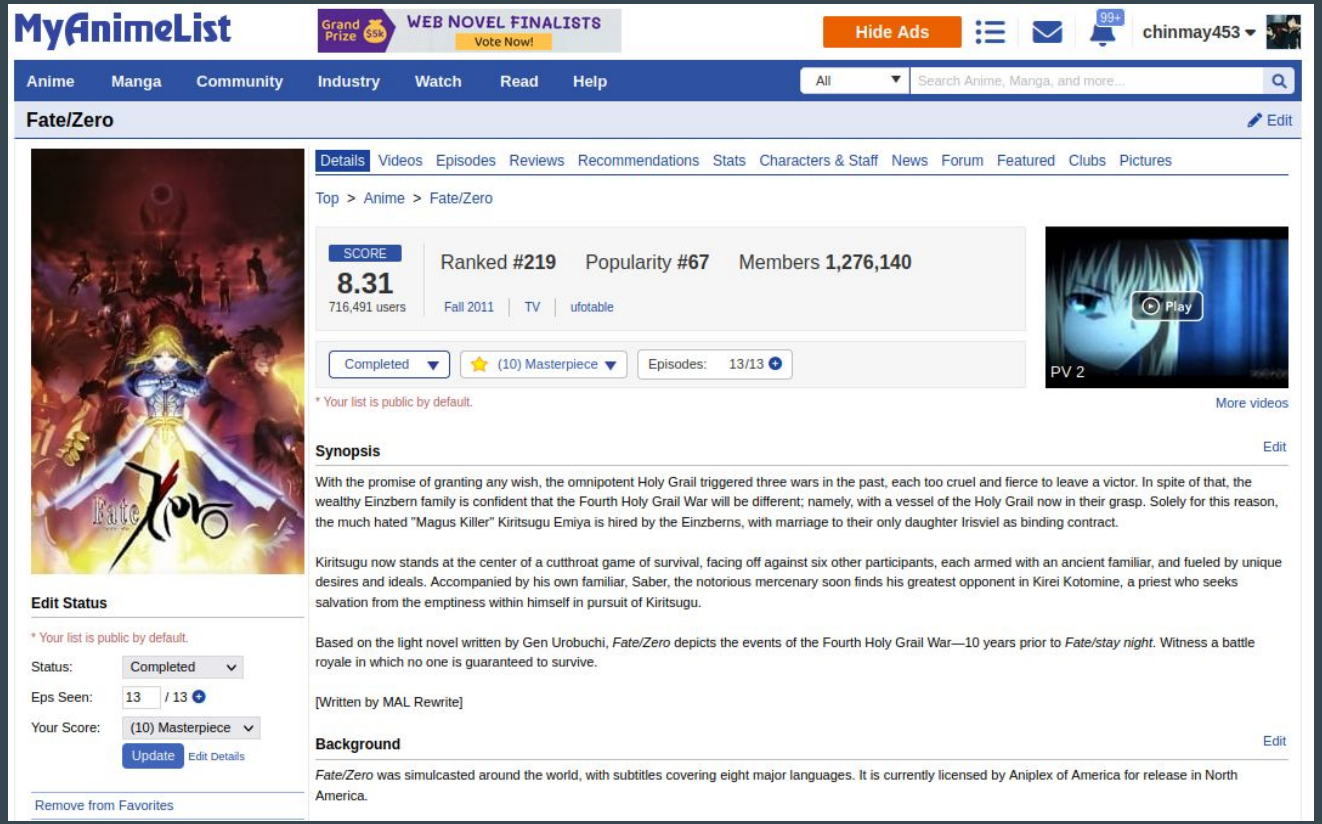
Codagami Anime Recommender

Data Mining Project

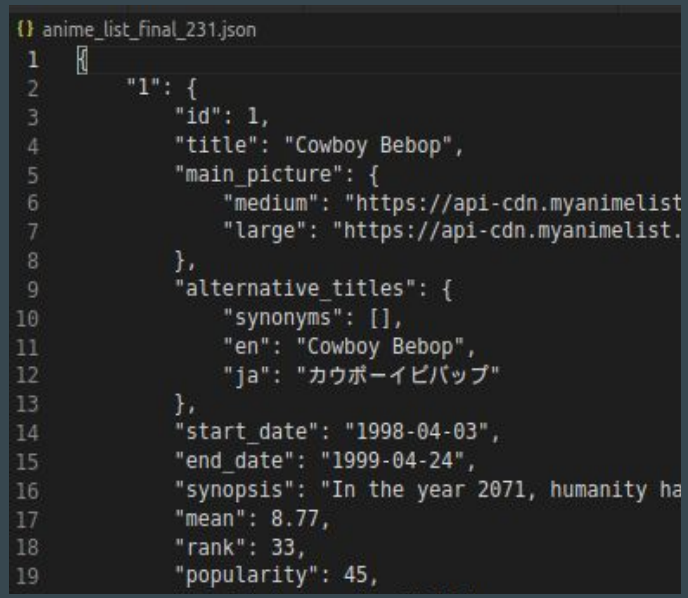
Group-15

Chinmaya Singal	180207	chinmaya@iitk.ac.in
Dipesh Khandelwal	180249	dipeshk@iitk.ac.in
Rythm Agarwal	180636	rythm@iitk.ac.in
Sakshi	180653	sakshisa@iitk.ac.in
Sarthak Dubey	180674	srthkdb@iitk.ac.in

