ABSTRACT

My project is on adding a wider selection of titles to Netflix in order to increase member viewership and effectively reducing churn. To try and provide quality suggestions to Netflix (my client in this case), I used Tableau and Google Sheets to do some analysis on how many shows are in each of their genres (ex: comedy, romantic, documentaries) and what their popularity is (by looking at IMDB ratings). Then, I looked at which categories had high average ratings and low title counts and made suggestions that fixing this problem will lead to more member satisfaction and therefore less churn. I finished the project by suggesting that future data science work can be in the form of a regression model to try and predict viewership hours using already collected data and some additional variables that still need to be collected in the future. Creating a model like this will be a good way to predict if these suggested routes will be effective or not.

DESIGN

This project was created because early this year Netflix lost a huge amount of momentum in their stock prices, predicted membership numbers and more. I thought it would be a cool idea to do some initial EDA and visualizations to think of some ways to get them back on track.

DATA

Kaggle dataset of all current United States Netflix titles in May 2022
https://www.kaggle.com/datasets/victorsoeiro/netflix-tv-shows-and-movies

Algorithms

Suggested regression models for future work predicting viewership hours

TOOLS

Google Sheets for analysis and filtering of data
Tableau for some filtering, analysis, and visualizations

Communication

The main way of communicating this project is through Google Slides