

Opinion Mining + Sentiment Classification

For the Top 10 Indian Web Series

(namely in 6 Genres - Action, Thriller, Drama, Comedy, Romance, Horror)

using Python



**BY
HARSH KUMAR**

(B.Tech Undergrad at Delhi Technological University, DTU

(formerly Delhi College of Engineering, DCE))

INTERN UNDER THE GUIDANCE OF

Prof. Sasadhar Bera, Ph.D

(Indian Institute of Management , Ranchi)

CERTIFICATE/LOR

ACKNOWLEDGEMENT

I am greatly indebted to Professor Sasadhar Bera, (Indian Institute of Management, Ranchi) who had accepted me for this Internship and had provided me with an excellent opportunity to work on this scalable project work on the topic of "*Opinion Mining + Sentiment Classification for the Top 10 Indian Web Series (namely in 6 Genres - Action, Thriller, Drama, Comedy, Romance, Horror) using Python*", which led me into doing a lot of research and learning new technical skills which diversified my knowledge to a huge extent.

I am grateful to him for having guided me throughout the research span of time and for providing his constructive criticism which made me bring out the best in me. Thus, I strongly record my deep sense of gratitude to my project mentor as he channelized me towards my first step into the field of research.

I express my sincere gratitude to my batch mates and my seniors, my family and friends for their wholehearted support.

I would like to convey my thanks to everyone who had been influential and supportive in this research cum intern work.

-HARSH KUMAR

ABSTRACT

Sentiment Analysis also known as Opinion Mining, refers to the use of natural language processing, text analysis to systematically identify, extract, quantify, and study affective states, subjective information and classify them into different classes called sentiments. Sentiment analysis is widely applied to user reviews and survey responses, online and social media, and healthcare materials for applications that range from marketing to customer service to clinical medicine.

In the past few years, new forms of entertainment, such as OTT (over-the-top) have emerged and become ubiquitous. OTT (over-the-top) is a means of providing television and film content over the internet at the request and to suit the requirements of the individual consumer. The term itself stands for “over-the-top”, which implies that a content provider is going over the top of existing internet services.

While there is no limit to the range of information conveyed by user reviews and texts, often these messages are used to share opinions and sentiments that people have. In our case, about the Web Series they'd watched, which we used to draw out insightful inferences for each genre.

In this project, we aim to perform Sentiment Analysis of user reviews of Top 10 Indian Web Series in different Genres (namely in 6 Genres - Action, Thriller, Drama, Comedy, Romance, Horror) shown on various OTT (over-the-top) platforms or media services (Amazon Prime Video, Netflix, ZEE5, Disney + Hotstar, Voot, ALTBalaji, SonyLiv etc).

Data used in this project(roughly 50k+) was collected by performing Web Scraping on review platforms/websites like IMDb, Amazon, Google Reviews etc.

We had worked on the following task :

OTT (over-the-top) platforms or media services (Amazon Prime Video, Netflix, ZEE5, Disney + Hotstar, Voot, ALTBalaji, SonyLiv) are showing different genres (categories) like Comedy, Drama, Action, Thriller, Horror, Devotional, Adult, etc. All the different categories contain different audience groups. The different audience has posted/posting their own views or reviews (supposing honest) on the particular category. Now the question is that , based-on those reviews, can we find a pattern of the success of any OTT content based on the audience likability by using text mining. What the audience wants to see (like emotion, real-life story, violence, etc) in different genres?

Given a review, we had classified whether the review was of positive, negative, or neutral sentiment. For reviews conveying both a positive and negative sentiment, whichever had the stronger sentiment was chosen.

KEYWORDS :

Opinion Mining, Sentiment Analysis, Machine Learning(ML), Natural Language Processing(NLP), Python, Text Mining, Data Preprocessing, Exploratory Data Analysis(EDA), Data Visualization, Web Scraping.

(To read the codes, refer the mentioned Github profile -
<https://github.com/HarshKumar1004/Opinion-Mining-Sentiment-Classification-InternWork-IIM-Ranchi>)

Table of Contents

Certificate.....	2
Acknowledgement.....	3
Abstract.....	4
Table of Contents.....	6
1. Introduction.....	7
1.1 Introduction to Sentiment Analysis.....	7
1.2 Introduction to Python.....	8
1.3 Introduction to TextBlob.....	9
1.4 Introduction to Naïve-Bayes (NB) Classifier.....	11
1.5 Need of Sentimental Analysis.....	12
2. Workflow.....	14
3. Implementation.....	15
3.1 Data Collection.....	16
3.2 Data Storage.....	23
3.3 Data Pre-Processing / Data Cleaning.....	24
3.4 Exploratory Data Analysis (EDA).....	27
3.5 Sentiment Analysis.....	28
3.6 Data Visualizations.....	30
4. Results and Analysis.....	35
4.1 Reviews Collected.....	35
4.2 Data Preparation.....	39
4.3 Sentiment Analysis.....	48
4.4 Data Visualizations.....	51
5. Conclusion and Future Scope.....	62
6. Bibliography.....	66

1. INTRODUCTION

In the past few years, there has been a huge growth in the use of OTT (over-the-top) platforms or media services such as Amazon Prime Video, Netflix, ZEE5, Disney + Hotstar, Voot, ALTBalaji, SonyLiv. Spurred by that growth, companies and media organizations are increasingly seeking ways to mine reviews and opinions for information about what people think and feel about their products and services.

Here, we performed a fair amount of research on how sentiments are expressed in genres such as in online reviews and critics' articles, how sentiments are expressed given the informal language and message-length constraints of user reviewing and then had drawn out useful inferences from them by identifying positive, negative, and neutral reviews based on their polarity and subjectivity. Additionally, created some data visualizations (bag of words, word clouds, plots etc) using the high frequency found sentiment keywords.

1.1 Introduction to Sentiment Analysis

Sentiment Analysis is a process of collecting and analyzing data based upon the person's feelings, reviews and thoughts. Sentimental analysis often called as opinion mining as it mines the important feature from people's opinion. From a user's perspective, people are able to post their own content through various media platforms, such as forums, micro-blogs, or online social networking sites. From a researcher's perspective, many social

media sites release their application programming interfaces (APIs), prompting data collection and analysis by researchers and developers.

Sentimental Analysis is done by using various machine learning techniques, statistical models and Natural Language Processing (NLP) for the extraction of feature from a large data.

Sentiment Analysis can be done at a document, phrase and sentence level. In document level, summary of the entire document is taken first and then it is analyzed whether the sentiment is positive, negative or neutral. In phrase level, analysis of phrases in a sentence is taken in account to check the polarity. In Sentence level, each sentence is classified in a particular class to provide the sentiment.

Sentimental Analysis has various applications. It is used to generate opinions for people of social media by analyzing their feelings or thoughts which they provide in form of text. Sentimental Analysis is used in many real life scenarios, to get reviews about any product or movies, to get the financial report of any company, for predictions or marketing.

1.2 Introduction to Python

Python is a high level, dynamic programming language which is used for this project report. Python3.9 version was used as it is a mature, versatile and robust programming language. It is an interpreted language which makes the testing and debugging extremely quickly as there is no compilation step. There are extensive open source libraries available for this version of python and a large community of users.

Python is simple yet powerful, interpreted and dynamic programming language, which is well known for its functionality of processing natural language data, i.e. spoken English. Other high level programming languages such as ‘R’ and ‘Matlab’ were considered because they have many benefits such as ease of use but they do not offer the same flexibility and freedom that Python can deliver.

1.3 Introduction to TextBlob

TextBlob is a Python (2 and 3) library for processing textual data. It provides a simple API for diving into common natural language processing (NLP) tasks such as part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, translation, and more.

TextBlob actively used Natural Language Tool Kit (NLTK) to achieve its tasks. NLTK is a library which gives an easy access to a lot of lexical resources and allows users to work with categorization, classification and many other tasks. TextBlob is a simple library which supports complex analysis and operations on textual data. NLTK is a perfect library for education and research, however it becomes very heavy and tedious for completing even the simple tasks. Additionally, TextBlob stands on the giant shoulders of NLTK and pattern, and plays nicely with both,hence had been used in this project.

For **lexicon-based approaches**, a sentiment is defined by its semantic orientation and the intensity of each word in the sentence. This requires a pre-defined dictionary classifying negative and positive words. Generally, a text message will be represented by bag of words. After assigning individual scores to all the words,

final sentiment is calculated by some pooling operation like taking an average of all the sentiments.

TextBlob returns **polarity** and **subjectivity** of a sentence. Polarity lies between [-1,1], -1 defines a negative sentiment and 1 defines a positive sentiment. Negation words reverse the polarity. TextBlob has semantic labels that help with fine-grained analysis. For example — emoticons, exclamation mark, emojis, etc. Subjectivity lies between [0,1]. **Subjectivity quantifies the amount of personal opinion and factual information contained in the text. The higher subjectivity means that the text contains personal opinion rather than factual information.** TextBlob has one more parameter — intensity. TextBlob calculates subjectivity by looking at the '**intensity**'. Intensity determines if a word *modifies* the next word. For English, adverbs are used as modifiers ('very good').

For example: We calculated polarity and subjectivity for "I do not like this example at all, it is too boring". For this particular example, polarity = -1 and subjectivity is 1, which is fair.

However, for the sentence "This was a helpful example but I would prefer another one". It returns 0.0 for both subjectivity and polarity which is not the finest answer we'd expect.

It is expected that if the library returns exactly 0.0 either if your sentence didn't contain any words that had a polarity in the NLTK training set or because TextBlob uses a weighted average sentiment score over all the words in each sample. This easily diffuses out the effect of sentences with widely varying polarities between words in our case : 'helpful' and 'but'.

So, TextBlob basically goes along finding words and phrases it can assign polarity and subjectivity to, and it averages them all together for longer text. More often than not, for simple cases, TextBlob works just fine and even here in our project it gave us satisfactory, reliable, confirming and corroborating results.

1.4 Introduction to Naive-Bayes (NB) Classifier

Naïve-Bayes classifiers are probabilistic classifiers which come under machine learning techniques. These classifiers are based on applying Bayes' theorem with strong (naïve) assumption of independence between each pair of features. Supervised machine learning is a technique whose task is to deduce a function from tagged training samples. The training samples for supervised learning consist of large set of examples for a particular topic. In supervised learning, every example training data comes in a pair of input (vector quantity) and output value (desired result). These algorithms analyze data and generate an output function, which is used to mapped new data sets to respective classes.

Let us assume, there is a dependent vector from x_1 to x_n , and a class variable 'y'. Therefore, according to Bayes' :

$$P(y | x_1, \dots, x_n) = \frac{P(y)P(x_1, \dots, x_n | y)}{P(x_1, \dots, x_n)}$$

Now according to assumption of independence

$$P(x_i | y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) = P(x_i | y),$$

For every 'i', this function becomes

$$P(y | x_1, \dots, x_n) = \frac{P(y) \prod_{i=1}^n P(x_i | y)}{P(x_1, \dots, x_n)}$$

In this $P(x_1, \dots, x_n)$ on given input is constant, hence we can apply classification rule as:

$$\begin{aligned} P(y | x_1, \dots, x_n) &\propto P(y) \prod_{i=1}^n P(x_i | y) \\ &\Downarrow \\ \hat{y} &= \arg \max_y P(y) \prod_{i=1}^n P(x_i | y), \end{aligned}$$

And for estimating we can use MAP (Maximum A Posterior) estimation $P(y)$ and $P(x_i | y)$; the $P(y)$ of class 'y' in training sample is relative frequency.

1.5 Need of Sentimental Analysis

1.5.1 Industry Evolution

Only the useful amount of data is required in the industry as compared to the set of complete unstructured form of the data. However the sentiment analysis done is useful for extracting the important feature from the data that will be needed solely for the purpose of industry. Sentimental Analysis will provide a great opportunity to the industries for providing value to their gain value and audience for themselves. Any of the industries with the business to consumer will get

benefit from this whether it is restaurants, entertainment, hospitality, mobile customer, retail or being travel.

1.5.2 Research Demand

Another important reason that stands behind the growth of SA deals with the demand of research in evaluation, appraisals, opinion and their classification. Present solutions for the purpose of sentiment analysis and opinion mining are rapidly evolving, specifically by decreasing the amount of human effort that will be required to classify the comments. Also the research theme that will be based in the long established disciplines of computer science like as text mining, machine learning, natural language processing and artificial intelligence, voting advise applications, automated content analysis, etc.

1.5.3 Decision Making

Every person who stores information on the blogs, various web applications and the web social media, social websites for getting the relevant information you need a particular method that can be used to analyze data and consequently return some of the useful results. It is going to be very difficult for company to conduct the survey that will be on the regular basis so that there comes the need to analyze the data and locate the best of the products that will be based on user's opinions, reviews and advices. The reviews and the opinions also help the people to take important decisions helping them in research and business areas.

1.5.4 Understanding Contextual

As human language is getting very complex day by day so it has become difficult for the machine to be able to understand human language that can be expressed in the slangs, misspelling, nuances, and the cultural variation. Thus, there will be a need of system that will make better understanding between the human and the machine language.

1.5.5 Internet Marketing

Another important reason behind the increase in the demand of sentimental analysis is the marketing done via internet by the business and companies organization. Now they regularly monitor the opinion of the user about their brand, product, or event on blog or the social post. Thus, we see that the Sentimental Analysis could also work as a tool for marketing too.

2. WORKFLOW

To achieve the objective discussed above in the Abstract section , the following Work Flow / Methodology was used :

- A thorough study of existing approaches and techniques in field of sentiment analysis.
- Collection / Extraction of relevant data from review platforms/websites like IMDb , Amazon , Google Reviews etc. with the help of Web Scraping techniques using Python or Web Scraping Google Extensions (like Instant Data Scraper).

- Pre-processing of data collected so that it can be fit for mining. Basically, the unstructured raw data undergoes various steps of preprocessing and cleaning which makes it more machine sensible than its previous form.
- Finally, the reviews were analyzed, visualized and classified with respect to sentiment classes based on the TextBlob provided Polarity and Subjectivity , such as Positive, Negative and Neutral.

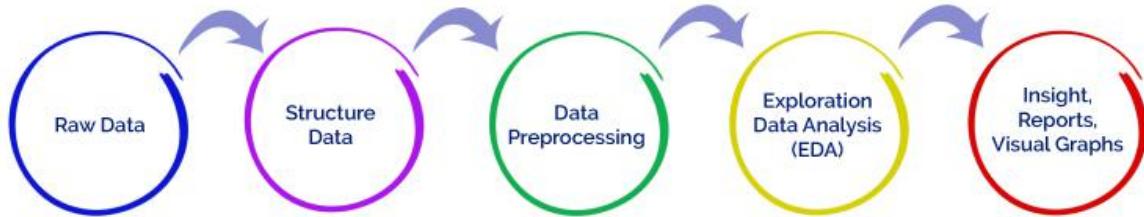


A typical sentiment analysis model.

3. Implementation

In this section we are going to study how a raw unstructured data was collected, stored, processed and then later on analysed as per our demands.

Data Preparation



3.1 Data Collection

Data collection is defined as the procedure of collecting, measuring and analyzing accurate insights for research using standard validated techniques. Data collection is not a simple task, as it may seem. Various decisions have to be made for collecting data. A researcher can evaluate their hypothesis on the basis of collected data. In most cases, data collection is the primary and most important step for research, irrespective of the field of research. The approach of data collection is different for different fields of study, depending on the required information.

For our project we had extracted the required datasets from review platforms/websites like IMDb , Amazon , Google Reviews etc. with the help of Web Scraping techniques by using Python and Web Scraping Google Extensions (like Instant Data Scraper).



3.1.1 Web Scraping by using Python

Web Scraping is an automated method used to extract large amounts of data from websites. The data on the websites are unstructured. Web scraping helps collect these unstructured data and store it in a structured form.

When you run the code for web scraping, a request is sent to the URL that you have mentioned. As a response to the request, the server sends the data and allows you to read the HTML or XML page.

The code then, parses the HTML or XML page, finds the data and extracts it.

To extract data using web scraping with python, you need to follow these basic steps:

- 1.Find the URL that you want to scrape
- 2.Inspecting the Page
- 3.Find the data you want to extract
- 4.Write the code
- 5.Run the code and extract the data
- 6.Store the data in the required format

As we know, Python has various applications and there are different libraries for different purposes.

In our written code, we used the following libraries:

1.Requests: Requests allows you to send HTTP, requests extremely easily. There's no need to manually add query strings to your URLs, or to form-encode your POST data.

2.Beautiful Soup: Beautiful Soup is a Python package for parsing HTML and XML documents. It creates parse trees that is helpful to extract the data easily.

3.Pandas: Pandas is a library used for data manipulation and analysis. It is used to extract the data and store it in the desired format.

Step 1: Find the URL that you want to scrape

For this example, we are going scrape Amazon website to extract reviews and other columns of an Indian web series called- The Family Man.

The URL for this page is

https://www.amazon.com/The-Family-Man-Season-1/product-reviews/B07XGLTVY7/ref=cm_cr_arp_d_paging_btm_next_2?pageNumber=2
(used this link as an example, can be changed accordingly.)

Step 2: Inspecting the Page

The data is usually nested in tags. So, we inspect the page to see, under which tag the data we want to scrape is nested.

To inspect the page, just right click on the element and click on “Inspect” or click Ctrl+Shift+I.

When you click on the “Inspect” tab, you will see a “Browser Inspector Box” open.

Step 3: Find the data you want to extract

Let's extract the Product,Title,Rating,Body/Reviews,which is in the "title", "a","i" and "span" tag respectively.

Step 4: Write the code

Step 5: Run the code and extract the data

Step 6: Store the data in a required format

After extracting the data, you might want to store it in a format.

This format varies depending on your requirement. For this example, we stored the extracted data in an Excel file.