Data Source


MyAnimeList
Grand Prize 35k
WEB NOVEL FINALISTS
Vote Now!
Hide Ads
99+
chinmay453
Anime
Manga
Community
Industry
Watch
Read
Help
All
Search Anime, Manga, and more...
Fate/Zero
Edit
Details
Videos
Episodes
Reviews
Recommendations
Stats
Characters & Staff
News
Forum
Featured
Clubs
Pictures
Top > Anime > Fate/Zero
SCORE
8.31
716,491 users
Ranked #219
Popularity #67
Members 1,276,140
Fall 2011
TV
ufotable
Play
PV 2
More videos
Completed
(10) Masterpiece
Episodes: 13/13
* Your list is public by default.
Edit Status
* Your list is public by default.
Status: Completed
Eps Seen: 13 / 13
Your Score: (10) Masterpiece
Update
Edit Details
Remove from Favorites
Synopsis
Edit
With the promise of granting any wish, the omnipotent Holy Grail triggered three wars in the past, each too cruel and fierce to leave a victor. In spite of that, the wealthy Einzbern family is confident that the Fourth Holy Grail War will be different; namely, with a vessel of the Holy Grail now in their grasp. Solely for this reason, the much hated "Magus Killer" Kiritsugu Emiya is hired by the Einzberns, with marriage to their only daughter Irisviel as binding contract.
Kiritsugu now stands at the center of a cutthroat game of survival, facing off against six other participants, each armed with an ancient familiar, and fueled by unique desires and ideals. Accompanied by his own familiar, Saber, the notorious mercenary soon finds his greatest opponent in Kirei Kotomine, a priest who seeks salvation from the emptiness within himself in pursuit of Kiritsugu.
Based on the light novel written by Gen Urobuchi, Fate/Zero depicts the events of the Fourth Holy Grail War—10 years prior to Fate/stay night. Witness a battle royale in which no one is guaranteed to survive.
[Written by MAL Rewrite]
Background
Edit
Fate/Zero was simulcasted around the world, with subtitles covering eight major languages. It is currently licensed by Aniplex of America for release in North America.

Data retrieval and processing



id	title	main_picture	alternative_titles	start_date	end_date
1.0	Cowboy Bebop	{'medium': 'https://api-cdn.myanimelist.net/im...	{'synonyms': [], 'en': 'Cowboy Bebop', 'ja': '...	1998-04-03	1999-04-24
6.0	Trigun	{'medium': 'https://api-cdn.myanimelist.net/im...	{'synonyms': [], 'en': 'Trigun', 'ja': 'トライガン'}	1998-04-01	1998-09-30
8.0	Bouken Ou Beet	{'medium': 'https://api-cdn.myanimelist.net/im...	{'synonyms': ['Adventure King Beet'], 'en': 'B...	2004-09-30	2005-09-29
16.0	Hachimitsu to Clover	{'medium': 'https://api-cdn.myanimelist.net/im...	{'synonyms': ['HachiKuro', 'Honey & Clover'], ...	2005-04-15	2005-09-27
18.0	Initial D Fourth Stage	{'medium': 'https://api-cdn.myanimelist.net/im...	{'synonyms': ['Initial D 4th Stage'], 'en': "...	2004-04-17	2006-02-18

s × 33 columns

Json Data collected from MAL API

Data converted and processed to pandas Dataframe for analysis

ALS based Recommender

Alternating Least Square (ALS) Matrix Factorization based Collaborative Filtering

Make recommendations

Enter your anime name in my_anime_list, and the recommender will generate recommendations

```
my_anime_list = ['Cowboy Bebop']

recommends = make_recommendation(
    best_model_params={'iterations': 10, 'rank': 20, 'lambda_': 0.05},
    ratings_data=rating_data,
    df_animes=animes,
    fav_anime_list=my_anime_list,
    n_recommendations=10,
    spark_context=sc)

print('Recommendations for {}:'.format(my_anime_list[0]))
for i, title in enumerate(recommends):
    print('{0}: {1}'.format(i+1, title))
```

