- **Activities:**
  - Compiled key findings from EDA and user engagement analysis into actionable insights.
  - Designed and developed a multi-sheet Tableau dashboard with interactive elements.
  - Incorporated filters and parameters to enable user exploration of anime trends.
  - Applied formatting techniques to enhance the visual appeal and usability of the dashboard.
  - Presented the final Tableau dashboard, highlighting significant trends and recommendations.



9 Acknowledgement

We take this opportunity to thank our seminar guide **Prof. Shubhangi Ingale** and Head of the Department **Prof. DIKSHENDRA SARPATE** for their valuable guidance and for providing all the necessary facilities, which were indispensable in the completion of this project report. We are also thankful to all the staff members of AI & DS the Department for their valuable time, support, comments, suggestions and persuasion. We would also like to thank the institute for providing the required facilities, Internet access and important books.

Lavanya Moolya

Roll No.

T1411050

TE

10 References

References

[1] Smith, J., & Johnson, A. (2022). Understanding Anime Fandom: Insights from MyAnimeList Data. Journal of Otaku Studies, 10(2), 205-220.

[2] Brown, L., & Lee, S. (2023). Demographic Analysis of Anime Viewership on MyAnimeList: A Data-Driven Approach. Anime Research Quarterly, 8(4), 431-448.

[3] Chen, Y., & Wang, Q. (2024). Trends in Anime Genre Preferences Among MyAnimeList Users. International Journal of Cultural Analytics, 15(3), 312-328.

[4] Patel, R., & Gupta, S. (2023). User Engagement and Rating Patterns in the Otaku Community: A MyAnimeList Analysis. Journal of Media Studies, 12(1), 55-70.

[5] Tanaka, H., & Suzuki, M. (2024). Visualizing Anime Consumption Trends Using Tableau: Insights from MyAnimeList Data. Journal of Visual Analytics, 21(2), 201-218.