(To read the codes, refer the mentioned Github profile -
<https://github.com/HarshKumar1004/Opinion-Mining-Sentiment-Classification-InternWork-IIM-Ranchi>)

The following was the code that I wrote for Web Scraping using Python :

```

Web Scraping with Python.py - C:\Users\Asus\AppData\Local\Programs\Python\Python39\Web Scraping with Python.py (3.9.6)
File Edit Format Run Options Window Help

...
Import the following libraries, make sure you have them installed in your system prior to importing them.
To install , press Win+R ,then write cmd to launch command line, and then write pip install libraryname
...
import requests
from bs4 import BeautifulSoup
import pandas as pd

''' Creating an empty list , where we will add/append the required data and then finally print it or export its data to a desired file format.'''
reviewlist = []

''' Defining a function to parse data '''
def get_soup(url):
    ''' Using requests library to get data from the url'''
    """ NOTE: We could've simply written r=requests.get(url),but this would show the following error:
        To discuss automated access to Amazon data please contact api-services-support@amazon.com.
        For information about migrating to our APIs refer to our Marketplace APIs at https://developer.amazonservices.com/ref=rm_5_sv,
        or our Product Advertising API at https://affiliate-program.amazon.com/gp/advertising/api/detail/main.html/ref=rm_5_ac for advertising use cases.

        The reason behind this error is that Amazon is very anti-scraping.
        There's an entire industry built around scraping data from Amazon, and Amazon has its own API access to sell,
        so it's in their best interest to stop people from freely grabbing data from their pages.
        Simply put,data is not to be sold for free.

        Based on that code, we make too many requests too quickly and thus our IP may get banned.
        When scraping sites, it's usually best to scrape responsibly by not going too fast, rotating user agents, and rotating IPs through a proxy service.

        To seem less programmatic, you can also try randomizing request timing to seem more human.
        Even with all of that, you'll still likely hit issues with this. Amazon is not an easy site to reliably scrape.

        So in order to counter this problem,the simplest and easiest way to manage it on windows is by using Splash and Docker Desktop.

        Splash is a lightweight browser with an API designed specifically for web scraping and rendering javascript and dynamic websites.
        We can quickly and easily send requests to Splash instance and have it render the JS for us, and return the HTML to parse.

        To download it -
        Docker Desktop -https://www.docker.com/products/docker-desktop
        Splash -https://splash.readthedocs.io/en/stable/install.html#linux-dock """
    #Using Splash from Docker Desktop. Creating a localhost at 8050 (default) and then using it to render the javascripts.
    r = requests.get('http://localhost:8050/render.html', params={'url': url, 'wait': 2})
    soup = BeautifulSoup(r.text, 'html.parser')
    return soup

#Extracting the required data from the HTML parsed data, by selecting the correct class names of selected tags.
def get_reviews(soup):
    ...

#Extracting the required data from the HTML parsed data, by selecting the correct class names of selected tags.
def get_reviews(soup):
    reviews = soup.find_all('div', {'data-hook': 'review'})
    try:
        for item in reviews:
            review = {
                'product': soup.title.text.strip(),
                'title': item.find('a', {'data-hook': 'review-title'}).text.strip(),
                'rating': item.find('i', {'data-hook': 'review-star-rating'}).text.strip(),
                'body': item.find('span', {'data-hook': 'review-body'}).text.strip(),
            }
            reviewlist.append(review)
    except:
        pass

#You can vary the range based on the no. of loops you want python to run(range(a,b)) , depending on the amount of data to be extracted/scraped.
for x in range(1,3): #Passing the Url of the amazon website of The Family Man reviews

    soup = get_soup(f'https://www.amazon.com/The-Family-Man-Season-1/product-reviews/B07XGLTVY7/ref=cm_cr_arp_d_paging_btm_next_2?pageNumber={x}')
    print(f'Getting page: {x}')
    get_reviews(soup)
    print(len(reviewlist))
    if not soup.find('li', {'class': 'a-disabled a-last'}):
        pass
    else:
        break

#Creating a dataframe from the scraped off data.
df = pd.DataFrame(reviewlist)

#Saving the data to an excel file, write the location where you want it to be saved, here the file will be stored on the Desktop
df.to_excel(r'C:\Users\Asus\Desktop\AMAZON REVIEWS OF- THE FAMILY MAN.xlsx', index=False)
print('DataFrame is written successfully to Excel Sheet.')

```

Web Scraping Code using Python

3.1.2 Web Scraping by using Google Extensions

A browser extension is a small software module for customizing a web browser. Browsers typically allow a variety of extensions, including user interface modifications, ad blocking, and cookie management.

Browser plug-ins are a separate type of module. The main difference is that extensions are usually just source code, but plug-ins are always executables (i.e. object code). As of 2021, plug-ins have been deprecated by most browsers, while extensions are widely used. The most popular browser, Google Chrome, has over 100,000 extensions available but no longer supports plug-ins.

Google Chrome extensions are programs that can be installed into Chrome in order to change the browser's functionality. This includes adding new features to Chrome or modifying the existing behavior of the program itself to make it more convenient for the user.

Google provides a lot of automated web scraping tools and extensions , but we have used Instant Data Scraper as it is relatively easier and beginner friendly to use. (requires no prior technical knowledge like sitemaps , HTML tags , working of web frameworks like JavaScript etc.)

[Home](#) > [Extensions](#) > Instant Data Scraper



Instant Data Scraper extracts data from web pages and exports it as Excel or CSV files.

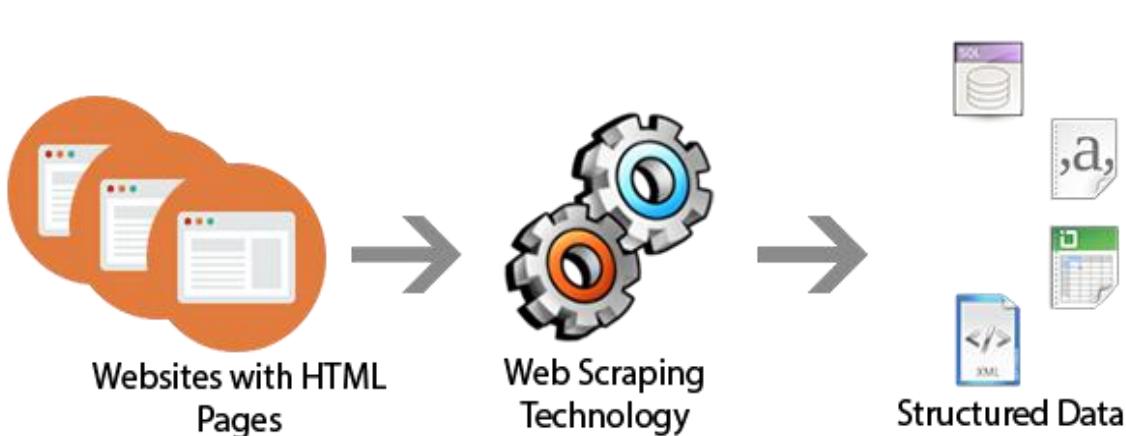
Instant Data Scraper is an automated data extraction tool for any website. It uses AI to predict which data is most relevant on a HTML page and allows saving it to Excel or CSV file (XLS, XLSX, CSV).

This tool does not require website specific scripts, instead it uses heuristic AI analysis of HTML structure to detect data for extraction. If the prediction is not satisfactory, it lets the user customize the selections for greater accuracy. This type of scraping technology is much more convenient, because it does not require large user created libraries of scraping scripts, which often become filled with outdated and redundant versions. This means that our scraping method works just as well with small and lesser known websites, as it does with global giants like Amazon. Also, our users do not need to have any coding, json or XML skills!

Installing the extension

Important: You can't add extensions when you browse in Incognito mode or as a guest.

1. Open the Chrome Web Store.
2. Find and select the extension you want. (Instant Data Scraper)
3. Click **Add to Chrome**.
4. Some extensions will let you know if they need certain permissions or data.
To approve, click **Add extension**.
5. To use the extension, click the icon to the right of the address bar to perform web scraping on the website/webpage.



3.2 Data Storage

Once, we started getting our data using web scraping techniques our next step was to store that data so that we could use it for sentiment analysis and it's prior needed steps. We ran our code for a period of time and collected the user reviews(roughly 50k+ data set) of Top 10 Indian Web Series in different Genres (namely in 6 Genres - Action, Thriller, Drama, Comedy, Romance, Horror) shown on various OTT (over-the-top) platforms or media services (like Amazon Prime Video, Netflix, ZEE5, Disney + Hotstar, Voot, ALTBalaji, SonyLiv etc).

Every time we ran the code shown in figure, an excel file was generated which consisted of user opinion and reviews about the series that were extracted. We used .xlsx format for our collected data files because they make it accessible and comprehend the particular field which consists of text.

We made 6 separate excel files to store user reviews of each genres. Each excel file has 10 sheets, for each web series of that genre. We stored them in our hard drive from where these could be easily imported to our Jupyter Notebook, used as an IDE (or any other platform) and further proceeded for analysis. Once we stored our reviews we had to pre-process the data stored before applying it to any classifier, if needed(we didn't, as we already got satisfactory results) because the data we collected was not fit for mining yet. Therefore, pre-processing / cleaning the data was our next step. Following figure shows glimpse of different files stored in hard drive.

The screenshot shows a Windows File Explorer window with the following details:

- File Explorer Title Bar:** ALL REVIEWS
- Toolbar:** File, Home, Share, View
- Breadcrumb Navigation:** This PC > Desktop > Intern Work-IIM Ranchi > ALL REVIEWS
- Left Sidebar (Quick Access):**
 - Desktop (selected)
 - Downloads
 - Documents
 - Pictures
 - ALL REVIEWS
 - Intern Work-IIM Rar
- Table View:** A list of files in the 'ALL REVIEWS' folder.

Name	Date modified	Type	Size
ACTION REVIEWS	14-07-2021 02:14 PM	XLSX Worksheet	760 KB
COMEDY REVIEWS	14-07-2021 10:26 AM	XLSX Worksheet	724 KB
DRAMA REVIEWS	14-07-2021 02:14 PM	XLSX Worksheet	539 KB
HORROR REVIEWS	14-07-2021 12:40 PM	XLSX Worksheet	371 KB
ROMANCE REVIEWS	14-07-2021 10:49 AM	XLSX Worksheet	508 KB
THRILLER REVIEWS	14-07-2021 02:12 PM	XLSX Worksheet	654 KB
TOP 10 INDIAN WEB SERIES IN VARIOUS ...	25-08-2021 06:06 PM	XLSX Worksheet	13 KB

Review files stored in Hard Drive

3.3 Data Pre-Processing / Data Cleaning

When dealing with numerical data, data cleaning often involves removing null values and duplicate data, dealing with outliers, etc. With text data, there are some common data cleaning techniques, which are also known as text pre-processing techniques.

Data preprocessing refers to manipulation or dropping of data before it is used in order to ensure or enhance performance, and is an important step in the data mining process. The phrase "garbage in, garbage out" is particularly applicable to data mining and machine learning projects.

Data-gathering methods are often loosely controlled, resulting in out-of-range values (e.g., Income: -100), impossible data combinations (e.g., Sex: Male, Pregnant: Yes), and missing values, etc. Analyzing data that has not been carefully screened for such problems can produce misleading results. Thus, the representation and quality of data is first and foremost before running any analysis. If there is much irrelevant and redundant information present or noisy and unreliable data, then knowledge discovery during the training phase is more

difficult. Data preparation and filtering steps can take considerable amount of processing time. Examples of data preprocessing include cleaning, instance selection, normalization, one hot encoding, transformation, feature extraction and selection, etc. The product of data preprocessing is the final training set.

With text data, this cleaning process can go on forever. There's always an exception to every cleaning step. So, we're going to follow the MVP (minimum viable product) approach - start simple and iterate.

Here are a bunch of things you can do to clean your data. We're going to execute just the common cleaning steps here and the rest can be done at a later point to improve our results.

Common data cleaning steps on all text:

- Make text all lower case
- Remove punctuation
- Remove numerical values
- Remove common non-sensical text (\n-new lines,\t-whitespaces etc)
- Tokenize text
- Remove stop words

More data cleaning steps after tokenization:

- Stemming / lemmatization
- Parts of speech tagging
- Create bi-grams or tri-grams
- Deal with typos
- And more...

```

# Applying a first round of text cleaning techniques
import re
import string

def clean_text_round1(text):
    '''Make text lowercase, remove text in square brackets, remove punctuation and remove words containing numbers.'''
    text = str(text)
    text = text.lower()

    text = re.sub('\[.*?\]', '', text)
    text = re.sub('[%s]' % re.escape(string.punctuation), '', text)
    text = re.sub('\w*\d\w*', '', text)
    return text

```

```

# Apply a second round of cleaning
def clean_text_round2(text):
    '''Get rid of some additional punctuation and non-sensical text that was missed the first time around.'''
    text = str(text)
    text = re.sub('["“”…]', '', text)
    text = re.sub('\n', '', text)
    return text

```

```

# Applying a third round of cleaning

import re
import string

text_translator = str.maketrans({ord(c): " " for c in string.punctuation})
def clean_text_round3(text, remove_punctuation_all=False):
    if not text:
        return ''
    try:
        text = text.replace(chr(160), " ")
        text = ''.join([i if ord(i) < 128 else ' ' for i in text])
    except Exception as e:
        try:
            text = text.encode('utf-8')
            text = text.decode('utf-8')
        except Exception as e:
            return ""
    try:
        text = text.encode('ascii', 'ignore').decode("utf-8")
        text = text.translate(text_translator)
    except Exception as e:
        return ""
    while ' ' in text:
        text = text.replace(' ', ' ')
    text = text.strip()
    return text

```

Python code for Data Cleaning

NOTE:

This data cleaning aka text pre-processing step could go on for a while, but we are going to stop for now. After going through some analysis techniques, if you see that the results don't make sense or could be improved, you can come back and make more edits such as:

- Mark 'cheering' and 'cheer' as the same word (stemming / lemmatization)
- Combine 'thank you' into one term (bi-grams)
- And a lot more...

3.4 Exploratory Data Analysis(EDA)

After the data cleaning step where we put our data into a few standard formats, the next step was to take a look at the data and see if what we're looking at made sense. Before applying any fancy algorithms, it's always important to explore the data first.

When working with numerical data, some of the exploratory data analysis (EDA) techniques we can use include finding the average of the data set, the distribution of the data, the most common values, etc. The idea is the same when working with text data. We are going to find some more obvious patterns with EDA before identifying the hidden patterns with machines learning (ML) techniques. Like :

- Most common words - find these and create word clouds

Organizing The Data

The output of this notebook was clean, organized data which can be done in two standard text formats:

1. Corpus - a collection of text
2. Document-Term Matrix - word counts in matrix format

3.5 Sentiment Analysis

So far, all of the analysis we've done had been pretty generic - looking at counts, creating wordcloud plots, etc. These techniques could be applied to numeric data as well.

When it comes to text data, there are a few popular techniques that we may go through, starting with sentiment analysis. A few key points to remember with sentiment analysis.

1. **TextBlob Module:** Linguistic researchers have labeled the sentiment of words based on their domain expertise. Sentiment of words can vary based on where it is in a sentence. The TextBlob module allows us to take advantage of these labels.
2. **Sentiment Labels:** Each word in a corpus is labeled in terms of polarity and subjectivity (there are more labels as well, but we're going to ignore them for now). A corpus' sentiment is the average of these.
 - **Polarity:** How positive or negative a word is. -1 is very negative. +1 is very positive.
 - **Subjectivity:** How subjective, or opinionated a word is. 0 is fact. +1 is very much an opinion.

For more info on how TextBlob coded up its sentiment function. (https://planspace.org/20150607-textblob_sentiment/)

Let's take a look at the sentiment of the various transcripts via our written Python code:

```
# Create quick lambda functions to find the polarity and subjectivity of each routine

from textblob import TextBlob

pol = lambda x: TextBlob(str(x)).sentiment.polarity
sub = lambda x: TextBlob(str(x)).sentiment.subjectivity

# Another way of writing the code , instead of using lambda parameter above.
...
def get_Subjectivity(text):
    return TextBlob(text).sentiment.subjectivity
def get_Polarity(text):
    return TextBlob(text).sentiment.polarity
...

data_clean_df['polarity'] = data_clean_df['transcript'].apply(pol)
data_clean_df['subjectivity'] = data_clean_df['transcript'].apply(sub)
data_clean_df
```

```
#classifying sentiments based on the reviews'score
def get_analysis(score):
    if score > 0:
        return "positive"
    elif score < 0:
        return "negative"
    else:
        return 'neutral'
data_clean_df["Analysis"] = data_clean_df.polarity.apply(get_analysis)
data_clean_df
```

```

j=0
k=0
for i in range(0,data_clean_df.shape[0]):
    if data_clean_df.Analysis[i]=='negative':
        j= j+1
    elif data_clean_df.Analysis[i]=='positive':
#The following code can be undocumented , if you're interested in reading that sentiments' reviews.
        #           print (k,data_clean_df.transcript[i])
        k+=1
neu= data_clean_df.shape[0]- (j+k)
print ('So,The following is our "Sentiment Analysis" for the Top 10 Indian Web Series(Action Genre) : ')
print ('\nNo of Negative Reviews from our Total DataSet(around 10k) ->',j)
print ('No of Positive Reviews from our Total DataSet(around 10k) ->',k)
print ('No of Neutral Reviews from our Total DataSet(around 10k) ->',neu)

neg_per=(j/data_clean_df.shape[0])*100
pos_per=(k/data_clean_df.shape[0])*100
neu_per=(neu/data_clean_df.shape[0])*100

print ('\nPercentage of Negative Reviews -> '+ str(neg_per) + " %")
print('Percentage of Positive Reviews -> '+ str(pos_per) + ' %')
print('Percentage of Neutral Reviews -> '+ str(neu_per) + " %")

```

3.6 Data Visualizations

Data Visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.

The advantages and benefits of good data visualization:

Our eyes are drawn to colors and patterns. We can quickly identify red from blue, square from circle. Our culture is visual, including everything from art and advertisements to TV and movies. Data visualization is another form of visual art that grabs our interest and keeps our eyes on the message. When we see a chart, we quickly see trends and outliers. If we can see something, we internalize it quickly. It's basically storytelling with a purpose.

Other benefits of data visualization include the following:

- Confirms our results derived from numeric data analysis.
- The ability to absorb information quickly, improve insights and make faster decisions;
- An increased understanding of the next steps that must be taken to improve the organization;
- An improved ability to maintain the audience's interest with information they can understand;
- An easy distribution of information that increases the opportunity to share insights with everyone involved;
- Eliminate the need for data scientists since data is more accessible and understandable; and
- An increased ability to act on findings quickly and, therefore, achieve success with greater speed and less mistakes.