[15] ✓ 11.6s

...

```
Recommendations for Cowboy Bebop:
1: Fullmetal Alchemist: Brotherhood
2: Hunter x Hunter (2011)
3: Shingeki no Kyojin Season 3
4: Shokugeki no Souma: San no Sara
5: "Violet Evergarden: Kitto ""Ai"" wo Shiru Hi ga Kuru no Darou"
6: Seishun Buta Yarou wa Yumemiru Shoujo no Yume wo Minai
7: Death Note
8: Kimi no Na wa.
9: Saiki Kusuo no Ψ-nan 2
10: Shingeki no Kyojin Season 3 Part 2
```

TF-IDF based Recommender

Term Frequency — Inverse Document Frequency (TF-IDF) based recommendation systems are content based recommenders

```python
def recommend_anime(title, max_reco = 10, cosine_sim = cos_sim, cosine_sim_g
    print("Anime Recommendations for: "+title)
    recommended_animes = []
    index = anime_names[anime_names == title].index[0]

    similar_scores = pd.Series(cosine_sim[index])
    similar_scores_genre = pd.Series(cosine_sim_genre[index])
    mean = anime_df['mean']
    mean = mean.apply(lambda x: x if x >= 7 else 0)
    anime_mean_score = pd.Series(np.array(mean))

    similar_scores_mul = similar_scores.mul(similar_scores_genre)

    similar_scores_mul = similar_scores_mul.mul(anime_mean_score)
    similar_scores_mul = similar_scores_mul.sort_values(ascending=False)

    top_animes = list(similar_scores_mul.iloc[1:max_reco+1].index)
    for anime_index in top_animes:
        anime_row = anime_df.iloc[anime_index]
        anime_name = anime_row['title']
        recommended_animes.append(anime_name)
    return recommended_animes


recommend_anime("Cowboy Bebop")
```

✓ 56.1s

```
Anime Recommendations for: Cowboy Bebop
['Cowboy Bebop: Yose Atsume Blues',
 'Seihou Bukyou Outlaw Star',
 'Cowboy Bebop: Tengoku no Tobira',
 'Space Adventure Cobra',
 'Waga Seishun no Arcadia: Mugen Kidou SSX',
 'Ginga Tetsudou Monogatari',
 'Iria: Zeiram The Animation',
 'Uchuu Senkan Yamato',
 'Freedom',
 'Sayonara Ginga Tetsudou 999: Andromeda Shuuchakueki']
```

KNN based Recommender

K- Nearest Neighbours (KNN) item based collaborative filtering

```
    reverse_mapper = {v: k for k, v in mapper.items()}
    print('Recommendations for {}:'.format(fav_anime))
    for i, (idx, dist) in enumerate(raw_recommends):
        print('{0}: {1}, with distance of {2}'.format(i+1, reverse_mapper[idx], dist))
```

✓ 0.1s

```
my_favorite = 'Cowboy Bebop'

make_recommendation(
    model_knn=model_knn,
    data=anime_user_mat_sparse,
    fav_anime=my_favorite,
    mapper=anime_to_idx,
    n_recommendations=10)
```

✓ 0.4s

```
You have input anime: Cowboy Bebop
Found possible matches in our database: ['Cowboy Bebop']

Recommendation system start to make inference
......

Recommendations for Cowboy Bebop:
1: FLCL, with distance of 0.31903903051406135
2: Akira, with distance of 0.31690885637703625
3: Monster, with distance of 0.3168795683168413
4: Samurai Champloo, with distance of 0.3042900227126648
5: Bleach, with distance of 0.30396440374767253
6: Cowboy Bebop: Tengoku no Tobira, with distance of 0.30382183444633903
7: Neon Genesis Evangelion: The End of Evangelion, with distance of 0.30313470443520196
8: Death Note, with distance of 0.28743321445921133
9: Fullmetal Alchemist: Brotherhood, with distance of 0.28349179908653377
10: Neon Genesis Evangelion, with distance of 0.28263286514808483
```

SVD based Recommender

```python
def cosine_similarity_sort(r_data, anime_id, top_n=10):
    ind = anime_id - 1
    anime_row = r_data[ind, :]
    magnitude = np.sqrt(np.einsum('ij, ij -> i', r_data, r_data))
    matrix_similarity = np.dot(anime_row, r_data.T) / (magnitude[ind] * magnitude)
    sorted_indices = np.argsort(-matrix_similarity)
    return sorted_indices[:top_n]
```

```python
def get_most_similar_anime(anime_df, anime_id, index_list):
    print('Best recommendations for {0}: \n'.format(
    anime_df[anime_df.anime_id == anime_id].title.values[0]))
    for id in index_list + 1:
        print(anime_df[anime_df.anime_id == id].title.values[0])
```

```python
k = 50      #k-principal components to represent anime, anime_id to find recommendat
anime_id = 2        # id for which we want recommendation
top_n = 10
rep_data = V.T[:, :k] # representative data
index_list = cosine_similarity_sort(rep_data, anime_id, top_n)

#Get the top N recommendations
get_most_similar_anime(anime_df, anime_id, index_list)
```

