**Anime Analysis Documentation**

**Introduction**

This document presents an in-depth analysis of anime data based on the dataset available up to 2023. The study involves data cleaning, exploratory data analysis (EDA), and key insights into anime trends, ratings, and categories. The analysis aims to provide a better understanding of anime popularity, viewer preferences, and the influence of different factors on ratings.

**Data Import and Cleaning**

* The dataset was imported using pandas for efficient data manipulation.
* Initial dataset inspections included checking the number of rows, columns, missing values, and data types.
* Missing values and inconsistencies were identified, and necessary preprocessing steps such as data imputation and removal of duplicates were performed.
* The dataset was structured to ensure that categorical and numerical data were properly formatted for analysis.

**Exploratory Data Analysis (EDA)**

**1. Name Representation**

* The dataset contains anime titles in multiple languages, reflecting its global reach.
* Distribution of anime titles based on language representation was analysed to understand regional variations.

**2. Category Representation**

* Anime were classified into different genres such as Action, Romance, Adventure, and Fantasy.
* A frequency analysis was conducted to determine the most common genres.
* Visual representation using bar charts illustrated the distribution of genres, providing insights into viewer preferences.

**3. Top Ranked Anime**

* The highest-rated anime were identified based on user ratings.
* Analysis included the average ratings for different categories to determine which genres or formats performed better.
* Trends in high-rated anime were examined to assess common themes and patterns among top-ranked shows.

**4. Anime TV Series vs Movies**

* The dataset was segmented to differentiate between anime TV series and movies.
* A comparative analysis was conducted to understand the differences in ratings, popularity, and audience reception between the two formats.
* Histograms and scatter plots were used to visualize the variations in ratings between TV series and movies.

**5. Anime Based on Manga**

* The dataset was examined to determine how many anime were adaptations of manga.
* A comparative rating analysis was performed to assess whether manga-based adaptations generally receive higher ratings than original anime.
* Additional insights were drawn regarding the influence of source material on the success of an anime.

**Visualizations**

* Various plots and graphs were generated using Matplotlib to better illustrate trends and insights in the dataset.
* Bar charts were used for genre distributions, while scatter plots and histograms helped visualize anime ratings and trends.
* Interactive plots allowed for better exploration of data, showcasing relationships between different anime attributes.

**Conclusion**

* The study revealed key insights into anime trends, popular genres, and the influence of source material on ratings.
* The findings highlight the importance of genre diversity and audience preferences in determining anime success.
* Future research can explore sentiment analysis on user reviews, prediction models for anime success, or content-based recommendations using machine learning.

**Prepared by:** Nirdesh Vishwakarma & Rajshekhar Mehra  
**Date:** 27-02-25