Following are the Python Code for Data Visualizations :

```
# Let's plot the results
import matplotlib.pyplot as plt

plt.rcParams['figure.figsize'] = [16, 12]

plt.scatter(data_clean_df['polarity'],data_clean_df['subjectivity'])

plt.title('Sentiment Analysis (Scatter Plot)', fontsize=30)
plt.xlabel('<-- Negative ----- Positive -->', fontsize=15)
plt.ylabel('<-- Facts ----- Opinions -->', fontsize=15)

plt.show()
```

```
plt.fill(data_clean_df['polarity'])
plt.title('Polarity of each Dataset', fontsize=30)
plt.xlabel('DataSets ( around 10K)', fontsize=15)
plt.ylabel('<-- Negative ----- Positive -->', fontsize=15)

plt.show()
```

```
plt.fill(data_clean_df['subjectivity'])
plt.title('Subjectivity of each Dataset', fontsize=30)
plt.xlabel('DataSets ( around 10K)', fontsize=15)
plt.ylabel('<-- Facts ----- Opinions -->', fontsize=15)

plt.show()
```

```
plt.plot(data_clean_df['polarity'],data_clean_df['subjectivity'])
plt.rcParams['figure.figsize'] = [14, 10]
plt.title('Sentiment Analysis (Line Plot)', fontsize=30)
plt.xlabel('<-- Negative ----- Positive -->', fontsize=15)
plt.ylabel('<-- Facts ----- Opinions -->', fontsize=15)

plt.show()
```

```
plt.plot(data_clean_df['polarity'])
plt.title('Polarity of each Dataset (Line Plot)', fontsize=30)
plt.xlabel('DataSets ( around 10K)', fontsize=15)
plt.ylabel('<-- Negative ----- Positive -->', fontsize=15)

plt.show()
```

```
plt.plot(data_clean_df['subjectivity'])
plt.title('Subjectivity of each Dataset (Line Plot)', fontsize=30)
plt.xlabel('DataSets ( around 10K)', fontsize=15)
plt.ylabel('<-- Facts ----- Opinions -->', fontsize=15)

plt.show()
```

```

plt.hist(data_clean_df['polarity'], rwidth=.969)
plt.title('Polarity of each Dataset (Histogram)', fontsize=30)
plt.xlabel('<-- Negative ----- Positive -->', fontsize=15)
plt.ylabel('Frequency', fontsize=15)

plt.show()

```

```

plt.hist(data_clean_df['subjectivity'], rwidth=.969)
plt.title('Subjectivity of each Dataset (Histogram)', fontsize=30)
plt.xlabel('<-- Facts ----- Opinions -->', fontsize=15)
plt.ylabel('Frequency', fontsize=15)

plt.show()

```

```

# Python program to generate WordCloud for POSITIVE SENTIMENTS

# importing all necessary modules
from wordcloud import WordCloud, STOPWORDS
import matplotlib.pyplot as plt
import pandas as pd

comment_words_pos = ''

# Add new stop words
selected_stop_words=['show','season','one','season','watch','story','web','episodes','bajpayee','bajpai','manoj',
                     'episode','review','actor','actors']
stopwords = list(additional_stop_words) + selected_stop_words + list(STOPWORDS)

# iterate through the file
for val in positive_df.transcript:

    # typecaste each val to string
    val = str(val)

    # split the value
    tokens = val.split()
    comment_words_pos += " ".join(tokens)+" "

wordcloud = WordCloud(width = 800, height = 800,
                      background_color ='white',
                      stopwords = stopwords,
                      min_font_size = 10).generate(comment_words_pos)

# plot the WordCloud image
plt.figure(figsize = (10,10), facecolor = None)
plt.imshow(wordcloud)
plt.axis("off")
plt.tight_layout(pad = 0)
plt.title('WordCloud for POSITIVE SENTIMENTS\n',fontsize=30)

plt.show()

```

```

# Python program to generate WordCloud for NEGATIVE SENTIMENTS

# importing all necessary modules
from wordcloud import WordCloud, STOPWORDS
import matplotlib.pyplot as plt
import pandas as pd

comment_words_neg = ''

# Add new stop words

selected_stop_words=['show','season','one','good','season','watch','story','bajpayee',
                     'bajpai','manoj','episode','review','actor','actors']
stopwords = list(additional_stop_words) + selected_stop_words + add_stop_words + list(STOPWORDS)

# iterate through the file
for val in negative_df.transcript:

    # typecaste each val to string
    val = str(val)

    # split the value
    tokens = val.split()
    comment_words_neg += " ".join(tokens)+" "

wordcloud = WordCloud(width = 800, height = 800,
                      background_color ='white',
                      stopwords = stopwords,
                      min_font_size = 10).generate(comment_words_neg)

# plot the WordCloud image
plt.figure(figsize = (10,10), facecolor = None)
plt.imshow(wordcloud)
plt.axis("off")
plt.tight_layout(pad = 0)
plt.title('WordCloud for NEGATIVE SENTIMENTS\n',fontsize=30)

plt.show()

```

```

# Python program to generate WordCloud for NEUTRAL SENTIMENTS

# importing all necessary modules
from wordcloud import WordCloud, STOPWORDS
import matplotlib.pyplot as plt
import pandas as pd

comment_words_neu = ''

# Add new stop words

selected_stop_words=['show','season','one','good','thriller','season','shame','watch',
                     'story','masterpiece','bajpayee','bajpai','manoj','episode','review','actor','actors']
stopwords = list(additional_stop_words) + selected_stop_words + add_stop_words + list(STOPWORDS)

# iterate through the file
for val in neutral_df.transcript:

    # typecaste each val to string
    val = str(val)

    # split the value
    tokens = val.split()
    comment_words_neu += " ".join(tokens)+" "

wordcloud = WordCloud(width = 800, height = 800,
                      background_color ='white',
                      stopwords = stopwords,
                      min_font_size = 10).generate(comment_words_neu)

# plot the WordCloud image
plt.figure(figsize = (10,10), facecolor = None)
plt.imshow(wordcloud)
plt.axis("off")
plt.tight_layout(pad = 0)
plt.title('WordCloud for NEUTRAL SENTIMENTS\n',fontsize=30)

plt.show()

```

4. Results and Analysis

In this section we are going to show various results that we had achieved in our implementation.(Basically this is for genre selective output and the generalized code had been written in the implementation part.)

4.1 Reviews Collected

User reviews were collected with the help of Web Scraping techniques from different OTT platforms (Amazon Prime, Netflix, ALT Balaji, ZEE5, Disney+Hotstar) for the top 10 Indian Web Series in 6 genres. The following are the web series that we considered for our analysis:

TOP 10 INDIAN WEB SERIES IN VARIOUS GENRES				
1) THRILLER				
S.NO.	NAME	OTT PLATFORM	Date Of Release	IMDb RATING
1	Sacred Games	Netflix	05-Jul-2018	8.6
2	Delhi Crime	Netflix	22-Mar-2019	8.5
3	Mirzapur	Amazon Prime	16-Nov-2018	8.4
4	Breath	Amazon Prime	26-Jan-2018	8.3
5	Apaharan	ALT Balaji	14-Dec-2018	8.3
6	Criminal Justice	Disney+ Hotstar	05-Apr-2019	8.1
7	Inside Edge	Amazon Prime	10-Jul-2017	8.0
8	Patal Lok	Amazon Prime	15-May-2020	7.8
9	The Final Call	ZEE5	22-Feb-2019	7.2
10	Hostages	Disney+ Hotstar	31-May-2019	6.9
2) ACTION				
S.NO.	NAME	OTT PLATFORM	Date Of Release	IMDb RATING
1	Mai Hero Boli Raha Hu	ALT Balaji	20-Apr-2021	8.9
2	The Family Man	Amazon Prime	20-Sep-2019	8.8
3	Special Ops	Disney+ Hotstar	17-Mar-2020	8.6
4	Jeet Ki Zid	ZEE5	22-Jun-2021	8.5
5	The Test Case	ALT Balaji	30-Apr-2017	8.4
6	The Forgotten Army	Amazon Prime	24-Jan-2020	8.1
7	Rangbaaz	ZEE5	22-Dec-2018	8.1
8	Abhay	ZEE5	07-Feb-2019	8
9	Jamtara - Sabka Number Ayege	Netflix	10-Jan-2020	7.3
10	Bard of Blood	Netflix	27-Sep-2019	6.9
3) DRAMA				
S.NO.	NAME	OTT PLATFORM	Date Of Release	IMDb RATING
1	Aspirants	TVFPlay	07-Apr-21	9.7
2	Kota Factory	TVFPlay	16-Apr-2019	9.2
3	Kaafir	ZEE5	15-Jun-2019	8.4
4	Made in Heaven	Amazon Prime	08-Mar-2019	8.3
5	Little Things	Netflix	25-Oct-2016	8.3
6	Laakhon Mein Ek	Amazon Prime	13-Oct-2017	8.2
7	Afsos	Amazon Prime	06-Feb-2020	7.6
8	Selection Day	Netflix	28-Dec-2018	7.4
9	Masaba Masaba	Netflix	28-Aug-2020	6.7
10	Four More Shots Please!	Amazon Prime	25-Jan-2019	6.4
4) COMEDY				
S.NO.	NAME	OTT PLATFORM	Date Of Release	IMDb RATING
1	Gullak	TVFPlay	27-Jun-2019	9.2
2	Pitchers	TVFPlay	10-Jun-2015	9.1
3	Yeh Meri Family	TVFPlay	12-Jul-2018	9.1
4	College Romance	TVFPlay	07-Aug-2019	8.9
5	Panchayat	Amazon Prime	03-Apr-2020	8.8
6	Hostel Daze	Amazon Prime	13-Dec-2019	8.6
7	Chacha Chaudhary Hain Humare	Amazon Prime	18-May-2018	7.7
8	PushpaValli	Amazon Prime	15-Dec-2017	7.5
9	Comictaan	Amazon Prime	13-Jul-2018	6.8
10	Mind the Malhotras	Amazon Prime	07-Jun-2019	6.5
5) ROMANCE				
S.NO.	NAME	OTT PLATFORM	Date Of Release	IMDb RATING
1	The Married Woman	ALT Balaji	08-Mar-2021	9.4
2	Broken But Beautiful	ALT Balaji	29-Nov-2018	9.2
3	Bandish Bandits	Amazon Prime	04-Aug-2020	8.7
4	Baarish	ALT Balaji	25-Apr-2019	8.3
5	Kehne Ko Humsafar Hain	ALT Balaji	16-Mar-2018	8.2
6	Bebaakee	ZEE5	30-Aug-2020	8.2
7	Karrle Tu Bhi Mohabbat	ALT Balaji	21-Apr-2017	8
8	Out of Love	Disney+ Hotstar	22-Nov-2019	7.1
9	A Suitable Boy	Netflix	26-Jul-2020	5.7
10	Mismatched	Netflix	20-Nov-2020	5.5
6) HORROR				
S.NO.	NAME	OTT PLATFORM	Date Of Release	IMDb RATING
1	Ghoul	Netflix	24-Aug-2018	7.1
2	Betaal	Netflix	24-May-2020	7.1
3	Typewriter	Netflix	19-Jul-2019	6.5
4	Booo Sabki Phategi	ALT Balaji	27-Jun-2019	5.3
5	Ghost Stories	Netflix	01-Jan-2020	4.3

Top 10 Indian Web Series in various Genres

When we ran our code, following Excel files were generated. The following are the reviews for each genres:

1. ACTION

REVIEWS OF FAMILY MAN	
stupid people dont watch their disclaimers and watch only content but just judge the series ignorantly. actually its a very gud series with brilliant cast and binge watch series. Disclaimers say not true story but based on real published news articles. Real published news a	There is a considerably visible yet veiled line between revolution and terrorism. One man's Freedom Fighter is another man's terrorist. These are some of the tenets the Action-packed Season 2 of The Family Man had to offer, to dwell on, by using the glorious visage of
Amazon original series is about a middle-class man who juggles up his responsibility as spy and family man. This secretive special cell has a task at hand to find out what the terrorists are planning and stop it before it takes lives of millions. Manoj Bajpai in lead is a treat.	Enjoyed Family Man season 1 and season 2 . Very well made and exciting and thrilling it was. All the actors did their roles well. Raji character was so good, powerful performance. Wow... Also loved the comedy in the police station Sri and Talpade locked up and Umay.
Bang on! A 10 on 10. What a series by Raj Nidimoru and Krishna D.K! All the hype around ended me in binge watching both the seasons of Family Man. It is an edgy action-drama series, which tells the story of a middle-class man who works for a special cell of the Nation	Well made, smooth, funny and action packed series ruined by biased content and an anti climax, without closure....!! 04/10.Good words first...An extremely well made series, with world class level production that can really compete to some series from US and other w
11 Oh my god!!! Never expected the second season to be this extra thrilling...Manoj Bajpayee is truly a "family man" the way he calmed her daughter in her hard time...It was very moving, the way he gave time to his family, the efforts he put in his marriage. This series is	There is a considerably visible yet veiled line between revolution and terrorism. One man's Freedom Fighter is another man's terrorist. These are some of the tenets the Action-packed Season 2 of The Family Man had to offer, to dwell on, by using the glorious visage of
12 Indian serials based or espionage are getting better & better, first with "Special Ops" & now this " The Family Man" which in my opinion rises even higher. I watched "Special Ops just three weeks ago & I was not fatigued even a l'il bit! read the rating & just binge watch	Amazon original series is about a middle-class man who juggles up his responsibility as spy and family man. This secretive special cell has a task at hand to find out what the terrorists are planning and stop it before it takes lives of millions. Manoj Bajpai in lead is a treat.
13 A True Masterpiece... 200% better than so called bollywood big budget movies. THE GOOD: Great talent, the creators and crew starting with the DOP, BGM, Screenplay, and Editing, absolutely bang on the Money. Manoj Bajpayee - A World Class Actor whose true potential	Enjoyed Family Man season 1 and season 2 . Very well made and exciting and thrilling it was. All the actors did their roles well. Raji character was so good, powerful performance. Wow... Also loved the comedy in the police station Sri and Talpade locked up and Umay.
14 5/5 stars!This is the best webseries I ever watched. If they make webseries like this...Bollywood is under serious threat. Tiger Zinda hai nahi rahega..maara jayega.Bollywood is making one side useless spy movies like Tiger Zinda hai and Baghi...and here a 7 hour web serie	Bang on! A 10 on 10. What a series by Raj Nidimoru and Krishna D.K! All the hype around ended me in binge watching both the seasons of Family Man. It is an edgy action-drama series, which tells the story of a middle-class man who works for a special cell of the Nation
15 One of the best shows coming from India.Amazon Prime have a winner here. The Family Man, is what the international audience deserves from the Indian makers. The direction, performance of the actors, score and pacing of the series are all topnotch. People may find 1	Well made, smooth, funny and action packed series ruined by biased content and an anti climax, without closure....!! 04/10.Good words first...An extremely well made series, with world class level production that can really compete to some series from US and other w
16 Prodigious - Unmissable!Worth watching and worth waiting.The Family Man Season 2 is absolutely treat to watch. Despite having 9 episodes concluding in approx 7 hours i completed last night i.e. 03.05.2021 thanks to prime subscription.Srikant Tiwari is all set to loose	11 Oh my god!!! Never expected the second season to be this extra thrilling...Manoj Bajpayee is truly a "family man" the way he calmed her daughter in her hard time...It was very moving, the way he gave time to his family, the efforts he put in his marriage. This series is
17 The Family Mantells us about a middle class family man who is also a spy and how he tries to maintain a balance between both. The Family Man comprises of 10 episodes; which in detail narrates the story of Srikant Tiwari, who works for TASC! Threat Analysis and S	12 Indian serials based or espionage are getting better & better, first with "Special Ops" & now this " The Family Man" which in my opinion rises even higher. I watched "Special Ops just three weeks ago & I was not fatigued even a l'il bit! read the rating & just binge watch
18 Last year when the season 1 released I didn't saw it but in this year when I saw the trailer of the season 2 I fell like I should watch it and for that I first saw the first season and the second season and I truly say it's a fabulous series that I have ever seen. It was very suspe	13 A True Masterpiece... 200% better than so called bollywood big budget movies. THE GOOD: Great talent, the creators and crew starting with the DOP, BGM, Screenplay, and Editing, absolutely bang on the Money. Manoj Bajpayee - A World Class Actor whose true potential
19 Wow! Just one word! POWERFUL.Just completed watching in one go.The whole series is edge on seat experience, gripping, nail biting and suspenseful. Can't predict the next move or scene.Excellent performance by all the cast especially Manoj Bajpal sir. He never fails in	14 5/5 stars!This is the best webseries I ever watched. If they make webseries like this...Bollywood is under serious threat. Tiger Zinda hai nahi rahega..maara jayega.Bollywood is making one side useless spy movies like Tiger Zinda hai and Baghi...and here a 7 hour web serie
20 Was in the first 20 odd minutes completely immersed...After all, it has Manoj Bajpayee, a close to perfection actor. Huge expectations and with the build up in the first 20 mins, it soared high. Then comes a turn-off moment...one that shuts down the sync audiences have	15 One of the best shows coming from India.Amazon Prime have a winner here. The Family Man, is what the international audience deserves from the Indian makers. The direction, performance of the actors, score and pacing of the series are all topnotch. People may find 1
21 #TheFamilyMan Story of an Indian Spy - with a Family! Having watched first 3 episodes, am so far thoroughly impressed with the humor though a little overwhelmed by the action or thrill factors. The best part is #ManojBajpayee as expected and his deadpan humor in	16 Prodigious - Unmissable!Worth watching and worth waiting.The Family Man Season 2 is absolutely treat to watch. Despite having 9 episodes concluding in approx 7 hours i completed last night i.e. 03.05.2021 thanks to prime subscription.Srikant Tiwari is all set to loose
22 Just those 4.5 minutes of starring Samantha Akkineni in the second episode of season 2 brought tears in my eyes. However, I don't know anything about LTTE. The way she killed that pervert in a street made me clap instantly. I bow down to the creator Raj & DK and tc	17 The Family Mantells us about a middle class family man who is also a spy and how he tries to maintain a balance between both. The Family Man comprises of 10 episodes; which in detail narrates the story of Srikant Tiwari, who works for TASC! Threat Analysis and S
23 Binge watched and finished the show in a day. Even though the show revolves around the life of a person working with TASC under the NIA, the show is not too heavy on the mind. Subtle comedy, different types of relationships, family bonding and issues have been so l	18 Last year when the season 1 released I didn't saw it but in this year when I saw the trailer of the season 2 I fell like I should watch it and for that I first saw the first season and the second season and I truly say it's a fabulous series that I have ever seen. It was very suspe
24 Showing Anti national & terrorism in a good light won't. Do anything, infact it's diabolical for framing same and not atall showing the real pain of those who actually suffered due to these terror attacks & chemical bust. Anyways the show is actually gripping no doubt ...	19 Wow! Just one word! POWERFUL.Just completed watching in one go.The whole series is edge on seat experience, gripping, nail biting and suspenseful. Can't predict the next move or scene.Excellent performance by all the cast especially Manoj Bajpal sir. He never fails in
25 Verdict - 3.5The Family Man is really a good series but it has some serious issues too. Did you see earlier a man is hiding something from his family, his equation with his wife getting rotten day by day, children suspect they are verge of divorce, wife want to do new job a	20 Was in the first 20 odd minutes completely immersed...After all, it has Manoj Bajpayee, a close to perfection actor. Huge expectations and with the build up in the first 20 mins, it soared high. Then comes a turn-off moment...one that shuts down the sync audiences have
26 one of the best of all the series I have ever watched. all the star cast is fantastic, their acting and the character they have got matched beautifully, the script the storyline all are brilliant. It lets u know many facts about the Secret Services, their jobs, their daily life, and also	21 #TheFamilyMan Story of an Indian Spy - with a Family! Having watched first 3 episodes, am so far thoroughly impressed with the humor though a little overwhelmed by the action or thrill factors. The best part is #ManojBajpayee as expected and his deadpan humor in
27 "The Family Man" is an amazing spy-thriller-action web series where Manoj Bajpayee does amazing acting as being an Indian Spy. The overall cast such as the wife, kids, colleagues, agents have done an amazing job too. Each episode goes with the flow and keeps the aud	22 Just those 4.5 minutes of starring Samantha Akkineni in the second episode of season 2 brought tears in my eyes. However, I don't know anything about LTTE. The way she killed that pervert in a street made me clap instantly. I bow down to the creator Raj & DK and tc
28 Hats off to the direction, actors, acting, story line.All the characters were perfect in their role. You will never feel that some one else should have done this or that role , all were damn awesome.Bonus points to Manoj Bajpal acting , it was superb, the way he changes quick	23 Binge watched and finished the show in a day. Even though the show revolves around the life of a person working with TASC under the NIA, the show is not too heavy on the mind. Subtle comedy, different types of relationships, family bonding and issues have been so l
29 I had watched all 10 episodes in one go and trust me , not even a single episode makes u feel pausing it for the next day.Every aspect of the series keeps u busy entertaining all the time.From family issues to the office problem , every dialogue, cinematography, situation	24 Showing Anti national & terrorism in a good light won't. Do anything, infact it's diabolical for framing same and not atall showing the real pain of those who actually suffered due to these terror attacks & chemical bust. Anyways the show is actually gripping no doubt ...
30 wow there is another show which displays the unintelligence of our country's intelligence as this one. if this is based on real stories and a premium intelligence agency is working without basic policing sense then India is really doomed. our main lead is India's top intellig	25 Verdict - 3.5The Family Man is really a good series but it has some serious issues too. Did you see earlier a man is hiding something from his family, his equation with his wife getting rotten day by day, children suspect they are verge of divorce, wife want to do new job a
31 What Indian serials should be doing web series are doing and here I am proudly saying Indian web series are the best (not comparing with other only with Indian serials). Highly recommended series. You will not regret watching it. First few episodes are slow but again it	26 one of the best of all the series I have ever watched. all the star cast is fantastic, their acting and the character they have got matched beautifully, the script the storyline all are brilliant. It lets u know many facts about the Secret Services, their jobs, their daily life, and also
32 Family Man is an amazing thriller starring Manoj Bajpayee as Shrivardhan a clonial middle class man working as a constent agent for Threat Analytic and Surveillance Cell (TASC) – an organisation which surveillance and identifies terror threats . In the haanline of the show 3 ISIS	"The Family Man" is an amazing spy-thriller-action web series where Manoj Bajpayee does amazing acting as being an Indian Spy. The overall cast such as the wife, kids, colleagues, agents have done an amazing job too. Each episode goes with the flow and keeps the aud

Review file for Action Reviews.

