Star Trek: The Analysis 🖖

Statistical Analysis of Episode Ratings across four TV shows for IDS702 class @ Duke University.

Star Trek is a Science Fiction media franchise spanning several TV shows, movies, books, and video games. This analysis focused on four Star Trek TV shows that ran in sequence from 1987 through 2005, with significant overlaps in the years they each ran.

Checkout ./reports/report.pdf for the full analysis.

Data

- IMDb Ratings and Episode information from https://datasets.imdbws.com (collected on Nov 23, 2021). This formed the source for the tabular data used in the analysis
- Script Data from www.chakoteya.net/StarTrek/. This was the source for the script of each episode in text form. This was used to calculate a proxy for each character's relative screentime in each episode. This proxy variable was then merged with the tabular data
- Four TV Shows:
    - Star Trek: The Next Generation (1987–1994)
    - Star Trek: Deep Space 9 (1993–1999)
    - Star Trek: Voyager (1995–2001)
    - Star Trek: Enterprise (2001–2005)
- 614 Episodes in all
- Relevant Variables:
    - Name of Director(s)
    - Name of writer(s)
    - Episode Rating
    - Proxy Character Screentimes for main characters (33 in all)

Research Questions 🏛️

- Which characters have the greatest influence on the ratings of episodes, either positive or negative?
- Are their differences in quality across the four shows and their constituent seasons?
- Are their differences in quality across the directors?

Methodology ⚒️

- Data Cleaning
- Exploratory Data Analysis
- Statistical Modeling
    - Heirarchical Linear Regression
    - Fixed Effects: Character Screentimes
    - Random Effects: Show:Season combination, Name of director(s)
- Tools Used: R 4.1.1

Results 📊

🔍 Influence of Characters

plot

- The figure illustrates the estimated average change in episode rating when a character's screentime is increased by 1 percentage point. This increase is at the expense of the screentime of "other" characters i.e. any character not included in the main character list. Corresponding 95% confidence intervals are also illustrated
- Most characters were not observed to have a significant impact on episode ratings
- Screentime of Sisko from Star Trek: Deep Space 9 and Paris from Star Trek: Voyager have the strongest positive impact on episode ratings
- Screentime of Dr. Crusher, Troi, and Wesley from Star Trek: The Next Generation, and Dax from Star Trek: Deep Space 9 have the strongest negative impact on the ratings

🔍 Differences across Shows and Seasons

Alt text

- Significant difference in quality across Show:Seasons was observed. Season 1 of The Next Generations was the worst while Season 6 of The Next Generations was the best

🔍 Quality of Directors

There was no evidence for difference in quality across directors. This is likely due to the small number of episodes per director.