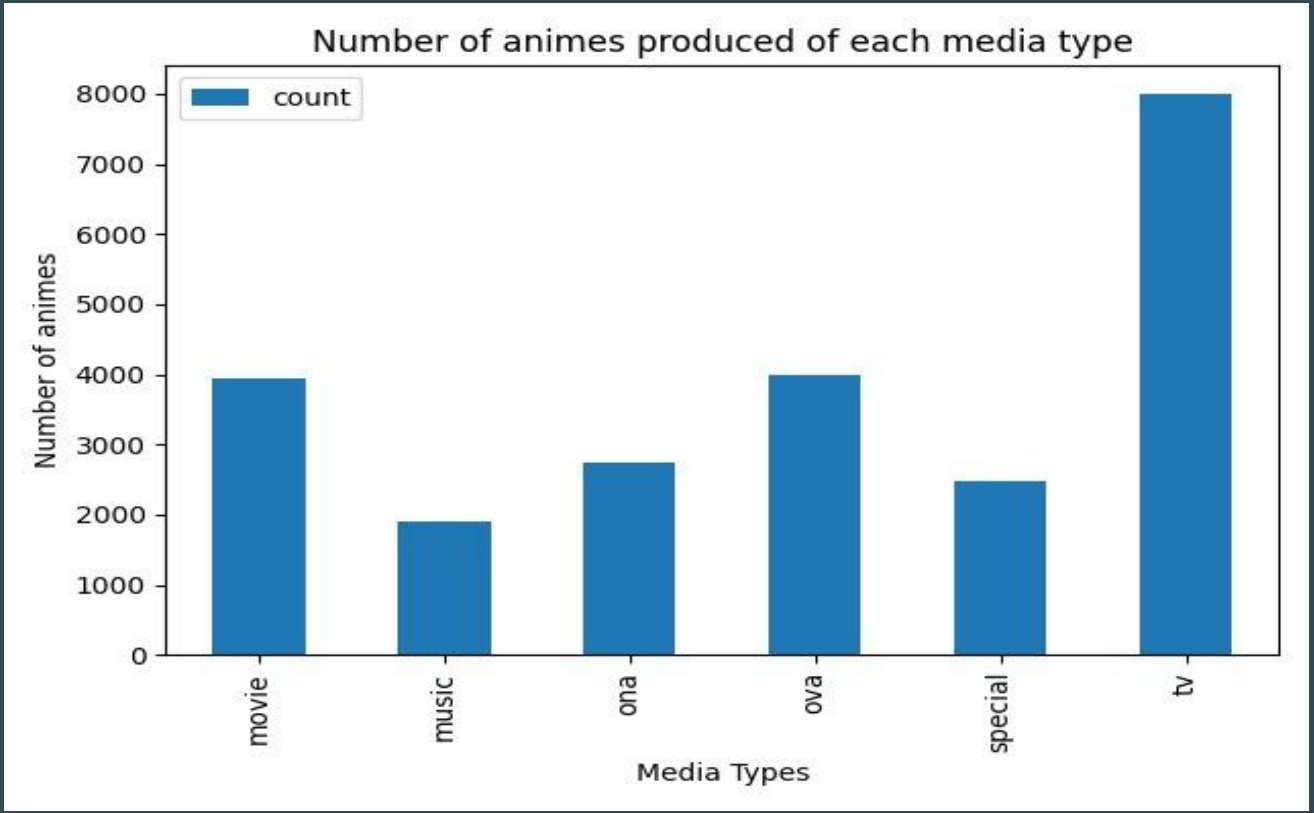
```
Best recommendations for Cowboy Bebop:

Initial D First Stage
Trigun
El Hazard: The Alternative World
Mobile Suit Gundam SEED
Beck
Saiyuuki Reload Gunlock
Yakitate!! Japan
Hunter x Hunter: Greed Island Final
Pita Ten
Green Green
```