2. COMEDY

A must watch with your Yaar and Pyaar	
A Best webseries	Best webseries
Bagga rocks	Bagga rocks
Bang on slow	Bang on slow
Hats off to The Timeliners	Hats off to The Timeliners
Must watch	Must watch
College romance	College romance
Best web series!!	Best web series!!
Best Indian Web series on YouTube	Best Indian Web series on YouTube
Solid	Solid
Best	Best
The Best	The Best
great one by timeliners	great one by timeliners
I Loved It	I Loved It
Amazing	Amazing
Epic!!!	Epic!!!
perfect comedy web series	perfect comedy web series
A must must watch webseries. One of the finest one from The Timeliners.	A must must watch webseries. One of the finest one from The Timeliners.
Superb	Superb
Mzaaaaaaa gya	Mzaaaaaaa gya
Best it is,best it will be	Best it is,best it will be
Great webseries	Great webseries
Do not watch Honest review	Do not watch Honest review
MUST WATCH!!!!	MUST WATCH!!!!
College Romance - the timeliners way	College Romance - the timeliners way
Most ridiculous show ever	Most ridiculous show ever
Not boring	Worth watching

Review file for Comedy Reviews.

3. DRAMA

The screenshot shows a spreadsheet titled "REVIEW OF ASPIRANTS" in cell A2. The content consists of 25 numbered reviews. The reviews are in a single column, each starting with a number from 3 to 27. The reviews are in a conversational tone, discussing various aspects of the series like acting, message, and production quality. Many reviews include emojis. The spreadsheet has a standard header row with columns A through T. The bottom navigation bar shows tabs for "ASPIRANTS", "KOTA FACTORY", "KAAFR", "MADE IN HEAVEN", "LITTLE THINGS", "LAHKON MAI EK", "AFSOS", "SELECTION DAY", "MASABA MASABA", and "FOUR MORE SHOTS PL". The status bar at the bottom indicates "Average=0 Count=2809 Sum=0".

REVIEW OF ASPIRANTS

3 This series is really inspiring, good message, light hearted and Sandeep bhaiya ❤️ ❤️ ❤️ ❤️ ❤️ ❤️ ❤️ ❤️ ❤️ ❤️ ❤️ ❤️ ❤️ ❤️ by Sunny Hinduja really he is getting appreciation what he deserves Sandeep bhaiya ko de
4 Just amazing, this series is just real it could ever get. everything is amazing, hats off to the team. If you haven't watched it yet then i would really suggest this banger series....
5 One of the best series I ever seen. I like the way the characters played there roles. This show is too much motivational for everyone. Especially the character of Sandeep bhaiya was too good.
6 So involving and absorbing. Amazing job ... All actors have done an amazing job .. indian film industry, (barring few theatre actors) shame on you ... So called hollow commercialized cinema clowns ... No hiding ... Public is ei
7 For someone who has spent prime of their life in the 'system' of competitive exams, this is so relateable. The drama between friends/ girlfriends gives snippets into the lives of the individuals, what all they go through. Over
8 This is my first ever review to any series or movie...I couldn't stop myself to write about this. It is a must watch series. Just awesome ❤️ ❤️
9 It's a very good series and a very good message for everyone in this series is "Don't give up ,just keep hard working and have some patients also..one of the day are you success in your life ❤️ just do it♦ ♦
10 Binge watched the whole thing! Brilliant ! 5/5 for Sandeep baiyah... Hell Hall Such sad eyes! Very well done show... Please make more of these kinda shows...
11 awesome web series. It will inspire you, make you laugh ☺ and at the same time make you cry 😢 in just it's 5 episode.
12 A perfect web series for 13+ and for non-IAS aspirants. Very inspiration, funny and has a good message. A must watch.
13 Very useful and enjoyable series for aspirants. I like this series. And after watching this I am motivate and thinking positive approach 😊
14 Five star for at least thinking to create this kind of series. Totally loved this aspirants and the lines taken totally do justice with the cinematography.
15 mind blowing .showing reality of our society heart touching story inspiring as well. we should appreciate such films proud of all characturs
16 Pretty Awesome!! Real struggle to get mentally prepared for UPSC!! Sunny Hinduja aka sandeep bhaiya is 1 number 🙏 🙏 🙏
17 Speechless!!! ❤️ ❤️ ❤️ Motivation...Friendship Goals...Ego...Success..Failure..Importance of Relationship... Hardworking...Plan B..Sub Kuch Sikha Diya Yaar...Fantastic job by all Star Cast 😍 😍 😍 Thanks TVF 😊 😊 😊
18 As I am also a student and I know very well about the routine of a student, after watching this series got a lot of motivation I think every student should watch this series.
19 Awesome webseries,a lot of things aspirants can understand by watching this great series and congrats all aspirants team for making this great content 🙏 🙏 🙏
20 This is a really inspiring, meaningful and motivating series. Highly recommended for all UPSC aspirants
21 Inspiration story...and one of the best motivational story for government job...Nd Last I'm heartly Salute for star cast... writer ... nd director...
22 Absolutely mind-blowing, ultimate series.Loved it.Roads we walked together with friends ❤️
23 Masterpiece 🙏 🙏 . Recommended for all UPSC aspirants. It'll never give you disappointment.
24 best motivational, inspirational and emotional series ever. Love TVF
25 There is a reason TVF shows are so amazing because the shows are not fiction, they are just lives !

Review file for Drama Reviews.

4. HORROR

The screenshot shows a spreadsheet with reviews for the horror series "Betaal". The reviews are in a single column, numbered 1 to 28. The reviews are in a conversational tone, discussing the plot, acting, and atmosphere of the series. The spreadsheet has a standard header row with columns A through T. The bottom navigation bar shows tabs for "GHOUЛ", "BETAAЛ", "TYPEWRITER", "BOO SAKI PHATEGI", and "GHOST STORIES". The status bar at the bottom indicates "Average=0 Count=1882 Sum=0".

first timer seeing indian original content -
1)Betaal Lok - WOW ! Great acting and amazing story telling. GOOD JOB !
2) Based on that, lets see another original series - Mirzapur - I am impressed by the original content and talent by Indian actors. I bated breath waiting for mirzapur 2....
3) OK...my expectations are really built up...lets do justice and see the latest offering - Betaal.....and my reaction
a) ARE YOU KIDDING ME????? the production team has taken a solid foundation zombi ...
4) After a very long time, I've watched a web series and that too in one go. Betaal seemed to intrigue me right from the time I discovered about it in the coming soon section of Netflix. Having just finished a few minutes ago, I
5) Betaal is our "Made in India" project for supernatural stories. For years, we have been either reprising the same old jump scares with little or no need to the story or we have been copying ones from the West. Either ways,
6) Betaal is a really good series to watch on Netflix. After Tumbbad, I think this series is the only mascot for Indian horror(although tumbbad was much superior). Except some direction and screenplay issues, this series is a g
7) I thoroughly enjoyed watching it. God knows what these news papers look into while judging and rating. It was full action packed horror with a wonderful storyline. Only thing which I feel could have been better is the costl
8) Zombie movies or series are one among the genre which is highly left untapped by Indian market probably because it is hard to get it right and in the midst of that makers have tried to tap that market.A story of village in w
9) Oh boy where do I begin. Spoilers galore ahead. Started off so promising, and then it turned into this massive frustration fest that made me sigh more than a few times. The premise was really great. Trapped evil waiting to
10) What was that!!! I mean really, what was the series trying to do! I know it's a new trend of folklore & myths content based webseries in India after era of biopics. But really, Just for the sake of myth, you can't just put a
11) I loved it. It was really engaging. First episode was superlative! The climax is scary though, alluding to the must-dreaded armageddon!Some self-proclaimed critics are surely gonna point out a gazillion flaws, but as a matter
12) Well this series is 3.5/5 where half a star is for the efforts of the actors. Well to be very honest the series started so well that I was surprised to have Hollywood experience in Bollywood premises and it was going so good un
13) I had really high expectation since folks compared it to Tumbbad, I mean Tumbbad was an epic, and Betaal is a JOKE. Just because you create a bleak dark atmosphere and rainy days doesn't mean it can get scary. For follow
14) I think Bollywood should never attempt to make something like this. So called Zombie, moreover it's really disappointing that Bollywood really tries to copy things from Hollywood, but haven't yet managed to successfully p
15) So, I have binged watch the whole show. And from the very first episode it shows you that it has the capability to engage from the whole mythological zombies story and from episode to episode it explores you the story be
16) Lame spoiler Alert! An O.K watch if you just want to indulge yourself into a tense web series (tries very hard to be one, but fails). I read some positive audience reviews on the same platform, ignored the critics reviews and
17) Don't know why there are bad reviews out there. I found it to be absolutely breathtaking and it had enough jump scares to qualify in the horror genre. The concept gets out secrets linked to the British era, which makes the
18) Wow..! 😊 Great start with THE BEST Zombie horror on screen from Bollywood it really deserves appreciation. First Indian horror web series that has horror story, & emotional scenes. A very good and successful attempt
19) Betaal is a perfect zombie thriller for Indian audiences. Binge worthy and flawless. 4 episodes but each episode keeps you entertained and wanting for more. This one truly deserves a second season. It is not at all anti natic
20) Nice Horror series by Netflix & red chilles. It is very short & fast paced series that introduce us to zombies & evil spirit. Bollywood doesn't know how to make a good Horror movie except few good movies like Tumbbad,13B,A
21) One star for the cast and the other one for art direction and sets. The beginning of the story was inquisitive and they could manage to continue that in the second episode also but as it happens in most horror movies or sh
22) This one is a good try from Indian filmmakers and should atleast receive appreciation but same can't be said for the series. It is dark but not gripping. It seems to have many unnecessary sub plots and therefore, story falls f
23) A great show created by one of the true emerging masters of the horror genre, this is Patrick Graham's second outing after the delectably suspenseful 'Ghoul'. Like in that show, Graham understands the innate principle of t
24) Red chilles entertainment's first attempt in this genre, and they did pretty well! First season of Betaal has 4 episodes. A crisp and tight plot of a curse in a tribal mountain which gets unleashed by a militant squad while ex
25) After watching great epic Zombie 🎬 , Tv Shows and Movies 🎬 , Here comes Indian version of the same with 3 Hours Long TV 🎬 Series , 4 Episode. Story line is good but could have been way better , Unique presentatc
26) JUST SIMPLY IGNORE ALL THE BAD REVIEWS AND WATCH FOR YOURSELF..film making ke baare me ghanta pata ni kisi ko aur aae bare judgement dene..Just grab your snacks and watch this finest Indian thriller/horror series!
27) Indeed this is an amazing series. After so much time finally we could see zombies getting some respect in Indian Film and Series industry (pun intended). But only three stars because there were certain loop holes in the sto
28) Watched Betaal. I must say this total Indian take on The Walking Dead by Shahrukh Khan's production didn't disappoint. Lot of characters with lots of possibilities and an open season ending. Special mention to the effect

Review file for Horror Reviews.

5. ROMANCE

CLASSICAL MUSIC MEETS FUSION		
From a Director cum Actor Anand Tiwari who you might recognise from Udaan and Go Goa Gone, a fairly less known face in Indian cinema and creator Amritpal Singh Bindra bring a series about Indian classical music that goes beyond your imagination and anything you have ever experienced in Indian Cinema.		
It is definitely the Magnum Opus of Anand Tiwari until the next season comes along.		
This series is absolutely a masterpiece and nothing short of a miracle to be pulled ...		
A web series that makes me feel proud of our very own Indian classical music,a web series that will touch your soul instantly & make you drool over its mind-blowing music,a web series that India deserved long back !		
Bandish Bandits is a masterpiece		
A web series that makes me feel proud of our very own Indian classical music,a web series that will touch your soul instantly & make you drool over its mind-blowing music,a web series that India deserved long back !		
Bandish Bandits, what a pleasant surprise ! In fact, this show turned out to be an exquisite and luxuriant treat to not just my ears, but also my eyesStarting with the basic theme, this series mets modernity. This is not some c		
Its a musical saga based on two opposite individuals who happen to meet each other but gradually re-discover themselves in the journey. This story takes us through a plethora of emotions - be it love, hatred, anger, grit, e		
MOVIES REVIEW AND FUN PAGE.BANDISH BANDITS: A mesmerising and inspiring musical experience. A treat to music lovers.After a long time i got engage to something this karismatic that i watched all 10 episodes about 4C		
ANTIQUE MUSICAL DELIGHT Series is gem for music lover who wants to understand the core of Indian Music. Before this series I only read about classical music for my UPSC preparation but know I started enjoying it.I would		
BANDISH BANDITS:A MUCH NEEDED COLLABORATION.I was Quite contemplating on Binge watching it. I was Quite drained watching Daily opera and wanted something New.something different.Finally Gave in and Watched		
Bandish Bandits on Amazon Prime.Start with the theme song and it rings for ever in the ears. What a great webseries whose music was carved by the musicians Shankar Ehsaan Loy. In particular the singers were outstanding;		
Since childhood in really fascinated towards our culture, i love the way our Indian mythology presents its ideology through story telling , i love music too but i never got that attachment with our classical music it might beca		
Classical music is a genre which usually isn't appealing for the current generation but this one proves it wrong with all the right reasons.A tale set at the backdrop of Rajasthan; two die hard musicians accidentally meets to		
In the competition of Netflix Vs Prime, the latter one really knows how to play the game, they know what Indian audience want to see on OTD. Bandish Bandits has its flaws but a big thumbs up on its fresh content. Extremely		
Bandish Bandits(Prime Webseries)Rating : ****Binge Worthy: Yes ✓ Unique story stitched in a retro metro setup where classical meets fusion, its one if its kind show that comes out as a winner for a never before touched		
It was suggested by my friend and i must thank her for suggesting this amazing web series. I was really delighted to watch this series, I watched total 10 episodes in one go, could not leave it in mid.As Tamanna(Sheyra Ch		
The title of this web series didn't attract me, but as I watched the episodes, the content was just awesome.. a tribute to our original Indian Classical music singers and those who have kept it intact and flowing to the commi		
Yes story has a lot of weight for a good series and though bandish bandit's plot is loosely woven, you can't miss the splendor by which each actor delivers. Cinematography is eye-catching and brings Rajasthan to your screen		
It's a musical treat for all those who appreciate classical Indian music. With the backdrop of beautiful city of Jodhpur, this story weaves a plot which at the onset looks like just an "east meets west" story but inside, it's a mc		
A teleseries made with so much passion that you feel you are watching a movie, one can marathon watch all 10 episodes of 40 minutes each in one go ! A mega treat for music lovers, an epic musical romance drama, class		
CLASSICAL AND FUSION LOVE STORY!This is the most amazing web series i have ever watched on Amazon! A must watch for all. Music and its importance is shown so beautifully. But western music has been pushed down o		
Beautiful experience of classical music. Great acting performance by entire cast, refreshing n best original music. Each song is magically created. Climax is unforgettable the music n acting n the feel will remain for long time.		
The wait has been worth it. Bandish Bandits reiterates that Anand Tiwari's success with Love Per Square Foot (LPSF), starring Vicky Kaushal, Angira Dhar and Raghubir Yadav was no flash in the pan. LPSF presented Vicky Kaus		
"BANDISH BANDITS" - Classical music meets fusion in the most miraculous, meritorious and passionate way.What a music. I have never enjoyed, loved, been touched and felted classical music the way I did after watching thi		
What a great escape from the ongoing world of crime thriller and suspense dramas. SuuuuuppperrR.From Story to Direction to Cinematography to Performances to Music. OMG, it is a completely satisfying package.Shank		
Just One of the best series of Amazon Prime ever made. In this OTT world of crime and thriller , this series makes a difference by itself in the whole OTT universe with a whole new concept which I believe that we haven't seen		
Outstanding,phenomenal, words are less to describe. Beautiful journey taking together classical and western music hand in hand. Rich music intertwined with exceptional performances by the entire cast makes it a must w		
27		
28 Naaseer salhab is an Institution...rightly said. Those aspiring actors should learn from his acting. The expressions were simply mind blowing...thanks to the DOP for such close shots. On the other hand Sheeba Chaddha has sho		
THE MARRIED WOMAN BROKEN BUT BEAUTIFUL BANDISH BANDITS BAARISH KEHNE KO HUMSAFAR HAI BEBAKEE KARRLE TU BHI MOHABBAT OUT OF LOVE A SUITABLE BOY		
Average=0 Count=2098 Sum=0		

Review file for Romance Reviews.

6. THRILLER

REVIEWS OF SACRED GAMES (THRILLER)		
2 REVIEWS OF SACRED GAMES (THRILLER)		
3		
4		
5 Had read the sprawling world building novel by Vikram Chandra nearly a year back; so was waiting with bated breath for the TV adaptation. Also met the author at a literary event; and asked him about the same.Chandra s		
Simply brilliant content to binge watch. Stellar performances, phenomenal score&beautiful writing. It takes you to the core essence of the series very quick.the story is very tight,no loose ends at all.. mind boggling directio		
7 It is a must watch series and for me ,it was the first time I was watching an Indian series after I watched Mahabharat when I was a kid in the 90s. The approach to the series seemed raw and gritty,with all the vagaries of		
8 It's indian storytelling at its peak. What I loved about the series was its originality and how writers have been able to skillfully adapt vikram chandra's novel into the series. What is commendable is that it is a pure desi indica		
9 Sacred Games explores a wide array of narratives and has multiple layers to the complexity of its characters. This is by far the only thing I could come up with if talking about the good points of Sacred Games season 2. After		
10 SACRED GAMES is an intense series. The plot, the twists, the action: everything says it well. All the episodes have something so unique that it's hard not to watch all of them back-to-back. After a long time, you get to see S		
11 Sacred Games is the premiere Hindi Netflix series in India. Produced by Phantom production who are known to invest in edgy stories.Based on the 1000 page Vikram Chandra novel by the same name. The story is, for the la		
12 Sacred Games. Netflix's first original Indian web series is Masala Chai but made for the international audience. Prepared with Indian ingredients of flavour, colour and emotion, steamed in western production values and ta		
13 The trailer of #Sacred games already had me hooked. I was thinking of #saifali Khan and when was the last time that he brought out his best and I was reminded of omkara. #sacredgames is another opportunity where he l		
14 A phenomenal debut by India since Sacred Games is her first series on Netflix. The amazing story, direction, camera works is just some of the very few elements that makes this show such a great one. The portraying of Mu		
15 I was looking forward to sacred games, adapted from the novel by Vikram Chandra and there was never a moment that let go of my attention. It's definitely binge worthy and I would rate it on par with the recent 'Unabon'		
16 I read the book about 5 years ago. It did put a certain demand on the reader to familiarize himself with the many Hindi words and songs referenced in the text, but overall I was highly impressed and counted it among my l		
17 What an debut by India with it very own Netflix series.Very well written, directed and well shot. Use of Mumbai is brilliant the story narrative has its own pace and is not rushed into but it holds the audience attention th		
18 Sacred Games is definitely one of the best thriller shows that I have seen. This is no ordinary story of a cop and a gangster, it has a twist.The show is very well written and the screenplay is amazing too . It truly lives up to it		
19 Go for it. Show ho tho, "Isa" horn chahiye!(Pardon my Hindi, I'm a Kannadiga)When you have got a crime story with beautiful screenplay, direction, and brilliant actors; you know the show is gonna be successful. Netflix ha		
20 It took a lot of peer reviewing for this series and that is the reason to watch it late by the way, there are some good criticisms in here on this series, an adaptation from and of a thousand page book. Obviously it has to be big f		
21 The content is factual in a fiction series. Sacred Games is an attempt to reunite people across faiths basis crime. Performances are just mesmerizing, the looks are cleverly looked at. Lot of work has gone into making it. It ta		
22 Sacred Games Season 1 is about rich direction great adaptation and strong performances more noticeably from Saif Ali Khan, you will see Saif giving maybe the best performance of his career.While Season 2 is very differen		
23 This is first Indian netflix original series directed by Anurag Kashyap and Vikramaditya Motwani.First 30 seconds in the first episode and you'll get to know that this series is going to be dark, edgy and controversial. This seri		
24 4 stars. Well written, well executed. Netflix running high on their previous success in the "crime-drama" genre produces sacred games is not a typical good cop - bad gangsta story. In fact the story is a great statement on t		
25 My Ratings: 8.5/10A Phantom production, Anurag Kashyap, Nawazuddin Siddiqui, Radhika Apte and Saif Ali Khan. Well I guess, that's all we need to create something out of the box. Sacred Games turns out to be a box whi		
26 As an ardent bibliophile, heavily inclined towards Bengali and English literature alongwith archaic yearn for Urdu poetry, I never understood Indian television serials. Barring a few webseries and some interesting literary ac		
27 Sacred games is such a star studded Indian based Netflix original series . sacred games is one of its kind of series which has many great well rounded hollywood stars coming together and acting . it was such a hunder coris		
SACRED GAMES(THRILLER) DELHI CRIME (THRILLER) MIRZAPUR (THRILLER) BREATHE (THRILLER) APAHARAN (THRILLER) CRIMINAL JUSTICE (THRILLER) INSIDE EDGE (THRILLER) PAAT/		
Average=0 Count=1610 Sum=0		

Review file for Thriller Reviews.

