Project Proposal → Analyzing Competing Streaming Services

By Madison Mihle and Kat Hardy

Ideas → streaming services (Netflix, Hulu, Apple, HBO Max, Amazon Prime)

What we want to look at:

- We will want to look at data trends of all the streaming services presented
    - Age analytics
    - Gender
    - Categories
    - Country
    - Movies vs TV series
    - Average number of profiles on each one
- Cost analysis → compare the cost of each streaming service
- We will figure out which streaming service is the most popular? Least popular? → from here we will dive deeper into why it is the most popular and assess the data presented.
    - If Netflix → see and pinpoint active accounts/cost
        - Type of subscription
        - Plan duration

Questions to ask:

1. Which streaming service has the most content? Is this a positive or a negative when looking at a streaming service?
2. Which streaming service has the highest IMDB rating? Which streaming service has the lowest? Why might that be?
3. What are the most popular TV shows and movies of the highest rated streaming service? What are the most popular titles of each?
4. How can we improve the lowest-rated streaming service? What are the future aspects we can look at regarding demographic information and membership analysis?

Where will we source our data?

- Kaggle
    - Netflix: -KH
        - https://www.kaggle.com/datasets/arnavsmayan/netflix-userbase-dataset
        - https://www.kaggle.com/code/arnavsmayan/netflix-userbase-visualization-notebook
        - https://www.kaggle.com/datasets/octopusteam/full-netflix-dataset - recent
            - Small summary
        - https://www.kaggle.com/code/darshandungrani/netflix-data-cleaning-analysis-and-visulization

    - HBO: - MM
        - https://www.kaggle.com/datasets/victorsoeiro/hbo-max-tv-shows-and-movies
        - https://www.kaggle.com/datasets/octopusteam/full-hbo-max-dataset - recent
        - https://www.kaggle.com/code/mattlong360/hbo-max-data-analysis
    - Amazon Prime: -KH
        - https://www.kaggle.com/datasets/victorsoeiro/amazon-prime-tv-shows-and-movies
        - https://www.kaggle.com/datasets/octopusteam/full-amazon-prime-dataset - recent
    - Apple TV -MM
        - https://www.kaggle.com/datasets/octopusteam/full-apple-tv-dataset
        - https://www.kaggle.com/code/nadeemkaggle123/apple-tv-movies-tv-series-dataset-analysis
    - Hulu: - MM
        - https://www.kaggle.com/datasets/victorsoeiro/hulu-tv-shows-and-movies
        - https://www.kaggle.com/datasets/octopusteam/full-hulu-dataset - recent
- IMDB
    - https://www.kaggle.com/datasets/octopusteam/full-imdb-dataset
- Rotten Tomatoes?

Breakdown

- Task 1: Clean the data, analyze, and place in a summary table of each streaming service and go into analysis.
    - Netflix
    - Amazon Prime
    - Apple TV
    - HBO Max
    - Hulu
- Task 2: Take each of the streaming services and analyze them deeper. Create new columns to be able to differentiate the streaming service.
    - 2a → See which one has more TV and movies.
    - 2b → See which one has the highest IMDB rating. Discuss why that might be.
- Task 3: From the data received, we will take the most popular streaming service and analyze from there. Create various charts and graphs.
    - 3a → look at the most popular genres of movies vs TV
    - 3b → look at the top 10 titles for each category
- Task 4: Take the least popular and analyze why that might be? What could be done to improve it? Want to look at demographics and membership. This will include the analytics of:
    - Age distribution, gender, country distribution, subscription types, membership length