

Final Project Outline

Satvik Kishore

Overview

“Star Trek: The Next Generations” is a Sci-Fi TV series that ran from 1987 through 1994 and had 7 seasons and a total of 174 episodes. It has a cult following and continues to be watched today. The series was on an episodic format, with each episode being contained within itself for the most part. For the final project, I propose an analysis into the ratings and perceptions of the episodes of the series. There was a main cast, but episodes often had guest cast, or recurring guest casts. Different episodes also typically had different writers and directors.

Research Questions

My main research questions will be whether episode quality differs across directors, writers, and seasons. The outcome variable used to measure quality will be the IMDB rating.

Data

The data will be pieced from multiple sources.

- <https://github.com/RMHogervorst/TNG> provides data on the episodes at the level of individual dialogues
- IMDB API
- https://en.wikipedia.org/wiki/List_of_Star_Trek:_The_Next_Generation_episodes

The variables included will be: IMDB rating, number of television viewers (in first airing), names of Writers, name of Director, season number, episode number, date of first airing, number of dialogues from the main characters (as proxy for who the main characters in the specific episode were), and other potential variables that can be extracted from the dialogue-level data. I will continue looking for other datasets that can augment this data. The size of the dataset will be 174 observations, that is the number of episodes. This could be a limitation as the sample size is small, and the number of predictor variables can be too many. Multicollinearity amongst predictor variables may make interpretation and inference difficult. Another difficulty is that there can be multiple writers for each episode, and I will need to look into how to construct a hierarchical model for this. If I cannot find a solution for this, I will have to drop the “writer” variable. Another potential route is including data from other Star Trek series as well, where there will be an overlap in the writers and directors.

Project Plan

I plan to use hierarchical models to analyze differences across writers, directors, and seasons. Timeline:

- November first week: Run preliminary analyses, data scraping, and finalize dataset.
- Rest of November: Modelling and writing.