4.2 Data Preparation

4.2.1 Reading the Data

All the web scraped review excel files were opened in Jupyter Notebook via Python (via Pandas library), where our further analysis were done.

1. ACTION

```
import pandas as pd #for working with dataframes

#Reading the webscraped reviews of all the top 10 webseries of action genre.
r1_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ACTION REVIEWS.xlsx",0)
r2_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ACTION REVIEWS.xlsx",1)
r3_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ACTION REVIEWS.xlsx",2)
r4_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ACTION REVIEWS.xlsx",3)
r5_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ACTION REVIEWS.xlsx",4)
r6_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ACTION REVIEWS.xlsx",5)
r7_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ACTION REVIEWS.xlsx",6)
r8_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ACTION REVIEWS.xlsx",7)
r9_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ACTION REVIEWS.xlsx",8)
r10_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ACTION REVIEWS.xlsx",9)
```

2. COMEDY

```
#Reading the webscraped reviews of all the top 10 webseries of comedy genre.
r1_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\COMEDY REVIEWS.xlsx",0)
r2_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\COMEDY REVIEWS.xlsx",1)
r3_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\COMEDY REVIEWS.xlsx",2)
r4_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\COMEDY REVIEWS.xlsx",3)
r5_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\COMEDY REVIEWS.xlsx",4)
r6_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\COMEDY REVIEWS.xlsx",5)
r7_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\COMEDY REVIEWS.xlsx",6)
r8_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\COMEDY REVIEWS.xlsx",7)
r9_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\COMEDY REVIEWS.xlsx",8)
r10_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\COMEDY REVIEWS.xlsx",9)
```

3.DRAMA

```
#Reading the webscraped reviews of all the top 10 webseries of DRAMA genre.
r1_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\DRAMA REVIEWS.xlsx",0)
r2_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\DRAMA REVIEWS.xlsx",1)
r3_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\DRAMA REVIEWS.xlsx",2)
r4_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\DRAMA REVIEWS.xlsx",3)
r5_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\DRAMA REVIEWS.xlsx",4)
r6_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\DRAMA REVIEWS.xlsx",5)
r7_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\DRAMA REVIEWS.xlsx",6)
r8_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\DRAMA REVIEWS.xlsx",7)
r9_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\DRAMA REVIEWS.xlsx",8)
r10_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\DRAMA REVIEWS.xlsx",9)
```

4.HORROR

```
#Reading the webscraped reviews of all the top 10 webseries of HORROR genre.  
r1_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\HORROR REVIEWS.xlsx",0)  
r2_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\HORROR REVIEWS.xlsx",1)  
r3_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\HORROR REVIEWS.xlsx",2)  
r4_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\HORROR REVIEWS.xlsx",3)  
r5_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\HORROR REVIEWS.xlsx",4)
```

5.ROMANCE

```
#Reading the webscraped reviews of all the top 10 webseries of ROMANCE genre.  
r1_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ROMANCE REVIEWS.xlsx",0)  
r2_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ROMANCE REVIEWS.xlsx",1)  
r3_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ROMANCE REVIEWS.xlsx",2)  
r4_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ROMANCE REVIEWS.xlsx",3)  
r5_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ROMANCE REVIEWS.xlsx",4)  
r6_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ROMANCE REVIEWS.xlsx",5)  
r7_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ROMANCE REVIEWS.xlsx",6)  
r8_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ROMANCE REVIEWS.xlsx",7)  
r9_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ROMANCE REVIEWS.xlsx",8)  
r10_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\ROMANCE REVIEWS.xlsx",9)
```

6.THRILLER

```
#Reading the webscraped reviews of all the top 10 webseries of THRILLER genre.  
r1_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\THRILLER REVIEWS.xlsx",0)  
r2_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\THRILLER REVIEWS.xlsx",1)  
r3_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\THRILLER REVIEWS.xlsx",2)  
r4_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\THRILLER REVIEWS.xlsx",3)  
r5_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\THRILLER REVIEWS.xlsx",4)  
r6_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\THRILLER REVIEWS.xlsx",5)  
r7_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\THRILLER REVIEWS.xlsx",6)  
r8_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\THRILLER REVIEWS.xlsx",7)  
r9_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\THRILLER REVIEWS.xlsx",8)  
r10_df=pd.read_excel(r"C:\Users\Asus\Desktop\Intern Work-IIM Ranchi\ALL REVIEWS\THRILLER REVIEWS.xlsx",9)
```

4.2.1 Data Concatenation

Combining all review data frames into one data frame for our further analysis on it.

1. ACTION

```
#combining all the review dataframes into one dataframe  
combined_df = pd.concat([r1_df, r2_df,r3_df,r4_df,r5_df,r6_df,r7_df,r8_df,r9_df,r10_df], ignore_index = True)
```

```
# Let's take a look at the updated df  
combined_df
```

	transcript
0	NaN
1	REVIEWS OF MAI HERO BOLL RAHA HU
2	NaN
3	Parth just nailed the role of nawab...as he re...
4	I love the series .it is amazing. Parth acting...
...	...
9896	Not seen yet
9897	entertainment is just for entertainment.
9898	AVERAGE
9899	Time pass series .
9900	Please avoid this one.
9901 rows × 1 columns	

```
combined_df.sample(10)
```

	transcript
1478	One of the best web series of recent time..and...
529	Best Indian web series. Family man season 2 is...
913	Simply outstanding nd Amazing..nd see..who has...
5469	Any chances of season 2?? 😊😊
5257	The best Indian web series...❤
6110	It was great to know about the bravery, courag...
731	*****SPOILERS***** \n\nit is one of the most ...
485	Merit\n• Manoj vajpayee's acting, he's flawles...
5449	want season 2 immediately.....
6765	The acting in the show is excellent. The show ...

2. COMEDY

```
# Let's take a look at the updated df  
combined_df
```

	transcript
0	Nostalgic
1	TVF has done it again, Brilliant series.
2	Please dont stop. Continue journey with Season...
3	Story which will seriously tickle its audienc...
4	Must watch web series
...	...
8006	2.5/5\n\nWatch for light hearted, rib tickling...
8007	time wasting series
8008	worst.....waste of time
8009	forced to act crazyy
8010	watch sumit sambhal lega instead

8011 rows × 1 columns

```
combined_df.sample(10)
```

	transcript
803	Really good I like it. waiting for next episode
3826	Out of all the web series i find this one as o...
2057	A not to be missed gem
431	The best series based on middle class family. ...
1995	Best TVF show ever
3127	Bombasting.....
3731	If u want to really enjoy ur time u should wat...
7934	super funny! really hope theres a season 2
6493	hilarious simplicity and comedy.... reminds m...
3575	A series you must watch

3. DRAMA

```
# Let's take a look at the updated df  
combined_df
```

	transcript
0	REVIEWS OF ASPIRANTS
1	This series is really inspiring, good message,...
2	Just amazing. this series is just real it coul...
3	One of the best series I ever seen. I like the...
4	So involving and absorbing. Amazing job ... Al...
...	...
6900	Where's 2nd part.....?
6901	Waiting for more shots.. season 3....?
6902	I want season 2!!
6903	One word "Worst!!!!!!!!!!!!!!!"
6904	Story pattern is same as season 1.

6905 rows × 1 columns

```
combined_df.sample(10)
```

	transcript
1066	Great show!
4690	It's very commonly known thing..... Very rela...
6754	Season 2 is a disaster , Nothing relatable , U...
2263	Series shows the truth about what the UPSC Asp...
5701	Many of us cannot keeping up with Kardashian. ...
2735	Tvf always make unique content.i love tvf .all...
396	Hats off to the makers.
6725	Nice wonderful series 🔥🔥😎😊
5914	To be honest, when i first saw season 1, it wa...
310	Masterpiece, Must Watch, Very Inspirational..

4.HORROR

```
# Let's take a look at the updated df  
combined_df
```

transcript

0 2/5\n\nGhoul is another offering by Netflix in...
1 Ghoul or Ghul is not a fascinating watch as a ...
2 It was a good horror miniseries, if not the be...
3 I came very late across this miniseries. And i...
4 Ghoul is a different kind of horror that makes...
...
4249 Please someone explain Anurag kashyap story?
4250 fuack this\n.....
4251 #JusticeForSSR, #BoycottBollywood
4252 Bakavass time vesting and illogical
4253 Average.....

4254 rows × 1 columns

combined_df.sample(10)

transcript

2644 Not much justice has been done to the story an...
4172 Booooooring!!! Really really bad.
1802 Mind blowing it was🔥🔥🔥🔥🔥\nIt was very scared ...
1321 A chunk of elegant occultist claptrap! If you ...
2861 All this nonsense makes you want to watch what...
1616 if u have watched hollywood zombie and horror ...
3020 The story is good and the chemistry between th...
2511 The first episode was really dull, I could bar...
1568 Not bad not good, just lockdown to, as far as ...
149 Don't expect too much out of it! It was okay! ...

5. ROMANCE

```
# Let's take a look at the updated df  
combined_df
```

transcript

0 Even though I am no expert but I found the cam...
1 I stumbled upon this series randomly while sur...
2 I was prejudice towards watching The Married W...
3 I am sure we will wait a long while till somet...
4 10/10 without thinking and would recommend to ...
...
6040 The only good thing about series is Rohit
6041 It was wasted of time 😭😭💔💔
6042 Hated it. total waste of time.
6043 Goddd..such a waste of time 🤦
6044 worst show ever!!!!really disappointed.

6045 rows × 1 columns

combined_df.sample(10)

transcript

1490 Oh my god it is just awesome and please come b...
5690 This story takes of great showing how there ar...
4093 I love imtize ❤️❤️❤️❤️❤️🥳🥳🥳
1354 Nothing less than Brilliant. Their music bring...
3209 It's very much lovely web series I never seen ...
5022 Absolutely brilliant. You will see without ge...
4725 Terrific story-line and all the artists who ar...
2649 loved everything about musical series Bandish ...
494 This is just so so beautiful... ❤️ I haven't g...
2794 Superb Series ever watched.. I really love it.

6. THRILLER

```
# Let's take a look at the updated df  
combined_df
```

	transcript
0	REVIEWS OF SACRED GAMES (THRILLER)
1	NaN
2	NaN
3	Had read the sprawling world building novel by...
4	Simply brilliant content to binge watch. Stell...
...	...
8443	Good one, but was it necessary to use the dirty...
8444	Waiting for s3 season 2 better than s1 \n...
8445	Worth watching but they could have done better.
8446	Series is good but too many loose holes...
8447	Very Interesting and Full of Twist show...

8448 rows × 1 columns

```
combined_df.sample(10)
```

	transcript
7652	Highly abusive, created by a sick mind,
664	Awesome must watch only those who like story an...
4437	One of the best series..fantastic
1604	To download
6129	Indian webseries on Netflix and primevideo see...
4122	This series is fantastic and fabulous 🙌🙌🙌😊😊😊❤...
7875	Crappy story, pathetic acting, poor direction ...
3269	Those who feel this as unrealistic are either ...
2667	Excellent web series.... Totally full of all t...
2757	Mirzapur is the true definition of what pankaj...

4.2.2 Data Cleaning

Let's take a look at the updated data frame after three rounds of data cleaning

1. ACTION

```
#Updated dataframe after three rounds of data cleaning  
data_clean_df
```

transcript	
0	nan
1	reviews of mai hero boll raha hu
2	nan
3	parth just nailed the role of nawabas he recei...
4	i love the series it is amazing parth acting i...
...	...
9896	not seen yet
9897	entertainment is just for entertainment
9898	average
9899	time pass series
9900	please avoid this one

9901 rows × 1 columns

data_clean_df.sample(6)	
transcript	
652	exciting good direction and script couple of e...
3647	mind blowing web series also saw such a high b...
7398	this is something extraordinary one can never ...
1289	just awesomemanoj bajpaei is just an awesome a...
8603	the budget and the star cast used in this seri...
6115	a treat to watch will fill your heart with pat...

2. COMEDY

transcript	
0	nostalgic
1	tvf has done it again brilliant series
2	please dont stop continue journey with season ...
3	story which will seriously tickle its audience...
4	must watch web series
...	...
8006	watch for light hearted rib tickling humor
8007	time wasting series
8008	worstwaste of time
8009	forced to act crazy
8010	watch sumit sambhal lega instead

8011 rows × 1 columns

data_clean_df.sample(6)	
transcript	
5191	if you want to know how to ruin sequel of a go...
7950	not goodnot relatableas indian we are more rel...
7242	refreshingly accurate depiction of a typical s...
3992	an awesome series with cool cast specially i e...
7040	nice good
1880	so glad this was made

3. DRAMA

transcript	
0	reviews of aspirants
1	this series is really inspiring good message I...
2	just amazing this series is just real it could...
3	one of the best series i ever seen i like the ...
4	so involving and absorbing amazing job all act...
...	...
6900	wheres part
6901	waiting for more shots season
6902	i want season
6903	one word worst
6904	story pattern is same as season
6905 rows × 1 columns	data_clean_df.sample(6)

4. HORROR

transcript	
0	ghoul is another offering by netflix in its at...
1	ghoul or ghul is not a fascinating watch as a ...
2	it was a good horror miniseries if not the bes...
3	i came very late across this miniseries and i ...
4	ghoul is a different kind of horror that makes...
...	...
4249	please someone explain anurag kashyap story
4250	fuack this
4251	justiceforssr boycottbollywood
4252	bakavass time vesting and illogical
4253	average
4254 rows × 1 columns	data_clean_df.sample(6)

transcript	
3076	the narration way of the story from the first ...
1703	best show story is great best part is there is...
2191	its funny not scary and total time waste
1824	is this something kind of joke paly by sharukh...
2936	the cast are playing the roles very well and t...
2430	decentthe plot is gripping and there is a soli...

5. ROMANCE

transcript

0 even though i am no expert but i found the cam...
1 i stumbled upon this series randomly while sur...
2 i was prejudice towards watching the married w...
3 i am sure we will wait a long while till somet...
4 without thinking and would recommend to every ...
...

6040 the only good thing about series is rohit
6041 it was wasted of time
6042 hated it total waste of time
6043 godddsuch a waste of time
6044 worst show everreally disappointed

6045 rows × 1 columns

data_clean_df.sample(6)

transcript

4848 just give this show a chance you wont be disap...
4008 omg season of bebakee is incredible
5465 stilted actingdialogues and dialogue delivery ...
882 its just amazing
3419 super
1326 while i am awed by the performances veteran...

6. THRILLER

transcript

0 reviews of sacred games thriller
1 nan
2 nan
3 had read the sprawling world building novel by...
4 simply brilliant content to binge watch stella...
...

8443 good onebut was it necessary to use the dirty ...
8444 wating for seasion better then
8445 worth watching but they could have done better
8446 series is good but too many loop holes
8447 very intresting and full of twist show

8448 rows × 1 columns

data_clean_df.sample(6)

transcript

7188 i cant explain i was amazing web series just g...
2481 best web series on amazon prime must watch thi...
1690 a well made serial giving fascinating glimpse ...
2454 this should be banned for its filthy dialogues...
7835 absolutely stunningso many characters so many ...
7503 not a very promising show to watch in these times

4.3 Sentiment Analysis

1. ACTION

	transcript	polarity	subjectivity	Analysis
0	nan	0.000	0.000	neutral
1	reviews of mai hero boll raha hu	0.000	0.000	neutral
2	nan	0.000	0.000	neutral
3	parth just nailed the role of nawabas he recei...	0.400	0.450	positive
4	i love the series it is amazing parth acting i...	0.525	0.625	positive
...
9896	not seen yet	0.000	0.000	neutral
9897	entertainment is just for entertainment	0.000	0.000	neutral
9898	average	-0.150	0.400	negative
9899	time pass series	0.000	0.000	neutral
9900	please avoid this one	0.000	0.000	neutral

9901 rows × 4 columns

	transcript	polarity	subjectivity	Analysis
3971	loved it recommended waiting for second season...	0.433333	0.566667	positive
8593	i cant believe that something which begin so b...	-0.048333	0.668333	negative
7382	zee never disappoints its audience what an ins...	0.500000	1.000000	positive
7777	not only for fastpaced execution but also for ...	0.000000	1.000000	neutral
8601	waste of timevery bad script all stories incom...	-0.450000	0.333333	negative
2993	i can clearly see it must be funded by some te...	0.100000	0.383333	positive
5728	one of the best web series	1.000000	0.300000	positive
6633	this series is too much fantastic series ever ...	0.466667	0.633333	positive
5699	series review	0.000000	0.000000	neutral
9577	awesome emraan hashmi is to good	0.850000	0.800000	positive

So,The following is our "Sentiment Analysis" for the Top 10 Indian Web Series(Action Genre) :

No of Negative Reviews from our Total DataSet(around 10k) -> 831

No of Positive Reviews from our Total DataSet(around 10k) -> 8163

No of Neutral Reviews from our Total DataSet(around 10k) -> 907

Percentage of Negative Reviews -> 8.393091606908392 %

Percentage of Positive Reviews -> 82.44621755378245 %

Percentage of Neutral Reviews -> 9.16069083930916 %

2. COMEDY

	transcript	polarity	subjectivity	Analysis
0	nostalgic	-0.500000	1.000000	negative
1	tvf has done it again brilliant series	0.900000	1.000000	positive
2	please dont stop continue journey with season ...	0.000000	0.000000	neutral
3	story which will seriously tickle its audience...	0.083333	0.777778	positive
4	must watch web series	0.000000	0.000000	neutral
...
8006	watch for light hearted rib tickling humor	0.400000	0.700000	positive
8007	time wasting series	0.000000	0.000000	neutral
8008	worstwaste of time	0.000000	0.000000	neutral
8009	forced to act crazyy	-0.300000	0.200000	negative
8010	watch sumit sambhal lega instead	0.000000	0.000000	neutral