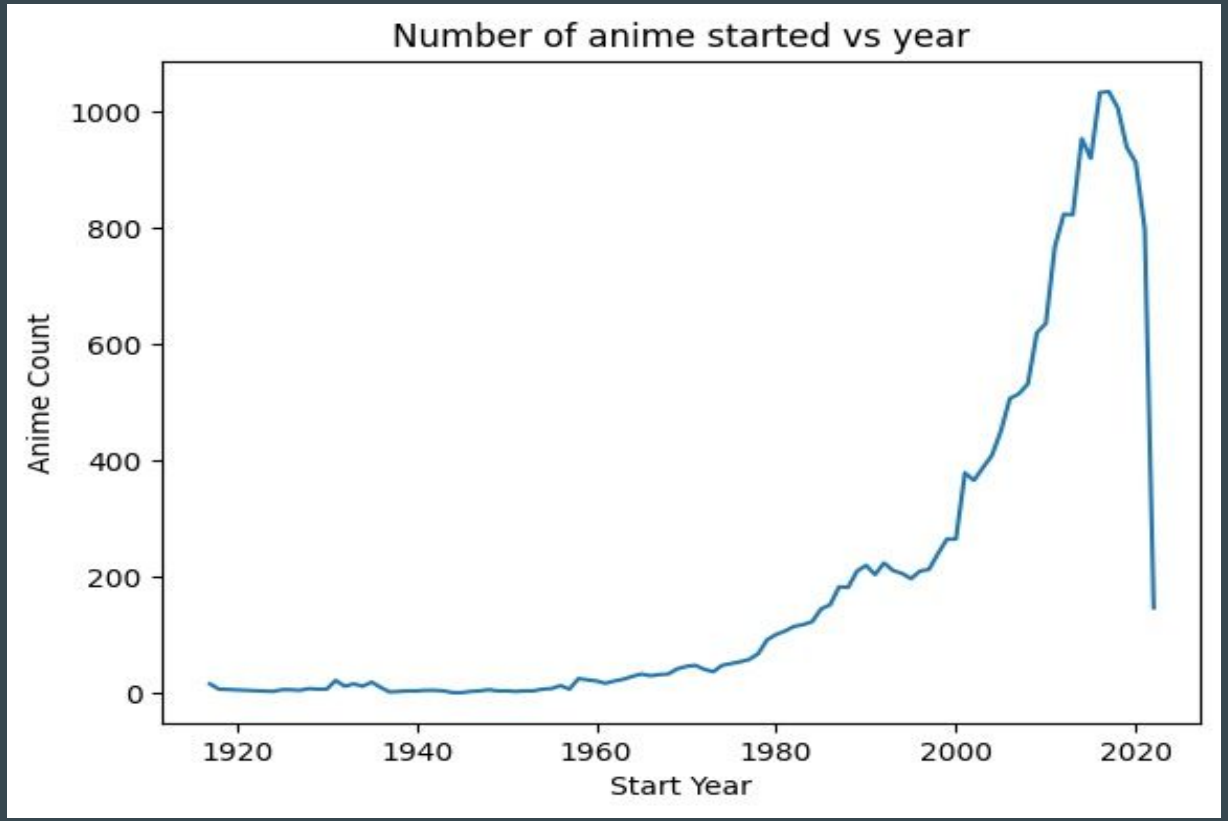
Insights and Analysis

...

