

BIG DATA VISUAL ANALYTICS (CS661)

PROJECT PROPOSAL REPORT

GROUP-09

AnimeLens: Visual Analytics System for Exploring Trends, Genres, and Viewer Preferences

1. Introduction

Anime has evolved from being a niche interest to a global cultural phenomenon, with millions of viewers actively following new releases, discussing their favorites online, and rating shows on platforms like MyAnimeList (MAL). Given this massive popularity, anime data is not only abundant but also rich in valuable insights that can tell engaging stories about evolving trends, genre preferences, and viewer habits.

In this project, our goal is to build a visual analytics system that explores and presents these insights through interactive visualizations. By combining data processing, analytics, and visualization techniques, we aim to create a system where users can visually explore the anime landscape, understand how genre popularity shifts over time, how seasons influence release schedules, and even how studio specializations shape the anime we watch.

2. Data Sources

We will use publicly available anime datasets, primarily sourced from Kaggle and the MyAnimeList API. These datasets include information such as:

- Anime title, type (TV, Movie, OVA, etc.)
- Release year and season
- Number of episodes and average episode duration
- Source material (Manga, Light Novel, Game, Original, etc.)
- Genre tags
- Average rating (MAL Score)
- Number of viewers who rated (Popularity metric)
- Studio name and producer details

Dataset Link : <https://www.kaggle.com/datasets/azathoth42/myanimelist>

3. Specific Tasks

Our system will perform a set of visual analytics tasks designed to extract meaningful stories from the data. These tasks aim to go beyond static visualization and allow interactive exploration, enabling users to filter, select, and focus on specific attributes of interest.

3.1 Genre Popularity Over Time

- Track how different anime genres (Action, Romance, Slice of Life, etc.) evolved in popularity over the years.
- Uncover trends such as the rise of Isekai or the decline of Mecha anime.

3.2 Seasonal Release Patterns

- Analyze the distribution of anime releases by season (Winter, Spring, Summer, Fall).
- Study how seasonal trends impact anime ratings and popularity.

3.3 Studio Specialization and Success

- Identify studios that dominate specific genres.
- Analyze which studios consistently produce top-rated anime.

3.4 Episode Count vs. Popularity

- Investigate whether long-running anime are more popular than shorter seasonal shows.
- Study whether shorter anime tend to receive higher ratings due to tighter storytelling.

3.5 Regional Anime Preferences

- Determine region-wise preferences in anime consumption.
- Analyze watch-time patterns across different countries and genre popularity in various regions.

3.6 Success Prediction Model

- Develop a machine learning model to predict an anime's success based on factors such as genre, episode count, animation studio, and user reviews.

3.7 Genre Co-occurrence Network

- Create a co-occurrence graph to reveal which genres tend to appear together (e.g., Romance + Comedy, Fantasy + Adventure).
- Understand popular formulae in anime storytelling.

4. Overall Solution

We plan to develop a web-based visual analytics system using a combination of backend data processing and interactive visualization techniques.

4.1 Tech Stack

- **Data Processing & Analysis:** Pandas, NumPy
- **Visualization Libraries:** Plotly, Bokeh
- **Dashboard Development:** Dash
- **Machine Learning:** Scikit-Learn, TensorFlow

4.2 Key Features

- Data filtering (filter by year, genre, studio, etc.)
- Interactive tooltips and hover effects
- Dynamic plots that update based on user selections
- Linked visualizations — selecting a genre highlights related studios and top anime in that genre

5. Team Members and Responsibilities

To ensure an efficient workflow, tasks have been categorized into four broad areas, with two team members assigned to each category:

Task Category	Team Members	Responsibilities
Data Collection & Preprocessing	Samarth, Yashanki	Gathering datasets, cleaning, handling missing values, and structuring data for analysis.
Data Visualization & Dashboard Design	Kartik, Sambhav	Developing interactive graphs, charts, and visual representations of anime trends.
Machine Learning & Success Prediction	Keshav, Satmeet	Building predictive models to assess anime success based on key attributes.
Web Development	Vedant, Shishir	Developing the web-based interface and implementing an anime recommendation engine.

6. Expected Outcomes

By the end of the project, we expect to deliver:

- A fully interactive web-based visual analytics dashboard.
 - Clear and visually appealing charts that tell the story of anime evolution.
 - Insights into changing viewer tastes, influential studios, and evolving storytelling patterns.
 - A comprehensive project report explaining our methodology, findings, and design choices.
-