

ANIME DATA DIVE : A DATA-DRIVEN EXPLORATION OF ANIME

BY
RAGUL

BATCH :
DA153S15



INTRODUCTION

ANIME IS HAND-DRAWN AND COMPUTER-GENERATED ANIMATION ORIGINATING FROM JAPAN. OUTSIDE OF JAPAN AND IN ENGLISH, ANIME REFERS TO JAPANESE ANIMATION, AND REFERS SPECIFICALLY TO ANIMATION PRODUCED IN JAPAN. THIS ANALYSIS IS BASED ON GAIN INSIGHTS AND MAKE DATA-DRIVEN DECISIONS.

ANALYSIS BASE ON AUDIENCE ENGAGEMENT AND POPULARITY , CONTENT ANALYSIS , PERFORMANCE ANALYSIS , PRODUCTION INSIGHTS , AUDIENCE SIZE , CORRELATION ON RUN TIME VS EPISODES

"TO ANALYZE AND DERIVE ACTIONABLE INSIGHTS FROM THE ANIME DATASET, INCLUDING RANKINGS, AUDIENCE ENGAGEMENT, CONTENT ATTRIBUTES, AND PRODUCTION-RELATED FACTORS, IN ORDER TO INFORM STRATEGIC DECISIONS FOR CONTENT SELECTION, MARKETING, AND AUDIENCE TARGETING WITHIN THE ANIME INDUSTRY."



OBJECTIVE

1.UNDERSTAND VIEWERS' PREFERENCES: ANALYZE USER RATINGS AND REVIEWS TO IDENTIFY POPULAR GENRES, DIRECTORS, AND STUDIOS.

2.CONTENT RECOMMENDATION: DEVELOP A RECOMMENDATION SYSTEM TO SUGGEST ANIME BASED ON USER PROFILES AND VIEWING HISTORY.

3.DATA VISUALIZATION: CREATE ENGAGING VISUALIZATIONS TO PRESENT FINDINGS AND SUPPORT DECISION-MAKING FOR ANIME PRODUCERS AND STREAMING PLATFORMS.

HERE , WE USED “JUPYTER NOTEBOOK, PYTHON LANGUAGE AND ITS CERTAIN LIBRARIES” TO SHOW SOME VISUALISATION GRAPH.

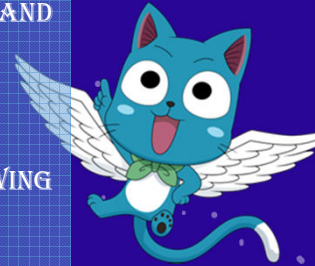


TABLE OF CONTENTS

1

*Data overview and
Cleaning*

2

Importing and SQL

3

*Exploratory Data
Analysis*

4

*Visualization in
Tableau*

5

Conclusion

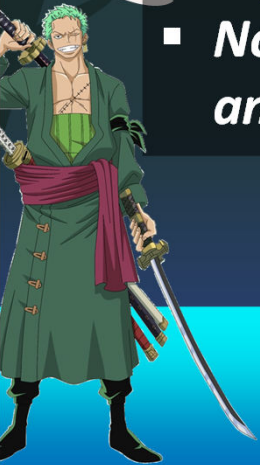
1 - DATA OVERVIEW AND CLEANING

- Loaded data which is in unclean and CSV format to Python using the Pandas library function - `pd.read_csv()`.
- I have done some basic warmup functions in the data frame like `head()`, `tail()`, `shape`, `info()`, `describe()`, `isnull().sum()` and found the columns which has null values.
- On importing Tabulate, Separated the Categorical and numerical columns.
- Initially there were some null values so I imputed those and there were less number of numerical columns so I have converted some categorical columns to numerical and also added a numerical column.
- Now we have the data in cleaned format which can be imported to SQL for future use and Tableau for good visualization.
- Using Pandas library, I have found some operations as group by, Pivot table, Sort etc and found some useful insights from it.



