**Natural Language Processing (NLP)**

**Business** **Problem**: Extract reviews of any product from ecommerce website like snap deal and Amazon.

**Data:**

Data source: Amazon

# Item : Samsung 6.2 kg Fully-Automatic Top load Washing Machine (WA62M4100HY/TL, Imperial Silver)

**Link :** <https://www.amazon.in/Samsung-Fully-Automatic-WA62M4100HY-TL-Imperial/product-reviews/B0747XV38N/ref=cm_cr_othr_d_paging_btm_1?showViewpoints=1&pageNumber>=

Data source: Snap deal

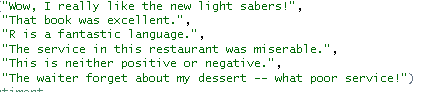
**Item :** E-Retailer Single PVC Classic Red Square Design Top Load 5KG

**Link :** <https://www.snapdeal.com/product/eretailer-single-pvc-5kg-to/672058046051/reviews?page>=

Model is build and tested and attachment contains the review table data for washing machine review and washing machine cover review from Amazon and Snap deal.

**Business** **Problem**: Perform sentimental analysis.

**Data:** Taking small set of data add in the program to perform Sentiment analysis.



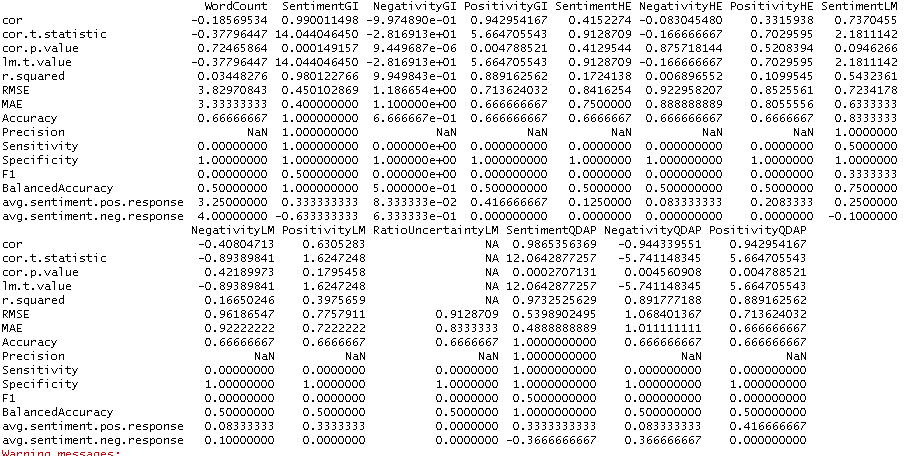
I have taken the cleaned data.

**Sentiment Output for above data:**

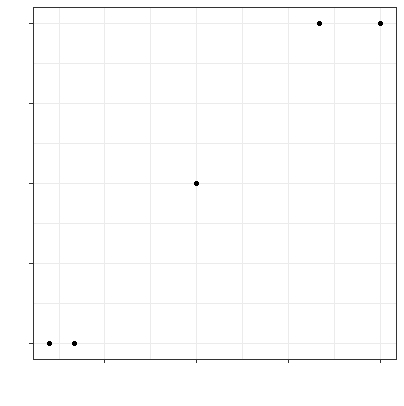
****

To processing in well shape assigning the +1 to Positive, 0 to Neutral and -1 to Negative.

**Compare Response:**



**Data Visualization:**

****

**Business** **Problem**: Extract movie reviews for any movie from IMDB and perform sentimental analysis.

**Data:** The dataset has 50000 observations and has 2 features called review and sentiment.

I’m considering review and cleaned and the dataset.

* Converted them into lower cases.
* Removed the punctuations marks.
* Removed numbers.
* Removed stop words.