8011 rows × 4 columns

	transcript	polarity	subjectivity	Analysis
5699	congratulations tvf for making such an awesome...	0.500000	0.750000	positive
5136	best series for youngster and college student...	0.833333	0.366667	positive
3461	siraa	0.000000	0.000000	neutral
2117	brings on waves of nostalgia	0.000000	0.000000	neutral
5180	have a great time watching this and had a fun ...	0.650000	0.437500	positive
5641	its simplicity is so alluring you will connect...	0.250000	0.200000	positive
7763	a very good entertaining smart and funny india...	0.468571	0.780714	positive
5397	nice light comedy and awesome surroundings of ...	0.518750	0.700000	positive
418	beautiful creation by tvf and the direction te...	0.625000	1.000000	positive
7766	very funny look forward to it each week indian...	0.275000	0.866667	positive

So,The following is our "Sentiment Analysis" for the Top 10 Indian Web Series(Comedy Genre) :

No of Negative Reviews from our Total DataSet(around 10k) -> 419

No of Positive Reviews from our Total DataSet(around 10k) -> 6364

No of Neutral Reviews from our Total DataSet(around 10k) -> 1228

Percentage of Negative Reviews -> 5.230308326051679 %

Percentage of Positive Reviews -> 79.44076894270378 %

Percentage of Neutral Reviews -> 15.328922731244537 %

3. DRAMA

	transcript	polarity	subjectivity	Analysis
0	reviews of aspirants	0.000000	0.000000	neutral
1	this series is really inspiring good message I...	0.508333	0.629167	positive
2	just amazing this series is just real it could...	0.400000	0.575000	positive
3	one of the best series i ever seen i like the ...	0.475000	0.525000	positive
4	so involving and absorbing amazing job all act...	0.185714	0.502381	positive
...
6900	wheres part	0.000000	0.000000	neutral
6901	waiting for more shots season	0.500000	0.500000	positive
6902	i want season	0.000000	0.000000	neutral
6903	one word worst	-1.000000	1.000000	negative
6904	story pattern is same as season	0.000000	0.125000	neutral

6905 rows x 4 columns

data_clean_df.sample(10)

	transcript	polarity	subjectivity	Analysis
2410	one of the best series based on common struggl...	0.350000	0.400000	positive
4698	on which channel did this come i really love this	0.500000	0.600000	positive
840	what a gem	0.000000	0.000000	neutral
3581	heart touching storymohit raina dia mirza live...	0.318182	0.500000	positive
5729	loved the show masaba definitely has the neena...	0.325000	0.550000	positive
5421	amazing and enjoyable	0.550000	0.750000	positive
6288	its a much watch series this show is the true ...	0.183333	0.283333	positive
2247	best show ever this boost your confidence on h...	1.000000	0.300000	positive
5998	its nice how this series builds upon you go gi...	0.228348	0.642262	positive
3430	the story has the innocence of old child purit...	0.266071	0.451071	positive

So,The following is our "Sentiment Analysis" for the Top 10 Indian Web Series(Drama Genre) :

No of Negative Reviews from our Total DataSet(around 10k) -> 530

No of Positive Reviews from our Total DataSet(around 10k) -> 5329

No of Neutral Reviews from our Total DataSet(around 10k) -> 1046

Percentage of Negative Reviews -> 7.675597393193338 %

Percentage of Positive Reviews -> 77.17595944967415 %

Percentage of Neutral Reviews -> 15.148443157132514 %

4. HORROR

	transcript	polarity	subjectivity	Analysis		
0	ghoul is another offering by netflix in its at...	-0.026488	0.569643	negative	data_clean_df.sample(10)	
1	ghoul or ghul is not a fascinating watch as a ...	0.119859	0.559253	positive		
2	it was a good horror miniseries if not the bes...	0.338889	0.577778	positive		
3	i came very late across this miniseries and i ...	-0.086250	0.382917	negative		
4	ghoul is a different kind of horror that makes...	-0.096825	0.600317	negative		
...		
4249	please someone explain anurag kashyap story	0.000000	0.000000	neutral		
4250	fuack this	0.000000	0.000000	neutral		
4251	justiceforssr boycottbollywood	0.000000	0.000000	neutral		
4252	bakavass time vesting and illogical	0.000000	0.000000	neutral		
4253	average	-0.150000	0.400000	negative		
974 rows × 4 columns						

4254 rows × 4 columns

So,The following is our "Sentiment Analysis" for the Top 10 Indian Web Series(Horror Genre) :

No of Negative Reviews from our Total DataSet(around 5k) -> 1133

No of Positive Reviews from our Total DataSet(around 5k) -> 2837

No of Neutral Reviews from our Total DataSet(around 5k) -> 284

Percentage of Negative Reviews -> 26.633756464503993 %

Percentage of Positive Reviews -> 66.69017395392571 %

Percentage of Neutral Reviews -> 6.676069581570286 %

5. ROMANCE

	transcript	polarity	subjectivity	Analysis		
0	even though i am no expert but i found the cam...	0.466667	0.725000	positive	data_clean_df.sample(10)	
1	i stumbled upon this series randomly while sur...	0.053571	0.619048	positive		
2	i was prejudice towards watching the married w...	-0.025000	0.525000	negative		
3	i am sure we will wait a long while till somet...	0.487500	0.722222	positive		
4	without thinking and would recommend to every ...	0.400000	0.500000	positive		
...		
6040	the only good thing about series is rohit	0.350000	0.800000	positive		
6041	it was wasted of time	-0.200000	0.000000	negative		
6042	hated it total waste of time	-0.366667	0.483333	negative		
6043	goddsuch a waste of time	-0.200000	0.000000	negative		
6044	worst show everreally disappointed	-0.875000	0.875000	negative		
6045 rows × 4 columns						

So,The following is our "Sentiment Analysis" for the Top 10 Indian Web Series(Romance Genre) :

No of Negative Reviews from our Total DataSet(around 10k) -> 396

No of Positive Reviews from our Total DataSet(around 10k) -> 5252

No of Neutral Reviews from our Total DataSet(around 10k) -> 397

Percentage of Negative Reviews -> 6.550868486352357 %

Percentage of Positive Reviews -> 86.88172043010752 %

Percentage of Neutral Reviews -> 6.567411083540116 %

6. THRILLER

	transcript	polarity	subjectivity	Analysis	
0	reviews of sacred games thriller	0.000000	0.000000	neutral	
1		nan	0.000000	0.000000	neutral
2		nan	0.000000	0.000000	neutral
3	had read the sprawling world building novel by...	0.220000	0.360000	positive	
4	simply brilliant content to binge watch stell...	0.356795	0.524930	positive	
...	
8443	good onebut was it necessary to use the dirty ...	0.033333	0.800000	positive	
8444	wating for seasion better then	0.500000	0.500000	positive	
8445	worth watching but they could have done better	0.400000	0.300000	positive	
8446	series is good but too many loop holes	0.600000	0.550000	positive	
8447	very intresting and full of twist show	0.275000	0.425000	positive	
8448 rows × 4 columns					
data_clean_df.sample(10)					
	transcript	polarity	subjectivity	Analysis	
48	congratulations to netflix and the team of sac...	0.103889	0.515556	positive	
4793	interesting but a low budget series	0.250000	0.400000	positive	
2850	realistic to an extent but the ease with polic...	-0.011111	0.177778	negative	
3676	awsome but only for should watch this	0.000000	1.000000	neutral	
4651	one of the best web series full of thriller	0.675000	0.425000	positive	
5938	this show is overwhelming the unique cast make...	0.537771	0.568615	positive	
457	the acting is simply awesome you should watch ...	0.000000	0.500000	neutral	
6582	unbelievably real cant remember any better ind...	0.350000	0.400000	positive	
187	really appreciate the series i used to watch o...	0.133333	0.316667	positive	
1483	very sensitive and thriller	0.130000	1.000000	positive	

So,The following is our "Sentiment Analysis" for the Top 10 Indian Web Series(Thriller Genre) :

No of Negative Reviews from our Total DataSet(around 10k) -> 1014

No of Positive Reviews from our Total DataSet(around 10k) -> 6594

No of Neutral Reviews from our Total DataSet(around 10k) -> 840

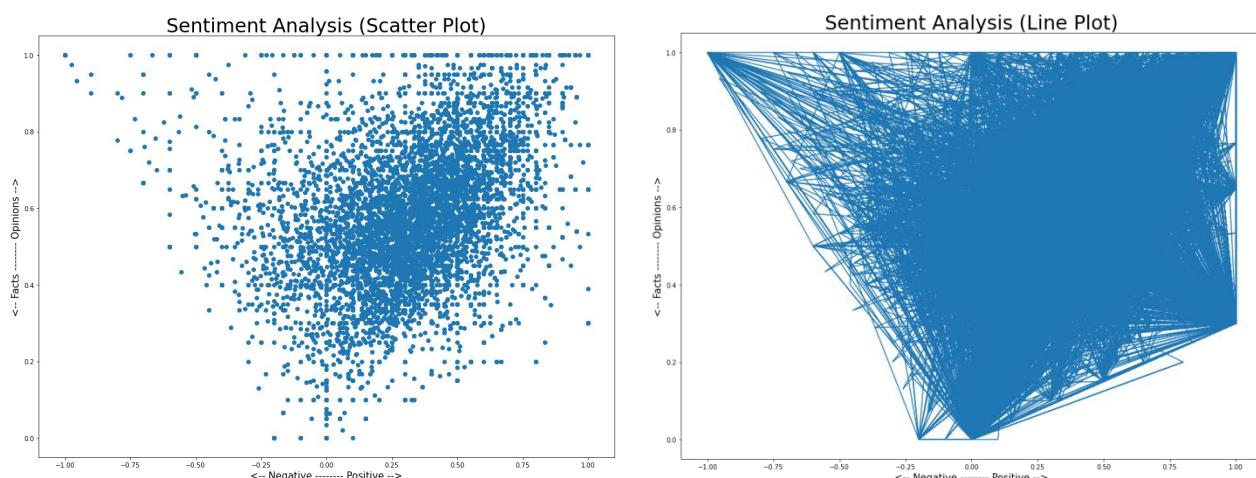
Percentage of Negative Reviews -> 12.002840909090908 %

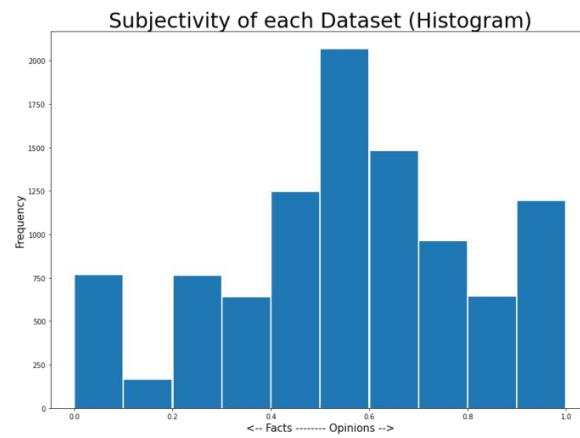
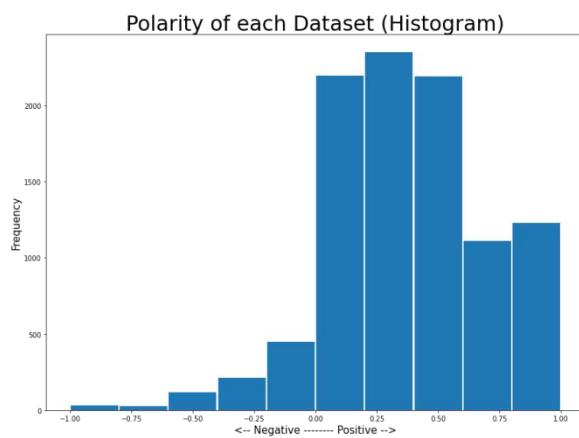
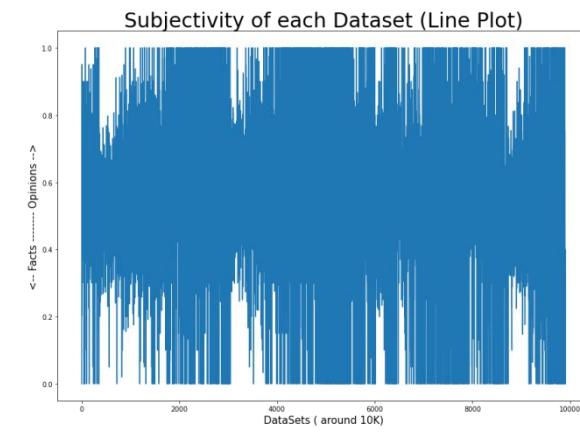
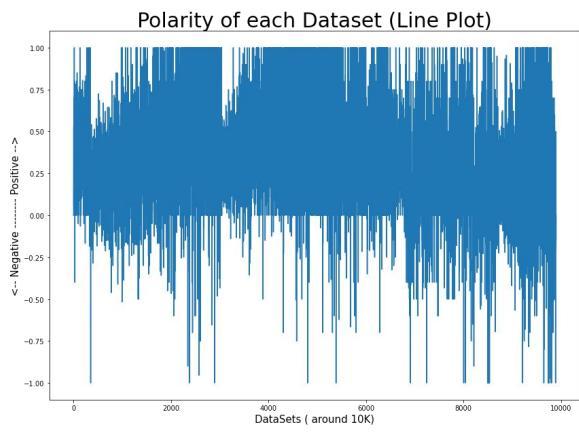
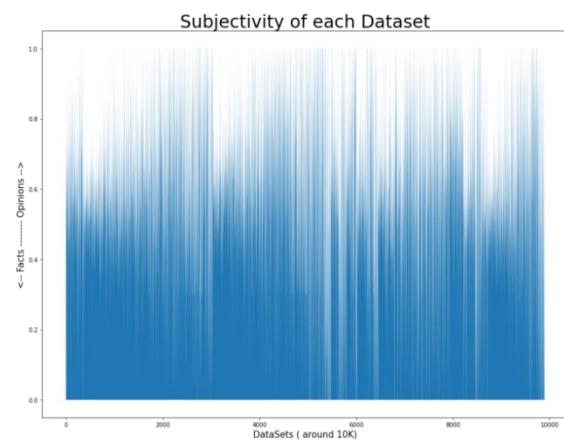
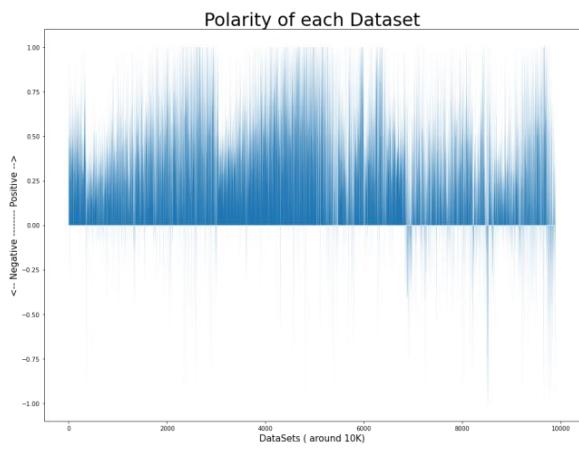
Percentage of Positive Reviews -> 78.05397727272727 %

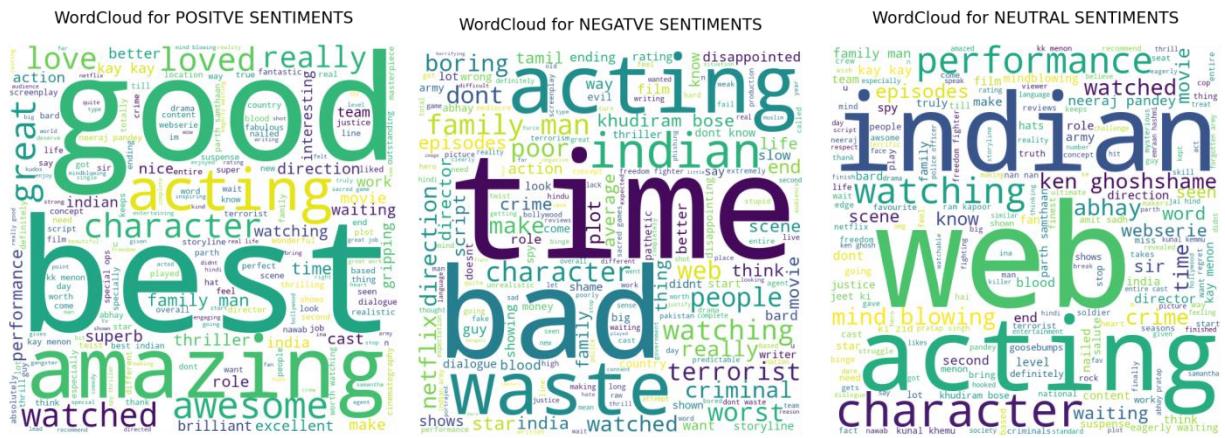
Percentage of Neutral Reviews -> 9.943181818181818 %

4.4 Data Visualizations

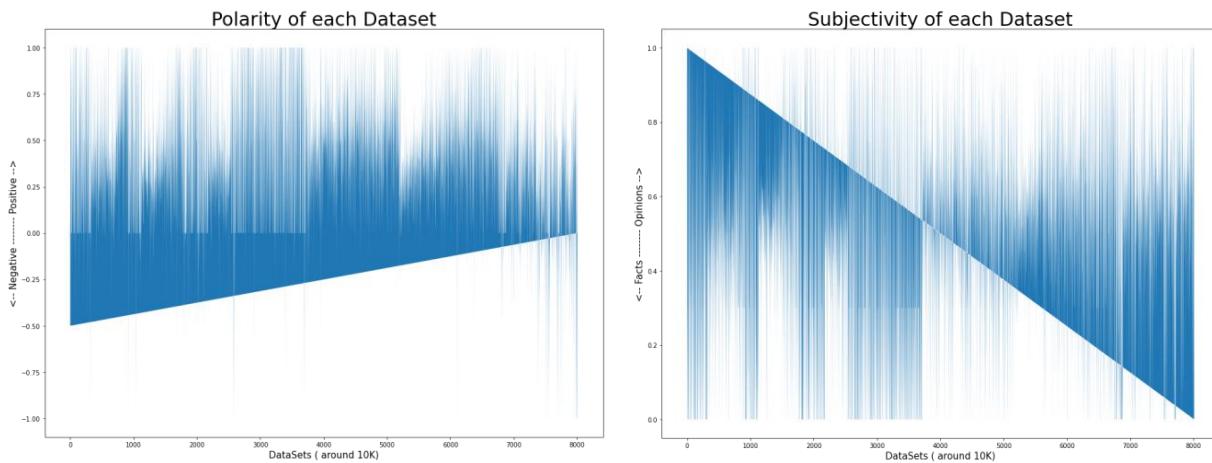
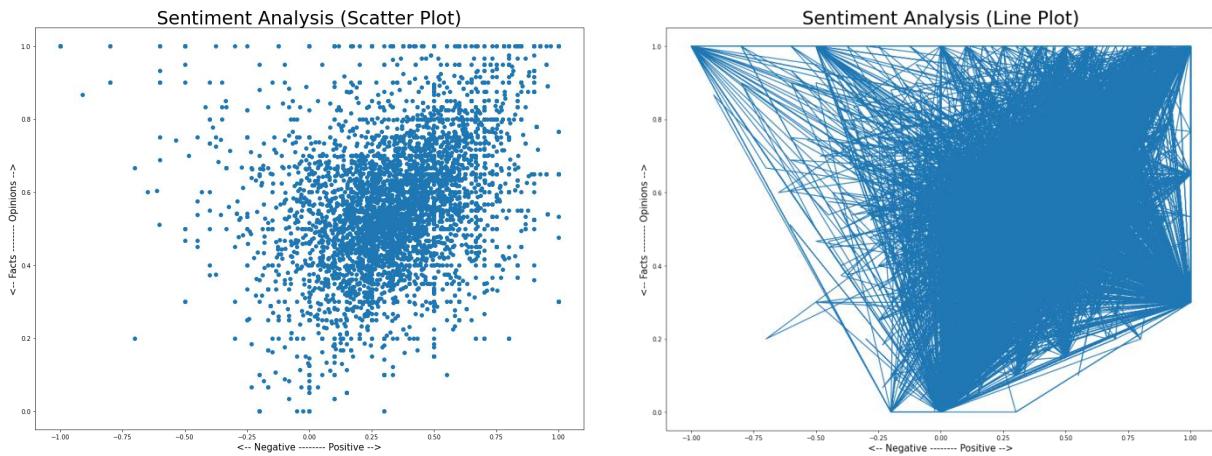
1. ACTION

