Analyzing Netflix's Strategic Shift Towards Diversity

Introduction

This project investigates Netflix's strategic shift towards diversity by analyzing the ethnic representation of actors over time. By showing the diversity index of actor's ethnicities and of the production countries and creating visualizations, we aim to demonstrate the increasing inclusivity in Netflix's casting choices. We also aim to show the shift in content, but with only the description to analyze, sufficient conclusions were not made.

The analysis process is shown in the Jupyter Notebook in our Github repository.

Methodology and Tech

Ethnicity Inference: Removed duplicate names and used a model trained on US census and voter data from the `ethnicolr` library to predict ethnicities based on names, and chose the race with the highest probability.

Calculating Diversity Index: the Simpson diversity index was used throughout this analysis.

Visualizations
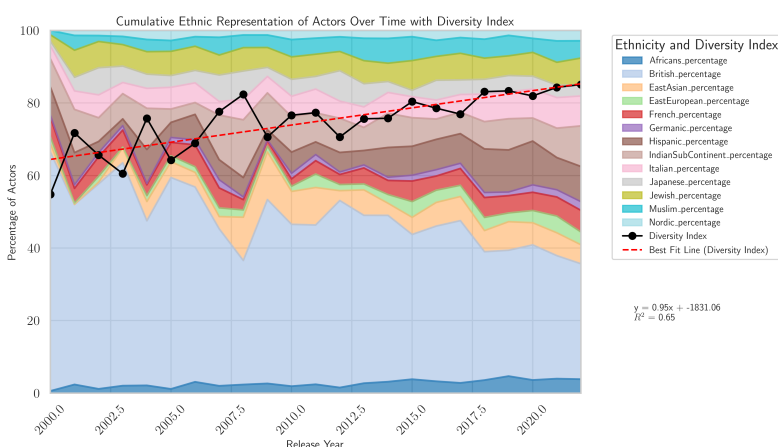
1. Ethnic Representation Over Time

An area chart was created to show the cumulative ethnic representation of actors from 2000 to 2022.

The diversity index was overlaid as a black line, with a red dashed best fit line representing the trend. As annotated on the graph, the R^2 value was 0.65, indicating a reasonably strong positive correlation.

2. Simpson Index of Production Countries Over Time

A line chart was created to display the Simpson Diversity Index of production countries.

The diversity trend is shown over time, providing insights into how Netflix has expanded their offerings. The fluctuations until around 1995 were caused by a small sample size, as Netflix typically streams newer titles, but afterwards there is a strong positive trend.