Number of animes produced of each media type

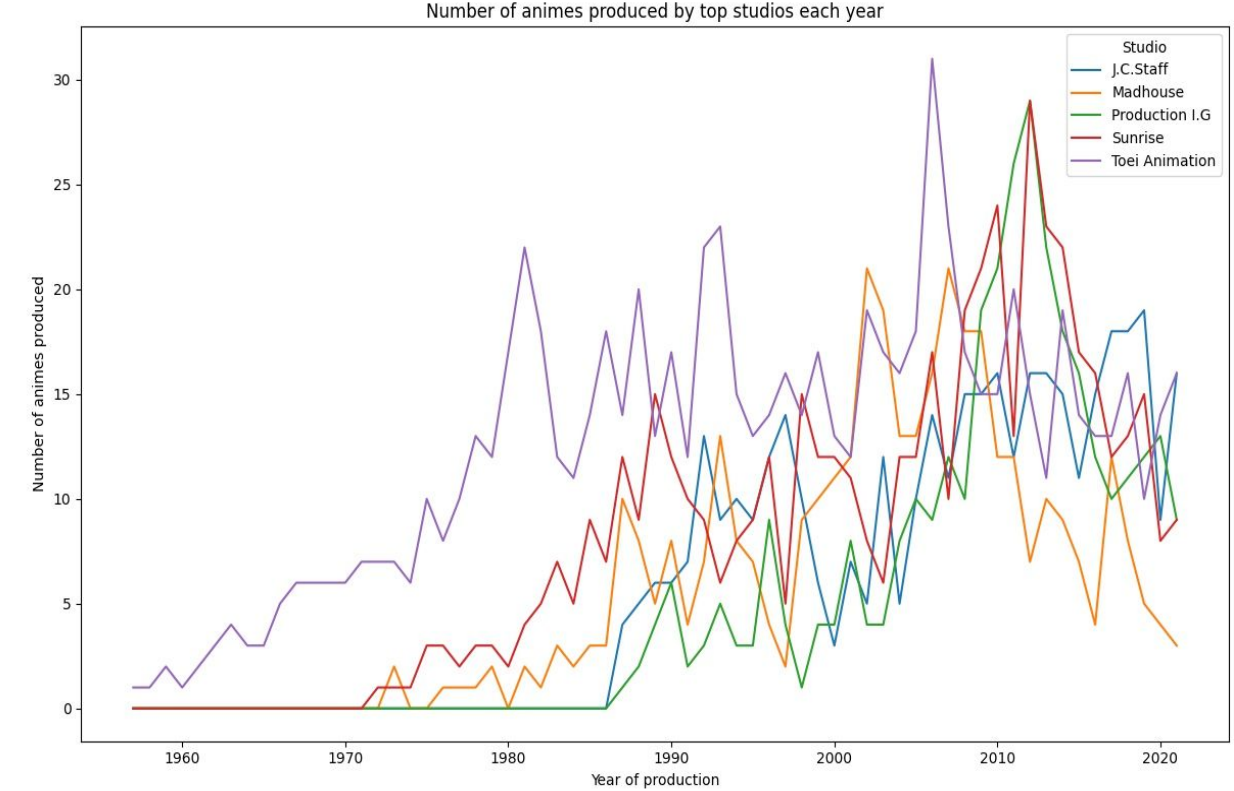


Number of animes aired with time



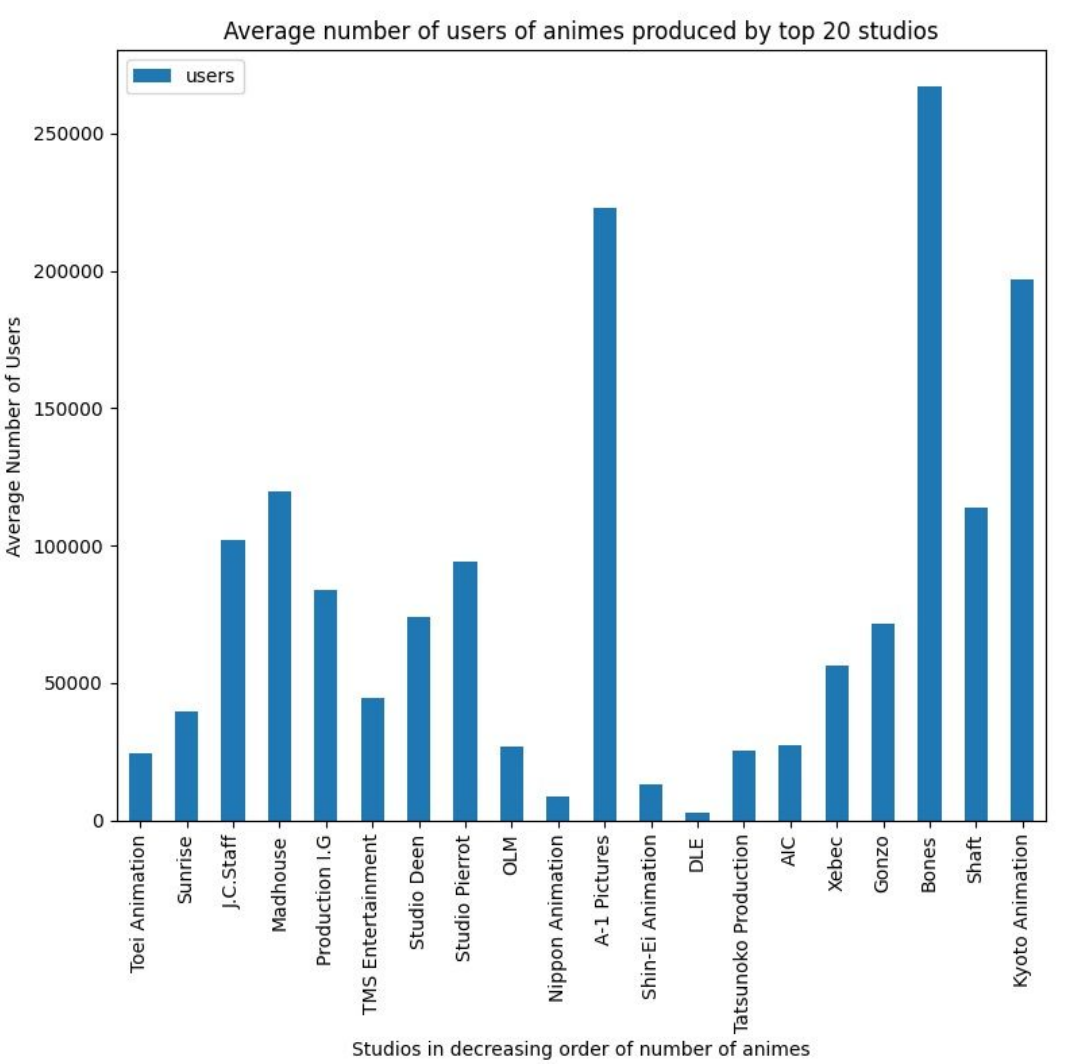
The dip in the number of animes started during 2020–21 is due to the COVID–19 pandemic.