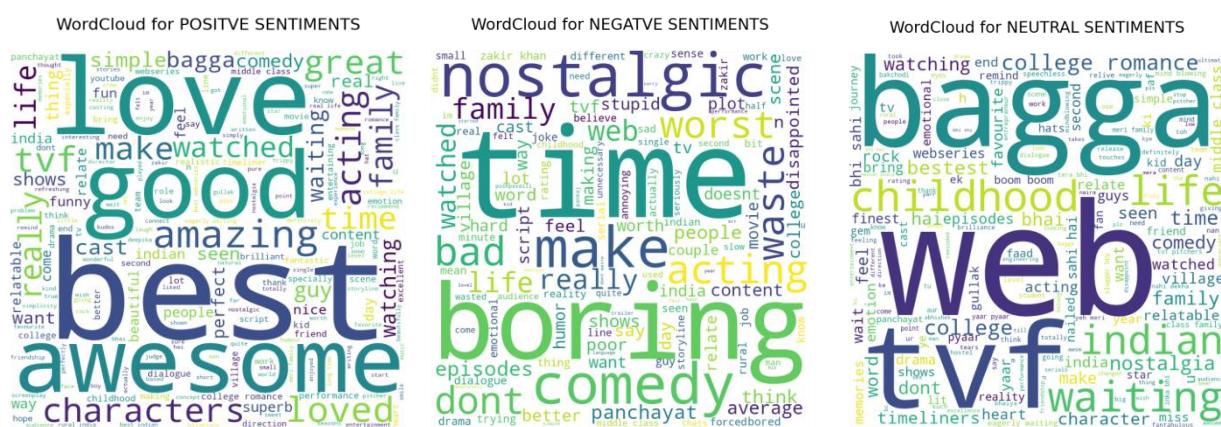
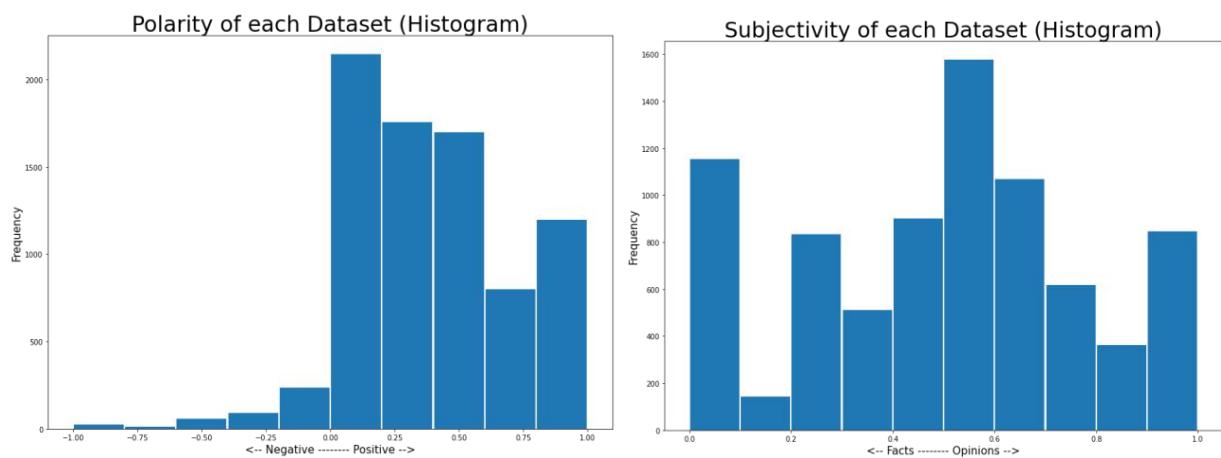
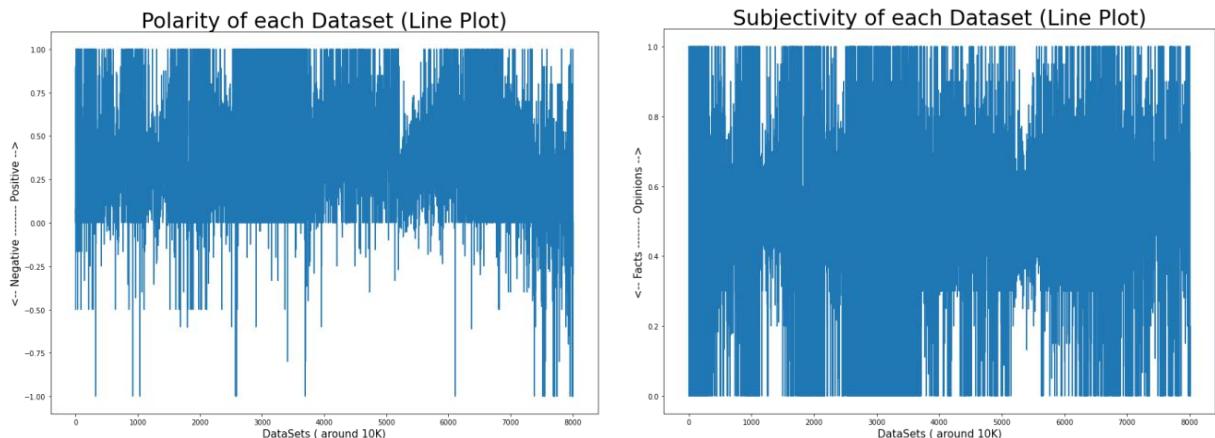




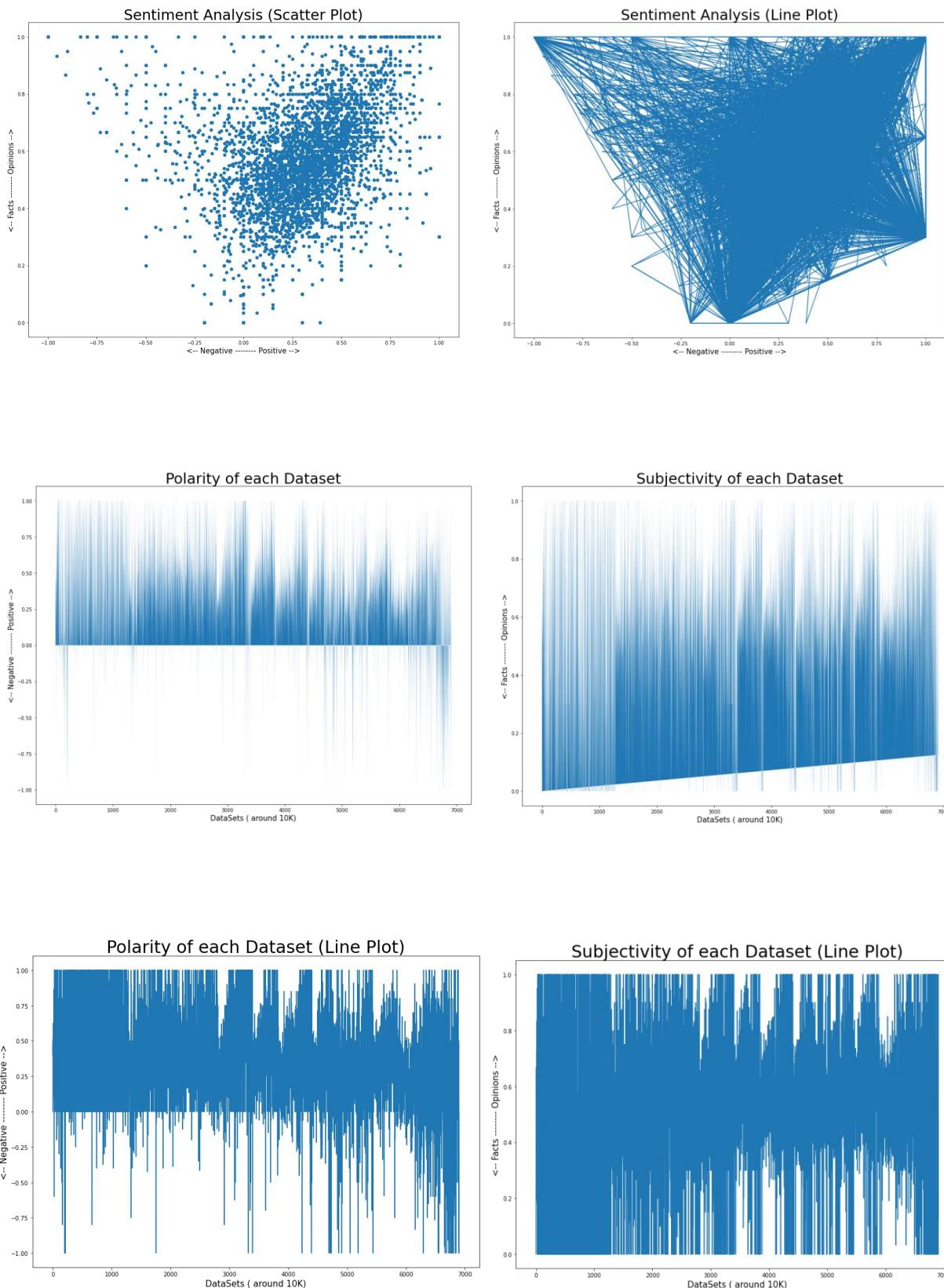


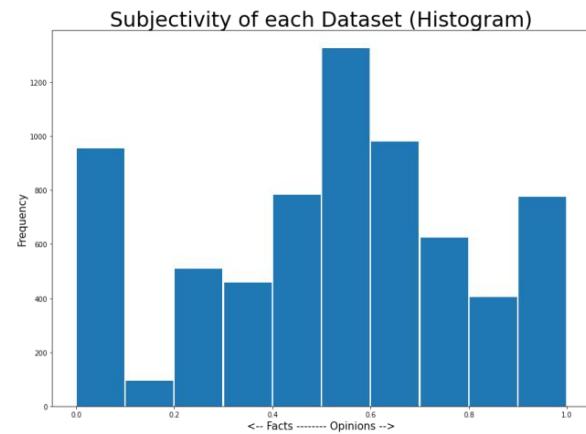
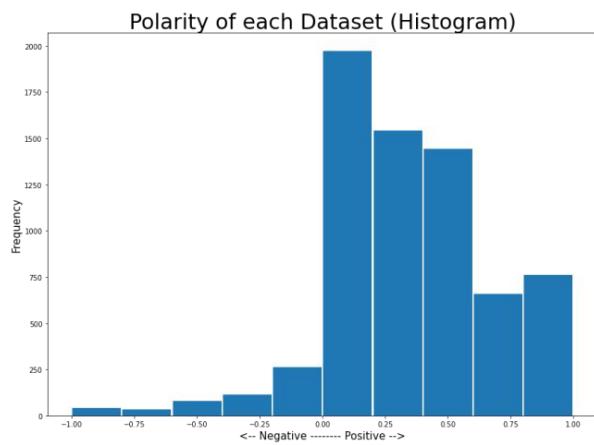
2. COMEDY





3. DRAMA





WordCloud for POSITIVE SENTIMENTS



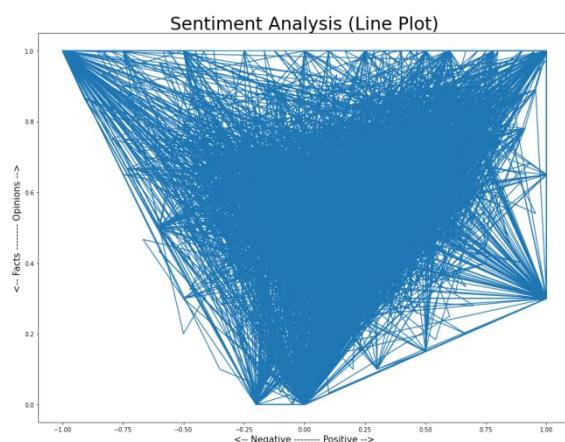
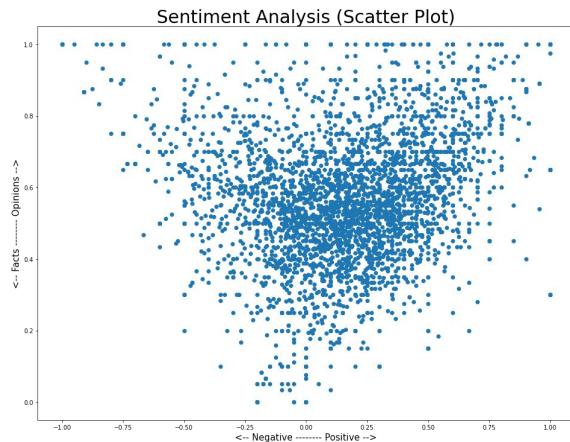
WordCloud for NEGATIVE SENTIMENTS

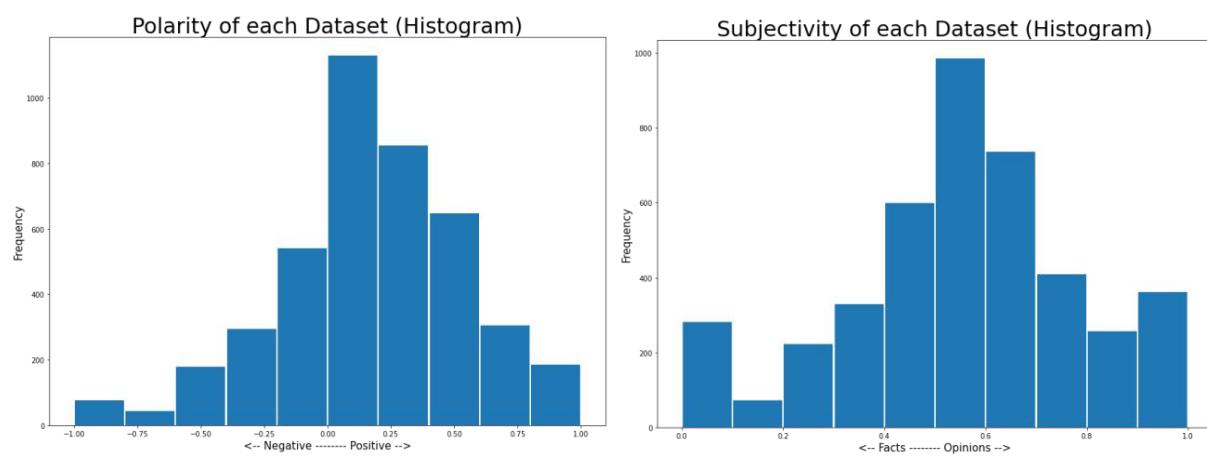
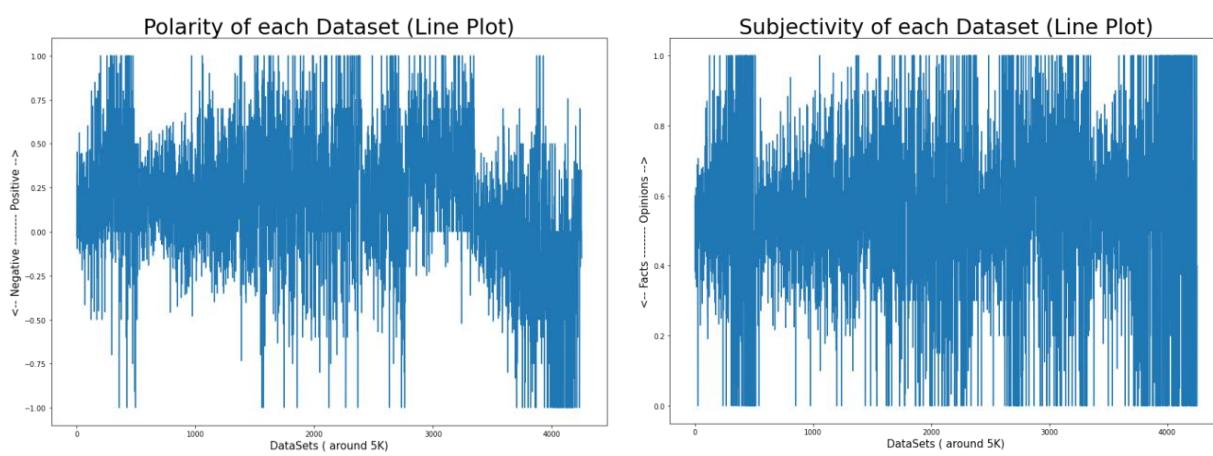
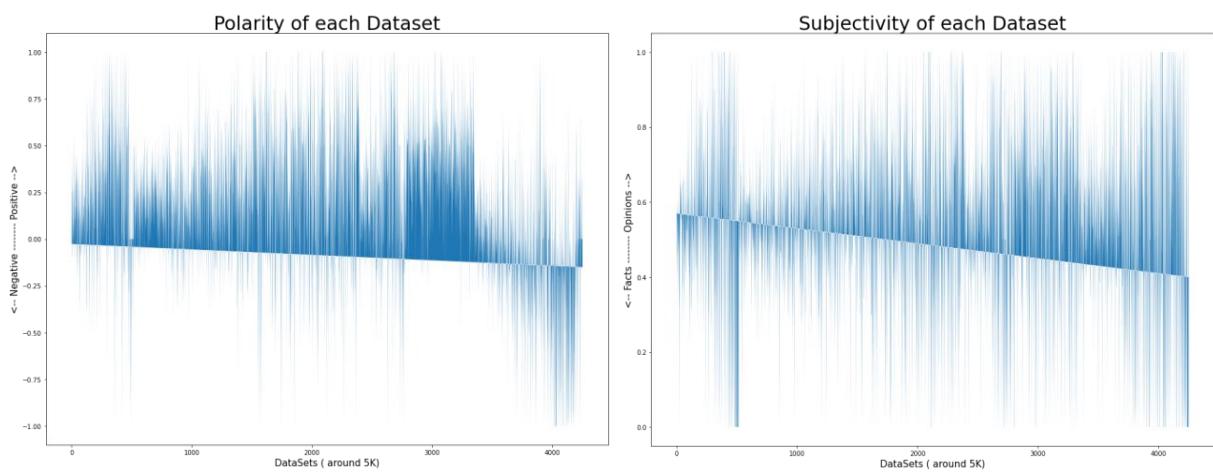


