

Analysis of Key Factors Influencing TV Shows' Success

Aakarsha LNU
alnu2@binghamton.edu
SUNY Binghamton
Binghamton, New York, USA

Riya Yeshwant Thakur
rthakur1@binghamton.edu
SUNY Binghamton
Binghamton, New York, USA

Aayushi Ahlawat
aahlawa1@binghamton.edu
SUNY Binghamton
Binghamton, New York, USA

Ishan Bagchi
ibagchi1@binghamton.edu
SUNY Binghamton
Binghamton, New York, USA

Tarun Tiwari
ttiware1@binghamton.edu
SUNY Binghamton
Binghamton, New York, USA

ABSTRACT

These days, television and movies are the two main forms of entertainment. Everybody enjoys watching and reading fascinating stories. When we watch television, we observe that many people discuss their experiences. On social media, anyone can publicly voice their opinions on any subject. First, big social networking sites like Twitter, Instagram, TMDB, and Reddit create a ton of information about user behaviour, news, rising trends, etc. The genre, cast, number of seasons, number of episodes, production companies, networks, episode run time all play a role in determining whether a television program is successful or not. Thus, by examining these elements for a range of programs, we aim to predict which TV show has a higher chance of becoming successful.

KEYWORDS

Tv shows, Genre, Cast, Budget, Reviews, Region, Twitter, Reddit, TMDB

1 DATASET

TMDB, Twitter, and Reddit are the datasets that have been considered for collecting, analysis and visualization of the data for gathering 4728 TV shows and over 1000 distinct TV series changes every day, including changes to their ratings, show information, number of episodes etc.

2 DATA COLLECTION

For this project, we have collected data from 3 data sources, namely, TMDB, Twitter, and Reddit which are used to find the essential factors which are contributing towards the success rate or the popularity of a TV Show.

3 DATA ANALYSIS

For this project, we have answered three research questions/areas based on our datasets and the type of data we have gathered. Our research objectives revolve around the factors influencing the popularity of TV shows based on data collected from TMDB, Twitter and Reddit.

4 METHODOLOGY

For creating live data dashboard, there are several tools, library packages, web applications and software's available.

From our data analysis project, we have answered and visualized

the results of all our research questions that were:

1. How are elements like the genre, cast, number of seasons, number of episodes, production firms, networks, and length of the episodes affecting the success of TV shows?

- From the TMDB data analysis it could be figured out that top 5 features that most affect the success of TV Shows are : returning series status, plot status, ended features status, canceled feature status and in production status with impact of greater than 2.0. And the least important feature is genre erotic with an impact of approximately 0.01.

2. To what extent are popular TV series on TMDB hyped on Twitter, and if this hype is positive or negative?

- On the basis of tweet frequency top 20 TV Shows were retrieved to calculate the polarity of each of the show. After calculating the polarity score hype was marked as positive, negative or neutral hype. On the basis of analysis made positive and negative hypes were in the range of 0-2000 with highest positive hype at 6000 and negative hype at 2000 respectively.

3. What phrases/terms from major TV series are becoming popular and most used or admired by the public on their respective subreddits?

- From the subreddit analysis 5 major phrases which are becoming popular are: BigBrother, thewalkingdead, House-OfCards, brooklynninenine, television

4.1 Dashboard

For this project, we will be creating an interactive dashboard that will be updated in real-time, showing results for the above-mentioned research questions.

- The dashboard will be implemented using the Flask framework, and all the plots and graphs will be built using various Python libraries.
- We will add the functionality to filter the results, based on the query entered by the user in our dashboard.

5 CONCLUSION

After examining all of these data sets for a range of programs, we were able to predict the factors influencing the popularity of TV shows. And for this project we will be adding some interactivity to basic data characterization by creating a live dashboard.