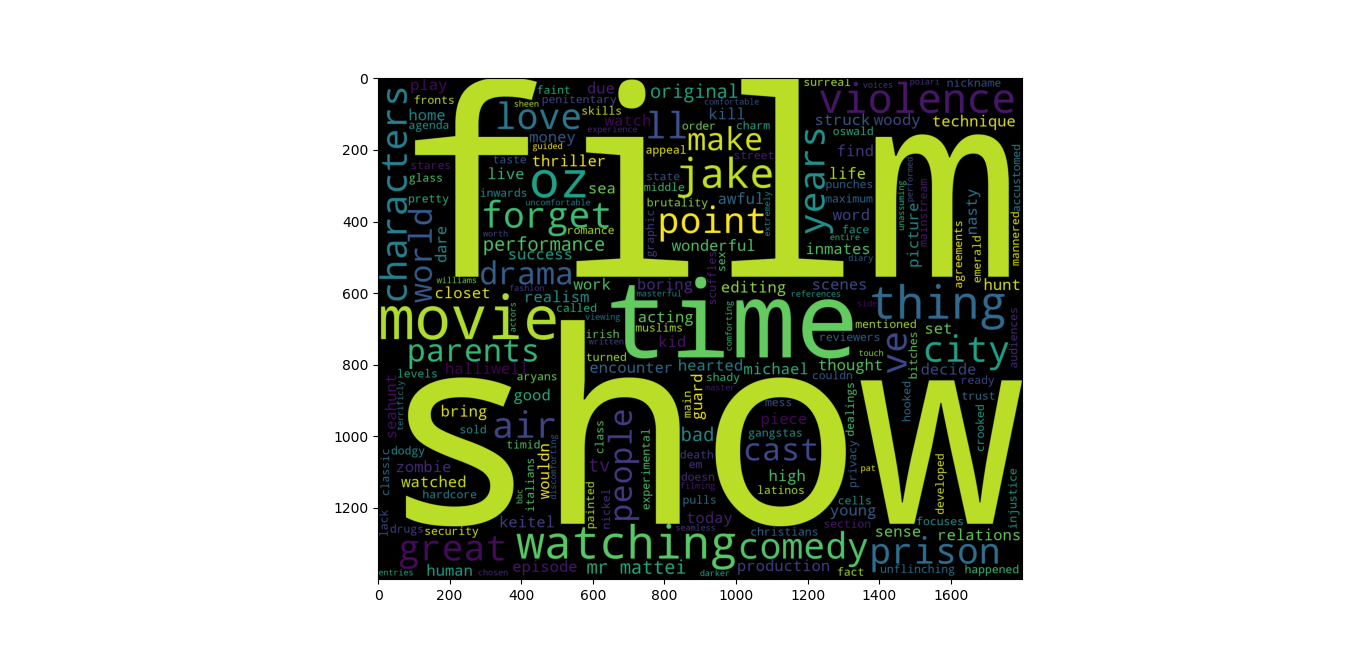
Now, data in clean format then tokenize them.

**Polarity:**

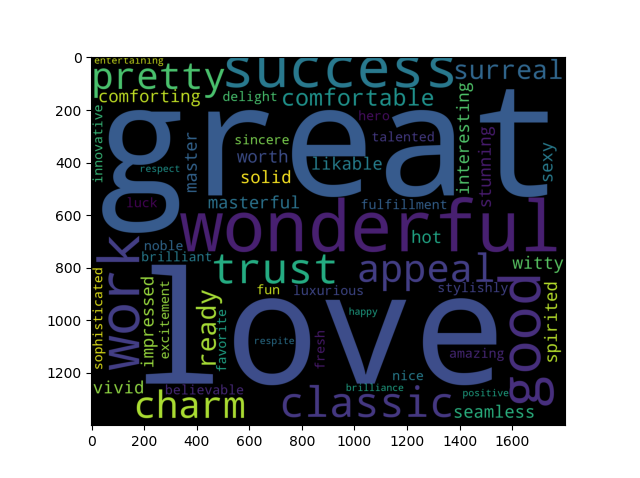
Calculated the polarity for each observation and kept them in the text and attached.