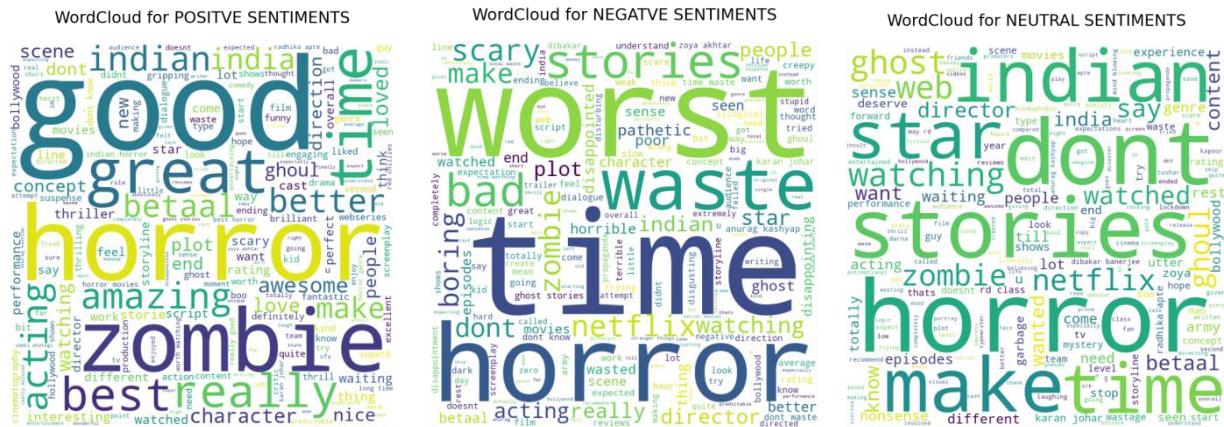
WordCloud for NEUTRAL SENTIMENTS



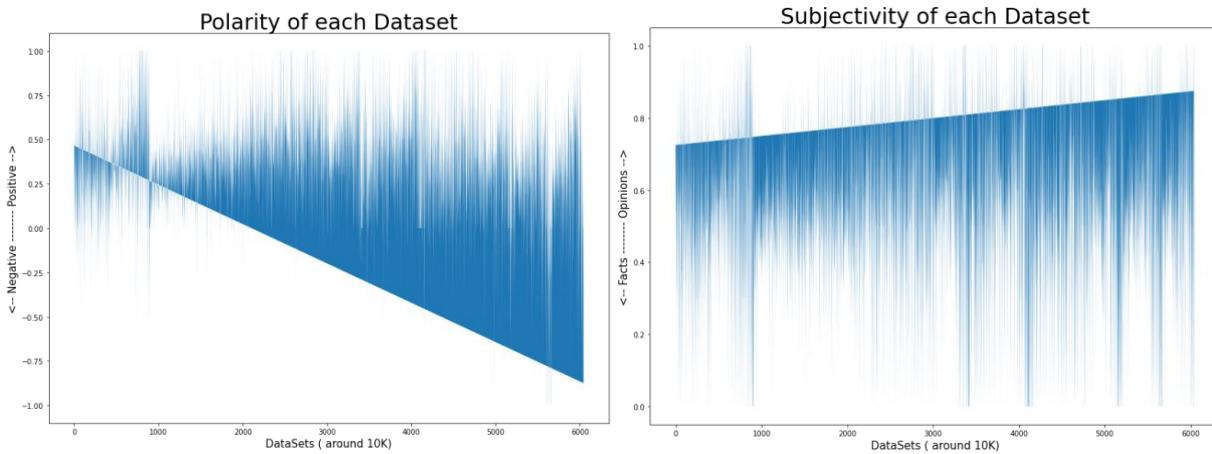
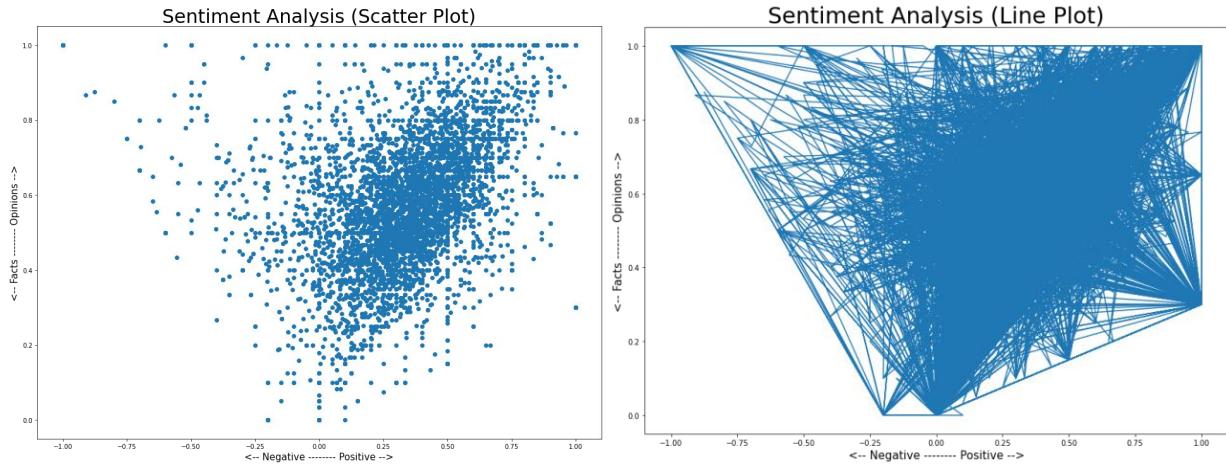
4. HORROR

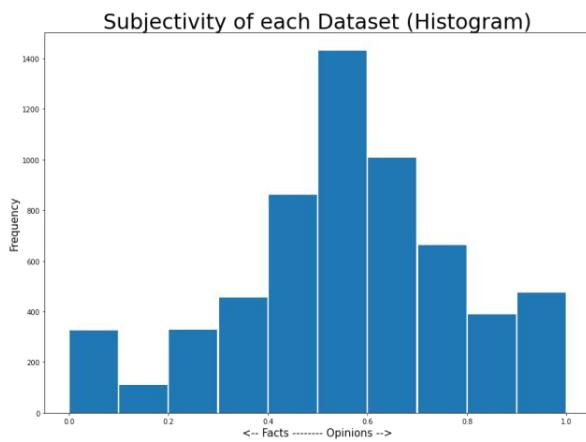
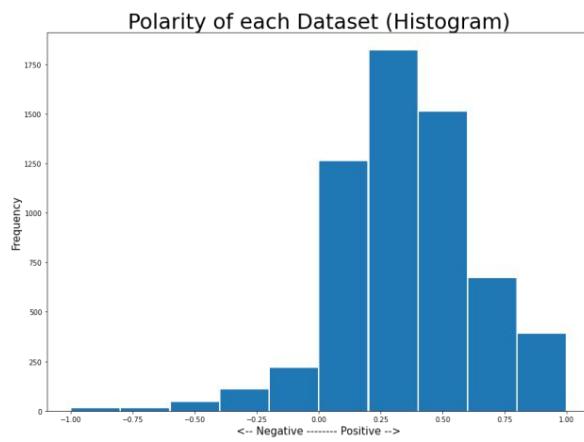
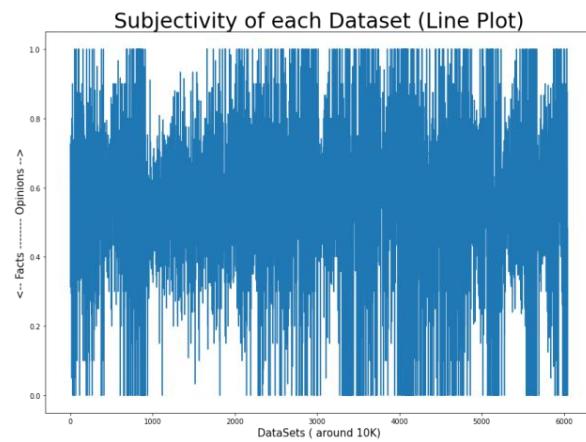
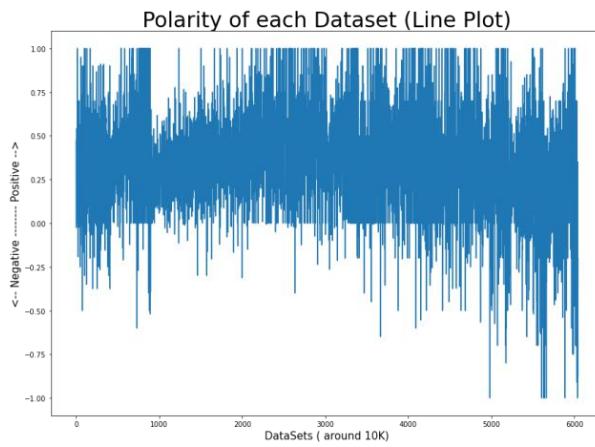






5. ROMANCE





WordCloud for POSITIVE SENTIMENTS



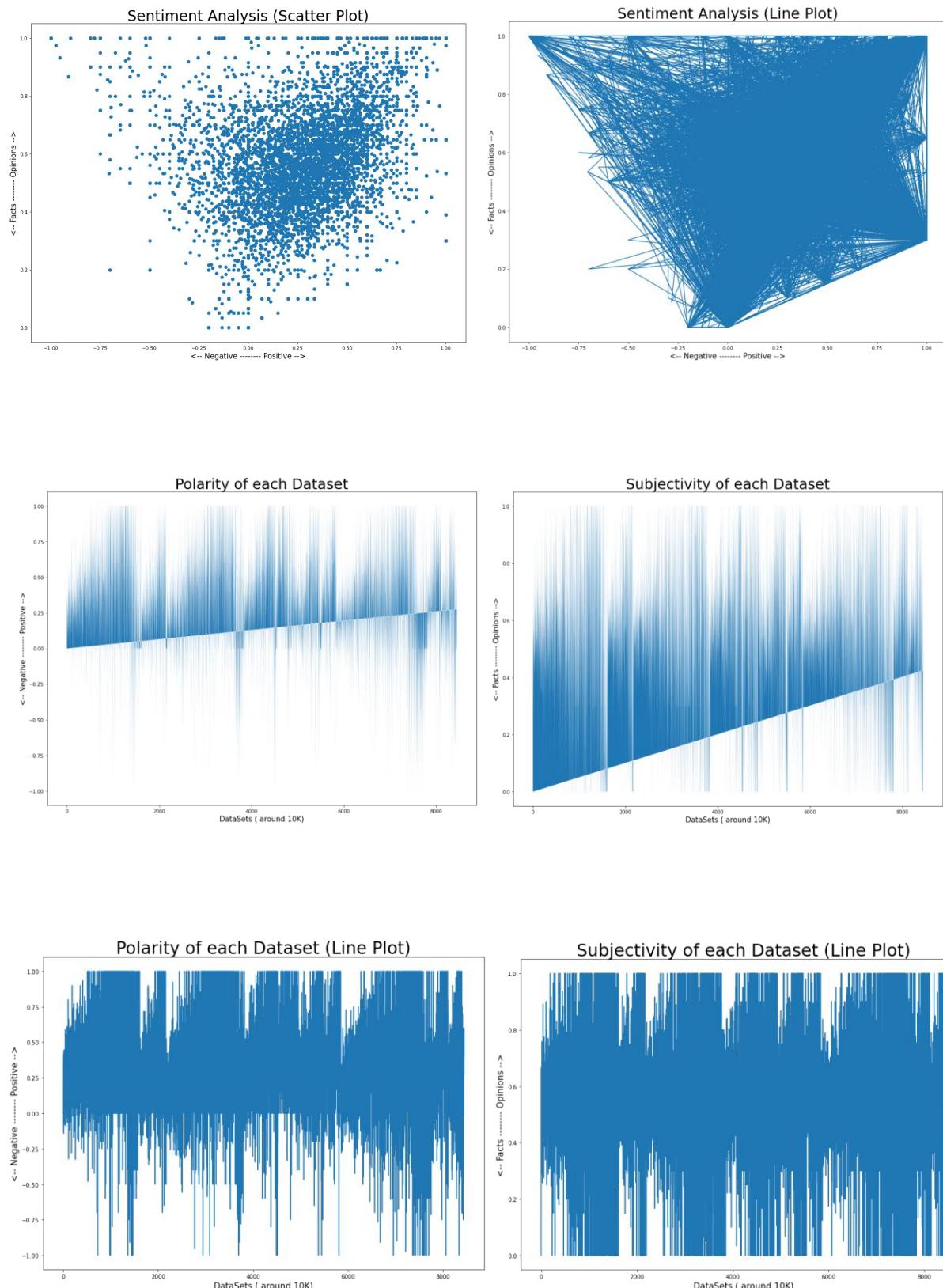
WordCloud for NEGATIVE SENTIMENTS

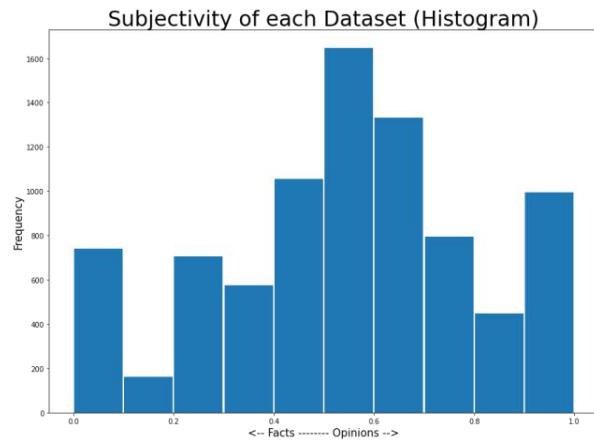
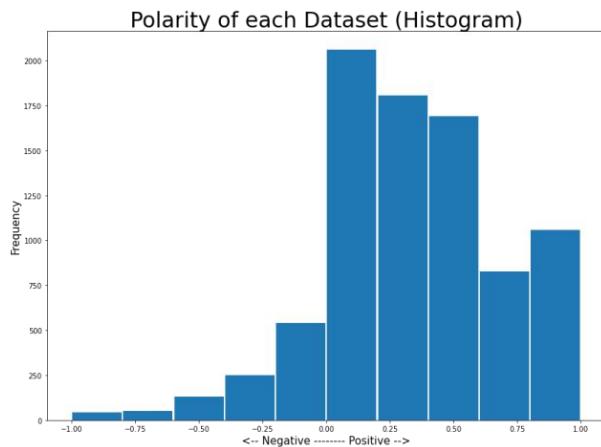


WordCloud for NEUTRAL SENTIMENTS



6. THRILLER





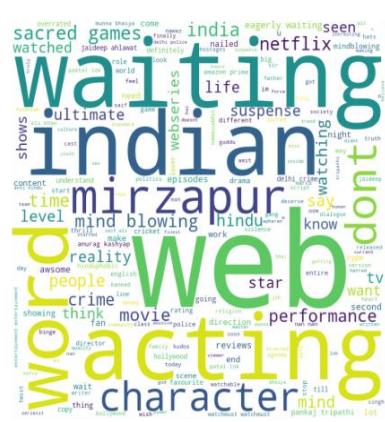
WordCloud for POSITIVE SENTIMENTS



WordCloud for NEGATIVE SENTIMENTS



WordCloud for NEUTRAL SENTIMENTS



5. Conclusion and Future Scope

Interest in sentiment analysis is growing tremendously. Although the field is emerging, but it is still fairly new and the researcher can be confronted with challenges. One possible challenge relates to the postings on the internet can be difficult to analyze given that they often reflect shorter versions of phrases.

The proliferation of user reviews sites offer an unprecedented opportunity to create and employ theories & technologies that search and mine for sentiments. The work presented in this research report specified a simple approach for sentiment analysis on web scraped data. The corpus-based method was used to find the semantic orientation of adjectives and the dictionary-based method to find the semantic orientation of verbs and adverbs using TextBlob and its features. This work was exploratory in nature and the prototype evaluated was a preliminary prototype. The initial results showed that it was a motivating technique.

Sentiment Findings:

So,The following was our "Sentiment Analysis" for the Top 10 Indian Web Series :

1. Action Genre:

- No of Negative Reviews from our Total DataSet(around 10k) -> 831
- No of Positive Reviews from our Total DataSet(around 10k) -> 8163
- No of Neutral Reviews from our Total DataSet(around 10k) -> 907

- Percentage of Negative Reviews -> 8.393091606908392 %

- Percentage of Positive Reviews -> 82.44621755378245 %

- Percentage of Neutral Reviews -> 9.16069083930916 %

2. Comedy Genre

- No of Negative Reviews from our Total DataSet(around 10k) -> 419

- No of Positive Reviews from our Total DataSet(around 10k) -> 6364

- No of Neutral Reviews from our Total DataSet(around 10k) -> 1228

- Percentage of Negative Reviews -> 5.230308326051679 %

- Percentage of Positive Reviews -> 79.44076894270378 %

- Percentage of Neutral Reviews -> 15.328922731244537 %

3. Drama Genre

- No of Negative Reviews from our Total DataSet(around 10k) -> 530

- No of Positive Reviews from our Total DataSet(around 10k) -> 5329

- No of Neutral Reviews from our Total DataSet(around 10k) -> 1046

- Percentage of Negative Reviews -> 7.675597393193338 %

- Percentage of Positive Reviews -> 77.17595944967415 %

- Percentage of Neutral Reviews -> 15.148443157132514 %

4. Horror Genre

- No of Negative Reviews from our Total DataSet(around 5k) -> 1133
 - No of Positive Reviews from our Total DataSet(around 5k) -> 2837
 - No of Neutral Reviews from our Total DataSet(around 5k) -> 284
-
- Percentage of Negative Reviews -> 26.633756464503993 %
 - Percentage of Positive Reviews -> 66.69017395392571 %
 - Percentage of Neutral Reviews -> 6.676069581570286 %

5. Romance Genre

- No of Negative Reviews from our Total DataSet(around 10k) -> 396
 - No of Positive Reviews from our Total DataSet(around 10k) -> 5252
 - No of Neutral Reviews from our Total DataSet(around 10k) -> 397
-
- Percentage of Negative Reviews -> 6.550868486352357 %
 - Percentage of Positive Reviews -> 86.88172043010752 %
 - Percentage of Neutral Reviews -> 6.567411083540116 %

6. Thriller Genre

- No of Negative Reviews from our Total DataSet(around 10k) -> 1014
- No of Positive Reviews from our Total DataSet(around 10k) -> 6594

- No of Neutral Reviews from our Total DataSet(around 10k) -> 840

- Percentage of Negative Reviews -> 12.002840909090908 %

- Percentage of Positive Reviews -> 78.05397727272727 %

- Percentage of Neutral Reviews -> 9.943181818181818 %

These also confirmed our vague analysis that we did using just the word cloud sentiments.

Future Scope

The task of sentiment analysis, especially in the domain of social-media reviews, is still in the developing stage and far from complete. So we propose a couple of ideas which we feel are worth exploring in the future and may result in further improved performance.

- We had currently only focused on unigrams, so the effect of bigrams and trigrams may be explored using TextBlob.ngrams() method.
- Spelling Corrections may be done using TextBlob's correct() method.
- Words Inflection/Stemming and Lemmatization may be done using TextBlob.words and TextBlob.lemmatize method.
- We can start working on multi languages like Hindi, Spanish, and Arabic to provide sentiment analysis to more local.
- We can use classifiers, so that we can get slightly better results.(eg. from textblob.classifiers import NaiveBayesClassifier)

6. Bibliography

- Unstructured Data Analytics: How to Improve Customer Acquisition, Customer Retention, and Fraud Detection and Prevention-Jean Paul Isson,Chapter1, 2 and 3.
- Automate the Boring Stuff with Python Programming-AI Sweigart
- Naive Bayes Classifiers Connectionist and Statistical Language Processing -Frank Keller
- Bayesian Reasoning and Machine Learning-David Barber
- S. ChandraKala1 and C. Sindhu2, “OPINION MINING AND SENTIMENT CLASSIFICATION: A SURVEY,”.Vol .3(1),Oct 2012,420-427
- Manju Venugopalan and Deepa Gupta ,Exploring Sentiment Analysis on Twitter Data, IEEE 2015
- Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews.|| In Proceedings of the Twelfth