Number of animes produced by top studios with time

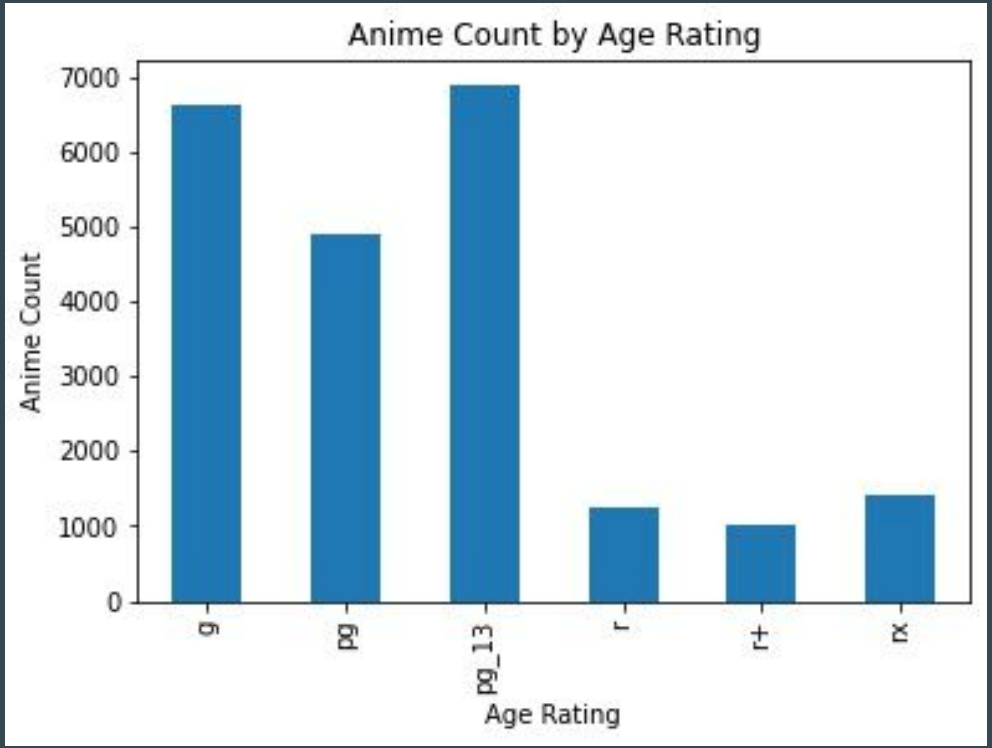


Viewership of Top Studios

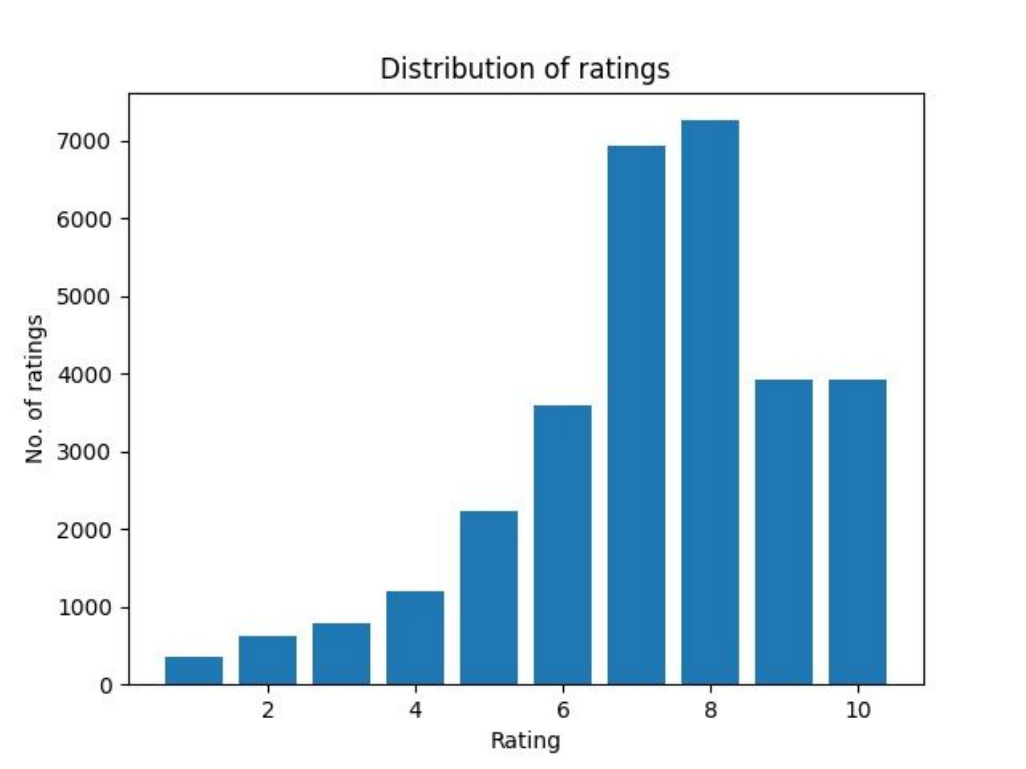
Bones has the largest viewership among top studios

