GEORGE WASHINGTON UNIVERSITY

Introduction to Data Science (DATS 6101)

Team: Data With Density

GitHub repo: ​​<https://github.com/evycampo/T1DATAWITHDENSITY.git>

Kaggle source: <https://www.kaggle.com/ruchi798/tv-shows-on-netflix-prime-video-hulu-and-disney>

Project Proposal

Title: Tv shows found on Netflix, Hulu, Prime video, Disney+

RESEARCH TOPIC

Some of the world's largest entertainment giants have ventured on streaming entertainment during the previous decade, including Netflix, Hulu, Prime Video, and Disney+. As more individuals are compelled to stay at home to prevent the spread of the new coronavirus, the idea of a bored, cable-cutting consumer looking for shows, documentaries or series to watch for weeks on end has become a reality. As a result, Tv shows found on Netflix, Hulu, Prime video, and Disney+ is our selected topic for the project. Figure 1 shows how the process to analyze the rate of tv shows that have been streaming over time, the most popular streaming platform, and targeted audience will be conducted.

**Work Flowchart:**



*Figure 1. Process of the Project*

DESCRIPTION OF AVAILABLE DATA:

The source of the dataset is Kaggle Sample Dataset where it was extracted as a CSV format. The data consists of 5367 observations and 11 variables (ID, Title, Year, Age, IMDb, Rotten Tomatoes, Netflix, Hulu, Prime Video) The dataset constitutes data of various types like numerical and categorical.

SMART QUESTIONS:

1.What are the most targeted age groups for the TV shows by Netflix, Hulu, Prime Video?

2.Which year published the highest number of TV shows?

3.Which streaming platform has the highest average rating (according to Rotten Tomatoes and IMDb)?

Input Data:

Independent variables being considered: age group, streaming platform,

Output Data:

Response being considered: ratings