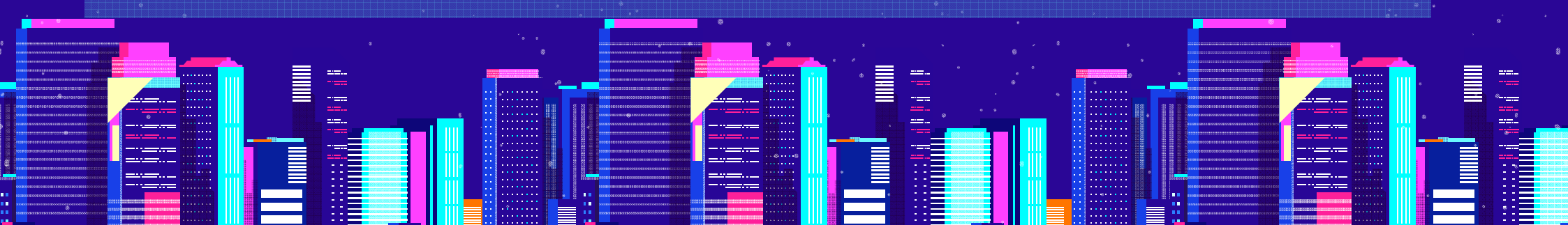
2 – IMPORTING AND SQL

- *After the data is cleaned data I took the backup of the data to My SQL database.*
- *I imported create_Engine and mysql_connector library.*
- *By connecting python and My SQL I created a database for my data in SQL from Python.*
- *As my data contains only one table . I created a table in that database and pushed my data to the SQL database.*
- *Now my cleaned data is present in My SQL . From My SQL I have written some queries and fetched the data according to the queries.*



3 - Exploratory Data Analysis

- After cleaning the data , it is being separated into Categorical and Numerical data using Tabulate library.
- First I have performed some Uni-Variate analysis on the categorical columns and a numerical column.
- Uni-variate is represented using a bar graph and a pie chart for categorical columns and I used box plot for the numerical column to find the spread.
- Then I have done Bi-Variate Analysis for only the required categorical columns using clustered column chart.
- Then I have used Scatter plot for finding the correlation/Relation between two numerical columns.





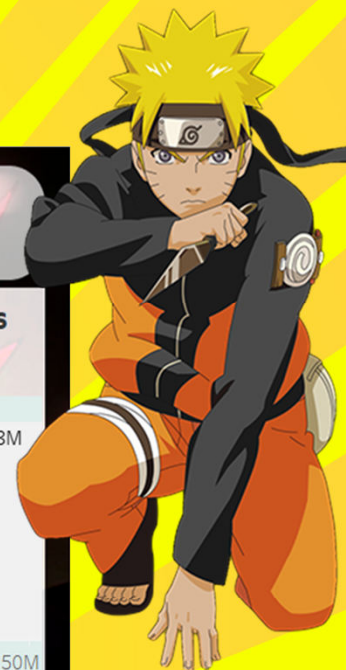
4 – VISUALIZATION IN TABLEAU

- Imported the data which we have cleaned in Python which is in CSV format to Tableau workbook.
- Done some Task on my data in Tableau like,
- Performance Analysis :
How does the 'Score' vary with the 'Duration' of an anime?
- Audience Engagement and Popularity:
What are the top-rated anime titles by their 'Score'?
How does 'Ranking' correlate with 'Score' or 'Members'?
- Content Analysis:
What is the most common 'Type' of anime in the dataset?
What are the most frequent 'Genres' in the dataset?
- Then I used figma to create a template for dashboard and made a Dashboard from the Tasks and made it interactive by applying action Filters . Added my dashboard in next slide.