* Polarity is equal to 0 – Neutral
* Polarity is less than 0 – Negative
* Polarity is greater than 0 – Positive

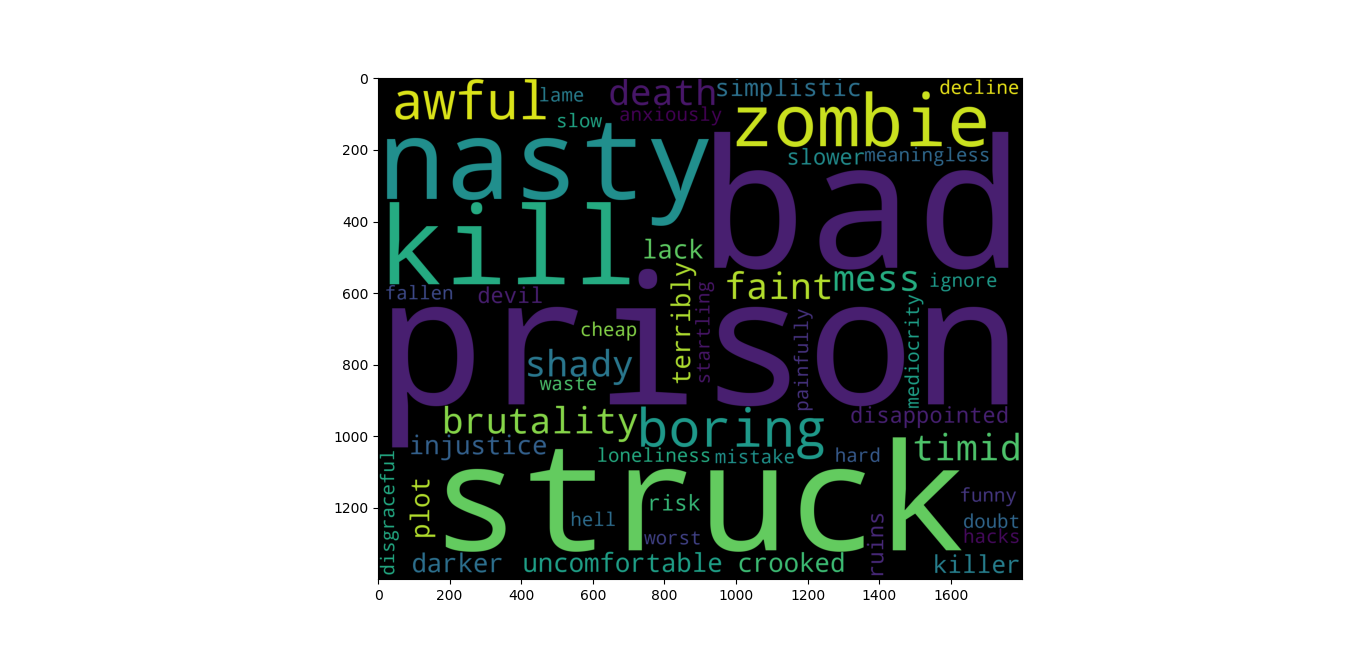
**Word Cloud:**



**Positive Word cloud:**

****

**Negative Word Cloud:**



**Business** **Problem**: Extract anything you choose from the internet and do some research on how we extract using **R Programming** and perform sentimental analysis.

**Data:** Copied the Star War Series film script.

Star Wars is a popular film franchise that takes place in a galaxy far, far away. This is a collection of script dialogue between characters for the first three movies (episodes 4-6).

SW\_EpisodeIV.txt.

SW\_EpisodeV.txt.

SW\_EpisodeVI.txt.

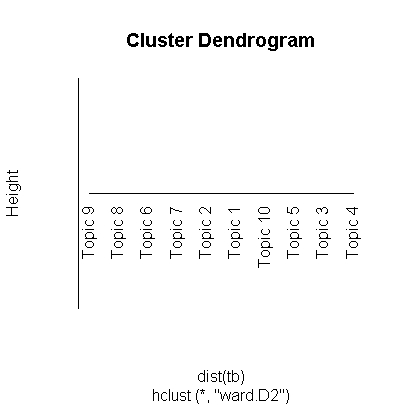
* Converted them into lower cases.
* Removed the punctuations marks.
* Removed numbers.
* Removed stop words.

Now, data in clean format then tokenize them.

