Overview

Over 100 television series are broadcasted across different networks in the US every year, and the question remains at the end of every tv season: which tv series will be renewed for a next season? Using the renewal status of tv series over the last 5 years (2014 to 2018), my analysis aims to identify factors influencing the odds of renewal of tv series and predict the renewal status of tv series in the 2019/2020 season.

Research questions

Inspiration of the project came from Spoilertv.com, where a numerical index was developed to predict the likelihood of cancellation. Instead of using a multinomial model on the cancellation likelihood, I wish to apply a logistic regression model on the odds of renewal.

The goal of the project is to explore the factors that influence the odds of renewal for tv series.

- How did the viewership, network ratings or total viewers, affect the odds of renewal? And other show-related factors, such as day or time of airing, length of series, availability of show on streaming platform?
- How did the odds of renewal differ by network, genre?
- How did production-related factors affect the odds of renewal, such as production cost, cast availability, DVD/Blu-Ray sales.

Data

The base dataset is obtained from spoilertv.com, where the yearly renewal status of all tv series, with viewing numbers and network information. Currently in contact with spoilertv.com to share the dataset on other factors included in their analysis. If unsuccessful, other factors will have to be merged from other publicly available websites, such as IMDB. There is roughly about 500 observations by compiling the yearly list of tv series for the past 5 years. The most significant problem expected will be errors in joining different sources of data, due to lack of unique tv series key/id.

Project plan

I plan to explore logistic regression models, both regular and hierarchical. Methods for missing value imputation may be explored depending on the completeness of the data. The data gathering and data cleaning process is expected to be done by 15 Nov 19, while model building process to be completed by 22 Nov 19 with remaining time for report writing and presentation.