**Project Title:**

To determine whether there is a link between the sentiment of a TV show and the viewer ratings.

**Student:**

Mark Baber - 17076749

**Supervisor:**

Dr Paul Jarvis

**Introduction:** In this project I am going to use sentiment analysis to determine if there is a link between the sentiment of a television shows episode and its review score.

I will also carry out some investigation into web scraping data and reviews from public domain websites (depending on the scope of this project).

A literature review of the Scopus database wasn’t able to find any literature about this topic, meaning it could be an under-explored area. By researching this area, I can add to the academic literature while also developing a method that could work for practitioners in the area.

**Background:**

At this point in time there are many television shows being commissioned for first and repeat seasons. People can watch this on multiple platforms such as Netflix, Prime or normal TV and advertisers can choose where to advertise based on factors such as the popularity of the show. The popularity of the show could be determined in multiple ways, but ratings is a clear way of demonstrating popularity. Websites such as rottentomatoes.com show scores determined by critics as well as scores voted for by an audience. This data could be a way to assess a popularity of a show. Combining this with sentiment analysis of individual episodes will demonstrate if there is a link between the sentiment and popularity.

**Rationale and Justification:**

As stated above there is no easily available literature on this subject, therefore carrying out the research can add value for practitioners and academics.

**Overall Objective:**

To determine whether there is a link between the sentiment of a TV show and the viewer ratings.

**Specific Objectives:**

* To perform a literature review of sentiment analysis.
* To investigate the sentiment of a TV show, per episode/season.
* To investigate the viewer rating of a TV show from reviewer websites.
* To investigate if there is a link between both results.

**Software:**

For keeping track of my data and the script I will use:

* Git
* GitHub

To carry out the sentiment analysis and write the scripts I will use:

* R / R-Studio
* Visual Studio Code

For writing up my project I use a mix of office suites, including:

* Microsoft Office
* LibreOffice
* Google Docs & Google Sheets

**Methodology:**

**Stage 1**

* Download TV show scripts
* Break down word for word
* Analyse sentiment per episode

**Stage 2**

* Get reviews from review sites
* Analyse ratings and sentiment
* Get public domain data for analysis

**Stage 3**

* Compare Findings

**Stage 4**

* Report findings

**Data analysis:**

For the Data Analysis I will look at using packages within R-Studio to tidy and analyse the data. This will allow the use of data manipulation to sort the outputs after converting the scripts into sentiments. This data will then be displayed on charts using ‘ggplot2’ which will make the data easier to read.

**Expected outcomes/anticipated results:**

Currently it is not possible to know if there is a link between sentiment and popularity. Some of the best remembered television episodes are ones which have been shocking or surprising in some way, therefore if there is a link it could be positive or negative. Potential outcomes of this include:

* A correlation between sentiment and popularity meaning advertisers could eventually push for more episodes of a certain sentiment.
* A lack of correlation between sentiment and popularity meaning script writers could potentially have more freedom.