Netflix Data Analysis

Problem Statement

The problem is to analyse the Netflix dataset using Pandas DataFrame. The goal is to understand the content on Netflix and to identify any patterns or trends.

Data

The data used for this analysis is a dataset of TV shows and movies available on Netflix as of 2021. The dataset includes information about the title, release date, category, type, country, rating, duration, and cast of each show or movie.

Analysis

The analysis was performed using the following steps:

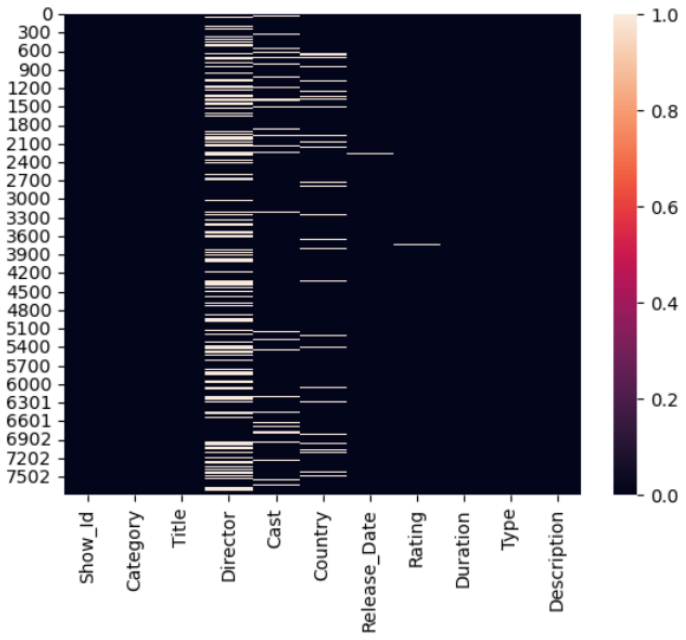
1. The data was loaded into a Pandas DataFrame.
2. The basic information about the data was explored, including the number of rows and columns, the data types, and the unique values in each column.
3. The data was filtered to find specific instances, such as when the show or movie was released in a particular year or when it had a particular rating.
4. The data was grouped by different categories, such as category, type, and country, to find the most popular shows and movies in each category.
5. The data was analyzed to find any patterns or trends, such as the number of shows and movies that have been released over time or the ratings of shows and movies from different countries.

By: HAFSA BAIG

```
:  import seaborn as sns

:  sns.heatmap(data.isnull())

[22]: <Axes: >
```



Results

The results of the analysis showed that the most popular category on Netflix is TV Shows, followed by Movies. The most popular type of TV Show on Netflix is Dramas, followed by Comedies. The most popular country for Netflix shows and movies is the United States, followed by the United Kingdom and Canada. The most popular rating on Netflix is TV-14, followed by PG-13 and R.

The analysis also showed that the number of shows and movies on Netflix has been increasing over time. In 2013, there were about 1,000 shows and movies on Netflix. In 2021, there were over 6,000 shows and movies on Netflix.

By: HAFSA BAIG