

Game of Thrones - Script Analysis and More..

Introduction

What makes a great movie “Great”? Or a great television series “Great”? We believe one of the most fundamental component behind a good movie or a television series is the script. With this simple idea we started exploring for commonalities between scripts of top rated movies and television shows.

As part of this project we explore the sentiment variability in raw scripts of top Game of Thrones episodes and top IMDB rated movies. We try to look for hidden patterns among the best of the best scripts and share our findings for aspiring screenwriters, movie buffs and Game of Thrones enthusiasts.

Key Findings

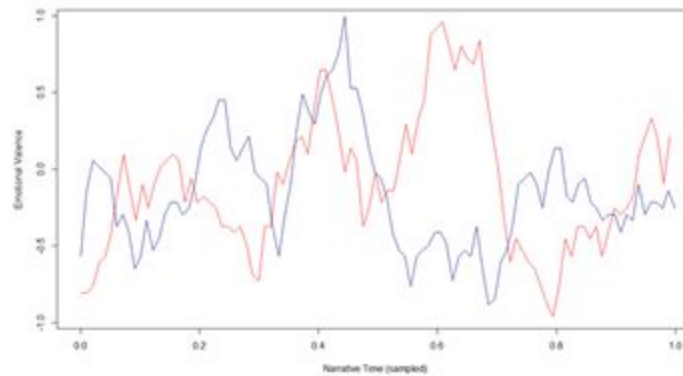
1. Popular and TV shows and movies have a template of sentiment variability in their scripts.
2. It is interesting to note that popular movies and shows have similar emotion variability across the length of the script which makes a strong argument that most audiences have a strong preference some of these templates.
3. Author/Show producers on Game of Thrones tend to kill characters as they get older or if they are affiliated to a particular house/group.

Analysis

1. Script Analysis in R
 - a. Identified and downloaded scripts for top movies and Game of Thrones episodes on IMDB.
 - b. Analysis done using the “syuzhet” package in R (“bing method”). After normalization, Fourier Transforms and Moving Average techniques for smoothing raw sentiment values were used. We used a bin size of 1% (of the total sentences in the script) to compare scripts of different length side by side.

Results

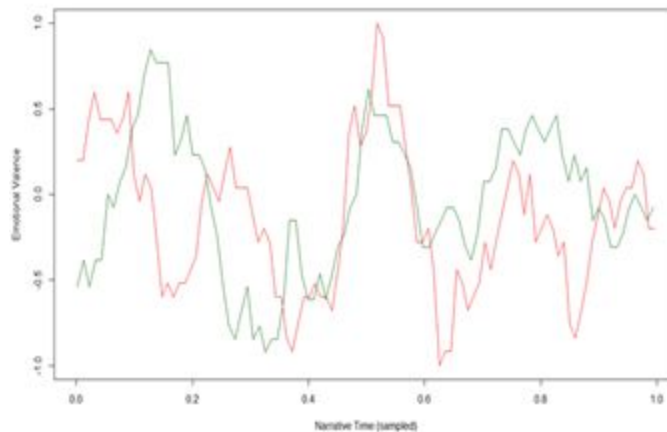
- a. Movies - Four top rated IMDB movies were analyzed; The Shawshank Redemption, The Godfather I and II, and The Dark Knight. Movie scripts exhibited similar patterns.



Blue: The Shawshank Redemption
 Red: The Godfather 2 (1974)

The above template shows how similar the emotional variability is between the two movies.

- b. Game of Thrones – Scripts of the following Game of Thrones episodes were analyzed; Hardhome, Blackwater, The Children, The Rains of Castamere.

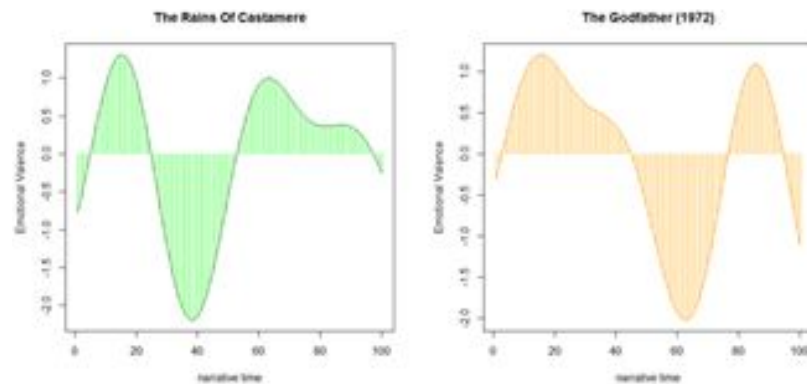


Red: Hardhome
 Green: The Rains of Castamere

There is also similarity seen between two Game of Thrones episodes. The emotional variance is almost identical for the latter 2/3rd of the adjacent visualization.

- c. Movies and Game of Thrones (“The Rains of Castamere” and “The Godfather”)

The sentiment variability of “The Rains of Castamere” and “The Godfather” are almost mirror images of each other.



Thus, we believe that a successful movie or TV series follows a significantly similar template for their script.

2. Visual Analysis in Tableau
 - a. Raw data was downloaded from [here](#).
 - b. Developed new data points from the existing data. New database can be found [here](#).
 - c. Visual analysis and representation done in [Tableau Public](#).
 - d. Character death analysis was done on two parameters - age and house affiliation.

Results

Visualization can be found [here](#).

- a. Characters associated with some houses have a higher death rate than the rest. For instance, House Bolton, House Targaryen etc. have a higher death rate as compared to other houses like Tyrell, Martell etc.
- b. Also, probability of a character dying increases as they get older. Age 51-60 is the most dangerous while age 11-20 is the safest for any character.

Data sources

1. <https://jordanscherner.wordpress.com/2014/08/06/valar-morghulis/>
2. [Game of Thrones Scripts](#)

PowerPoint Presentation

The PowerPoint presentation for this project can be found [here](#).