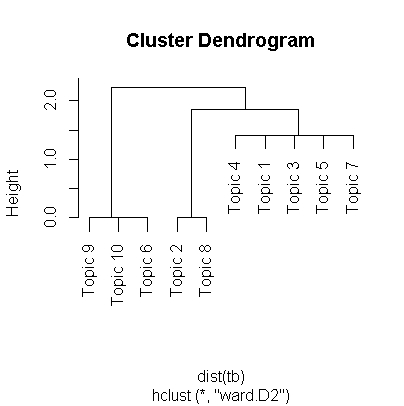
**Topic Extraction:**

I have created 10 topic for each episode, please find them in the diagram.

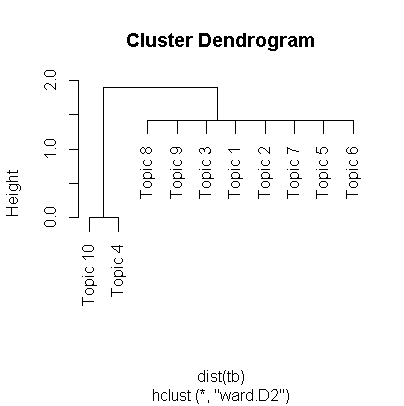
Episode 4 :



Episode 5:

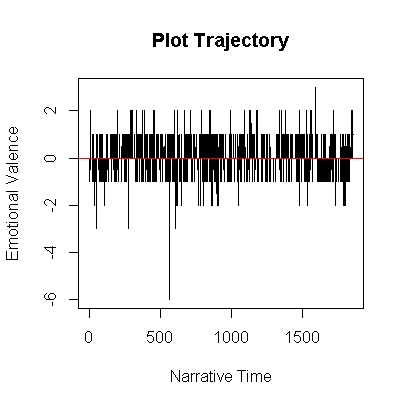


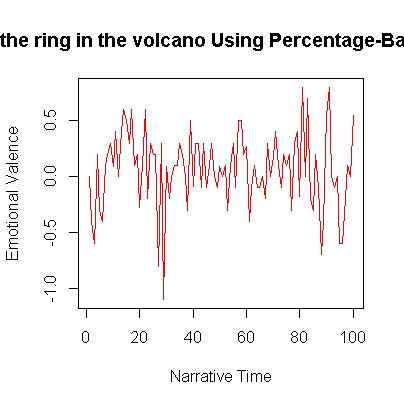
Episode 6:



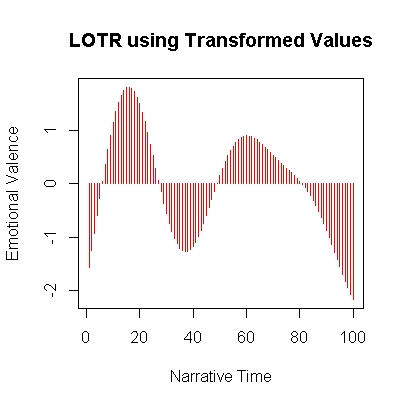
**Sentiment Analysis for episode 4:**

**Plotting for Trajectory:**

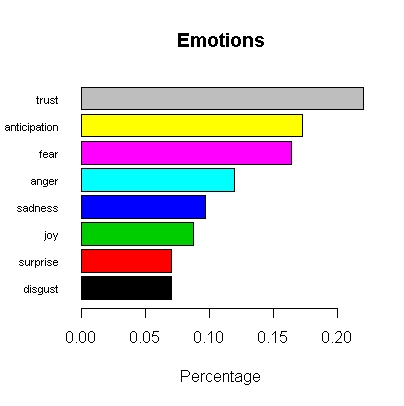
****

****

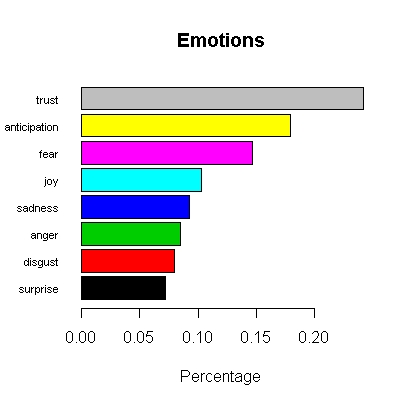
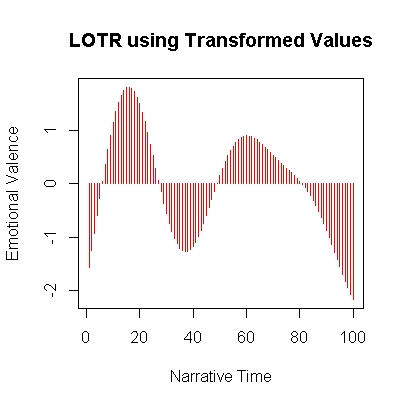
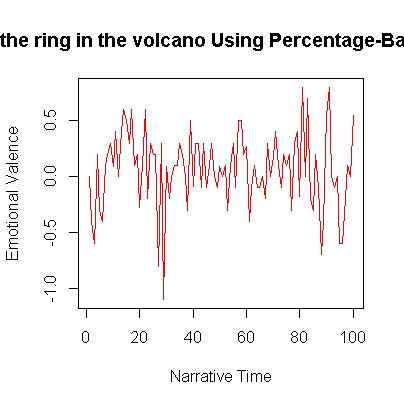
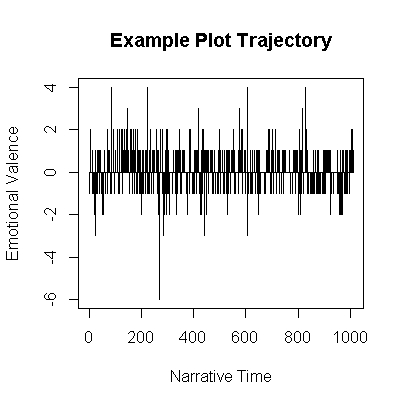
**Plot by using Bing:**

****

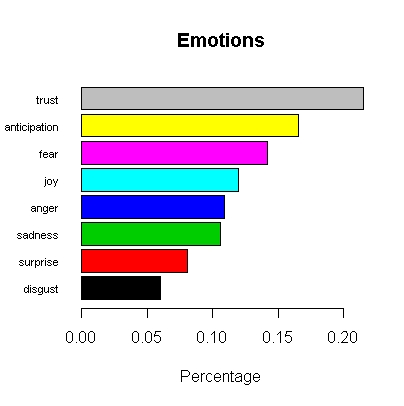
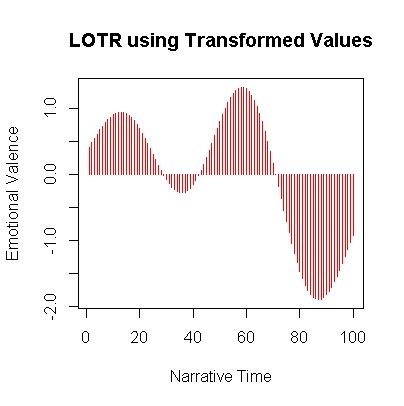
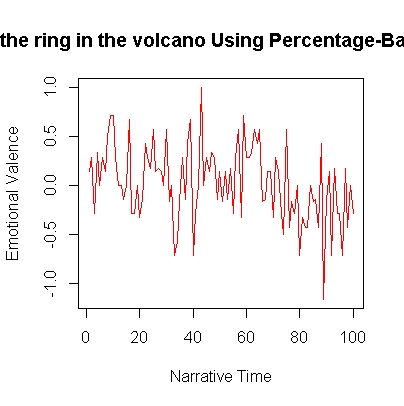
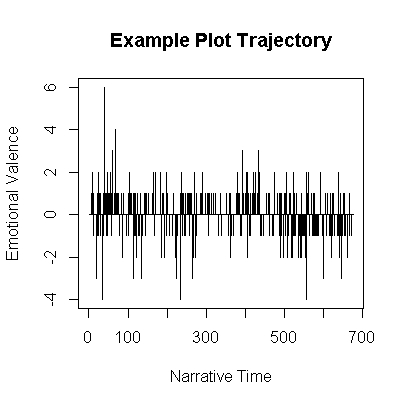
**Emotional Plot:**



**Sentiment Analysis for Episode 5:**



**Sentiment Analysis for Episode 6:**

****