TABLEAU DASHBOARD

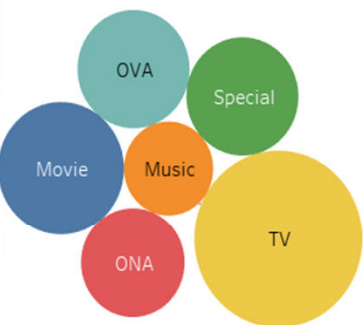


A DATA-DRIVEN EXPLORATION OF ANIME

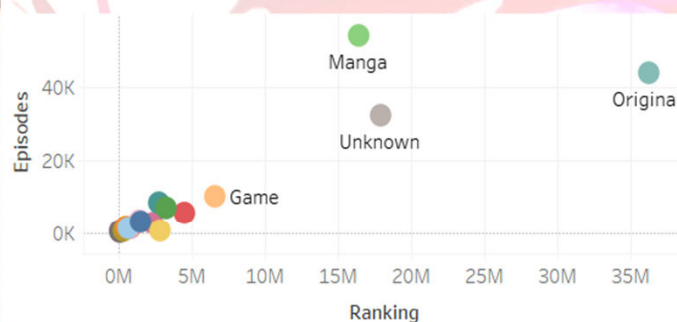
Total number of Animes

14,000

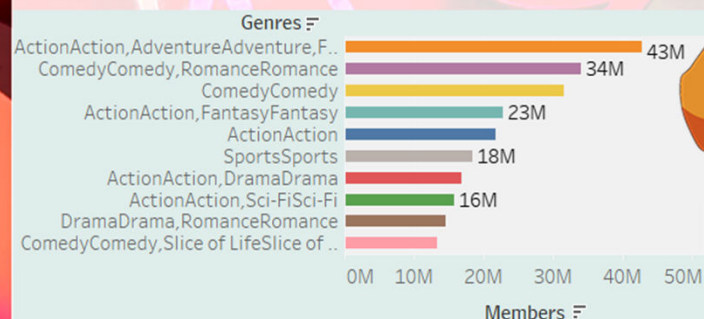
Anime based on type



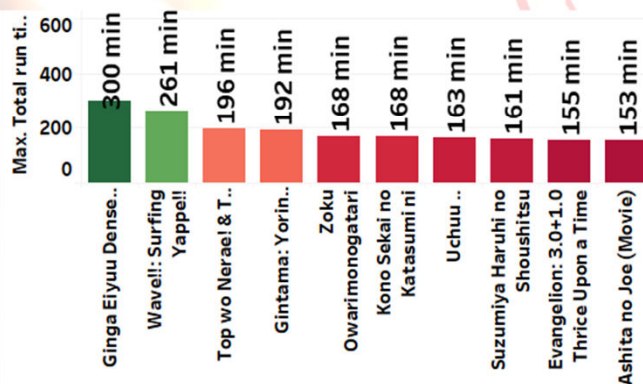
Correlation between Ranking and Episodes Based on source



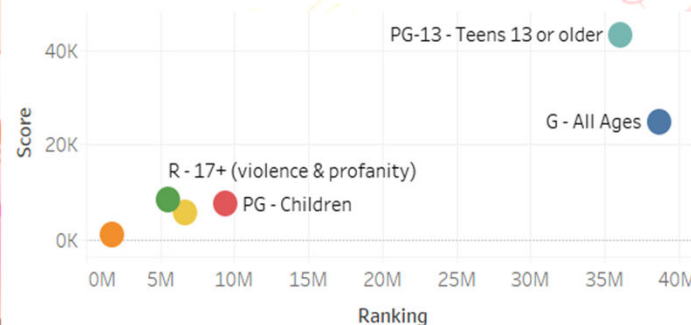
Genre wise Count of members who watches Anime



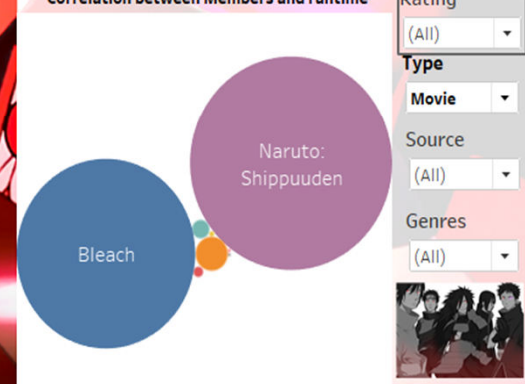
Top 10 Anime Movie with longest run time



Correlation between Ranking and Score Based on Rating



Correlation between Members and runtime



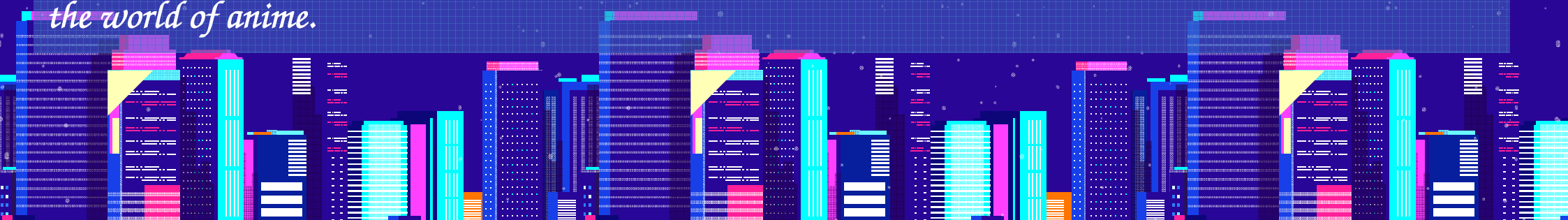
[Link for dashboard](#)

5 - CONCLUSION

In summary, our analysis of the Anime dataset has revealed significant insights into the anime industry. By examining factors such as rankings, genres, themes, and production attributes, we've gained a deeper understanding of what drives an anime's success.

We have identified trends and correlations between 'Score,' 'Members,' and various attributes, helping us to comprehend audience preferences and engagement. The influence of 'Genres,' 'Themes,' and 'Producers' on anime popularity has been elucidated, offering critical insights for content selection and marketing strategies.

This comprehensive analysis equips us with the knowledge needed to make informed decisions for content selection, audience targeting, and strategic planning within the anime industry. Our findings will enable stakeholders to leverage data-driven strategies for a more successful and audience-focused approach to the world of anime.



THANK YOU