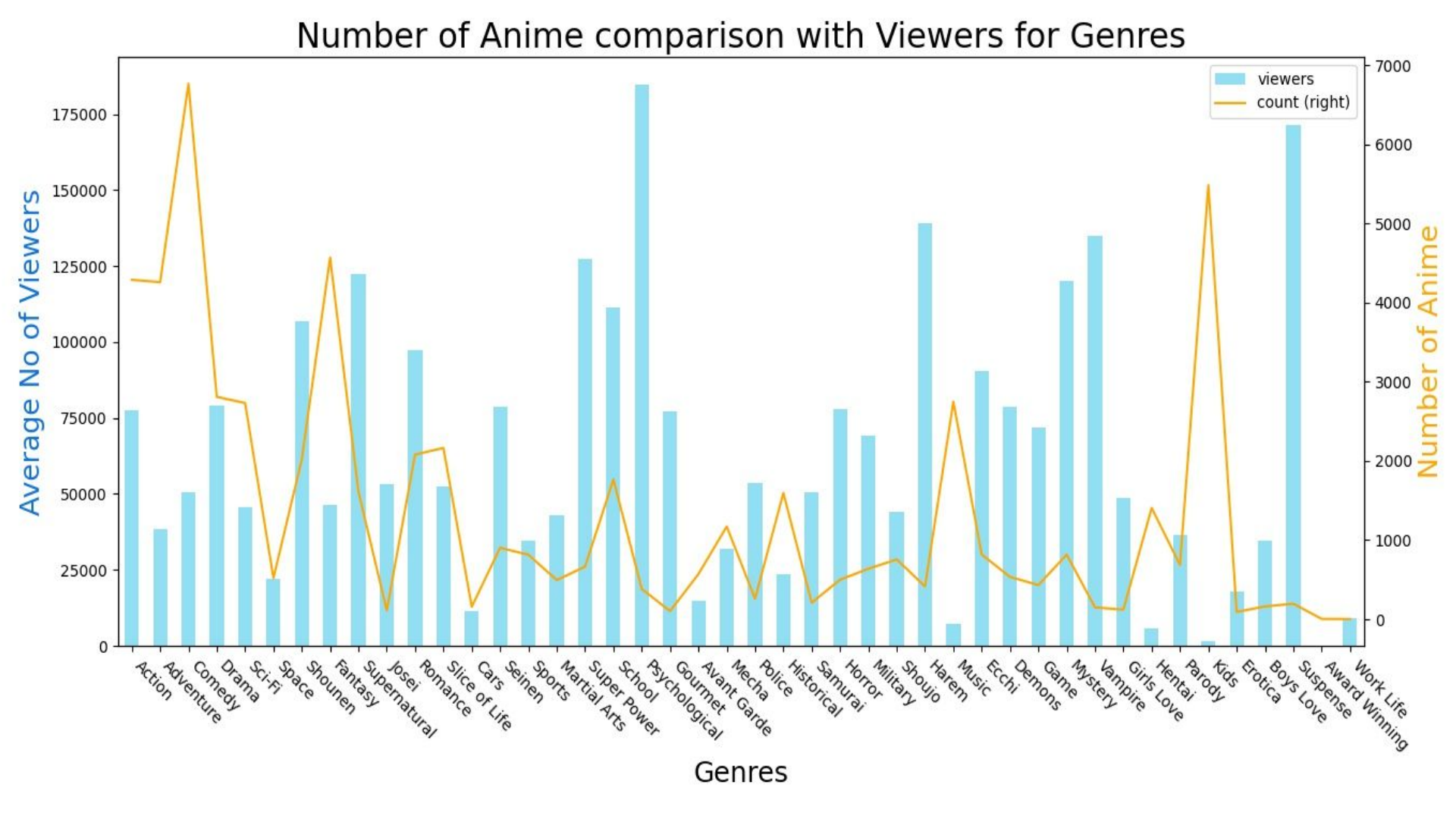


Distribution of animes based on age rating



Frequency of various ratings given by viewers




Number of Anime comparison with Viewers for Genres
viewers
count (right)
Average No of Viewers
Number of Anime
Genres
0
25000
50000
75000
100000
125000
150000
175000
0
1000
2000
3000
4000
5000
6000
7000
Action
Adventure
Comedy
Drama
Sci-Fi
Space
Shounen
Fantasy
Supernatural
Josei
Romance
Slice of Life
Cars
Seinen
Sports
Martial Arts
Super Power
School
Psychological
Gourmet
Avant Garde
Mecha
Police
Historical
Samurai
Horror
Military
Shoujo
Harem
Music
Ecchi
Demons
Game
Mystery
Vampire
Girls Love
Hentai
Parody
Kids
Erotica
Boys Love
Suspense
Award Winning
Work Life

Arigatō Gozaimasu!