Report

on

Understanding Netflix Content Trends Through Data Science

Abstract

This report presents an end-to-end data science analysis performed on the Netflix Movies and TV Shows dataset. The purpose of this work is to clean, analyze and visualize data to extract meaningful insights. Various data pre-processing, descriptive statistics and visualization techniques were used to understand content distribution and trends.

Introduction

Data Science involves extracting meaningful information from raw data using statistical and analytical techniques. It plays an important role in understanding patterns and trends in large datasets. This assignment focuses on applying data science techniques to a real-world dataset.

Dataset Description

The Netflix Movies and TV Shows dataset contains information such as title, type, director, cast, country, release year, rating, duration, genre, and description. The dataset provides insights into the content available on Netflix.

Data Preprocessing

The dataset was initially explored to understand its structure. Unnecessary columns were removed, and missing values were handled using appropriate replacement techniques. Duplicate records were also removed to ensure data consistency and quality.

Descriptive Analysis and Visualization

Descriptive statistics were used to analyze content types, ratings, countries, and release years. Visualizations such as bar charts and line graphs were created to identify trends.

The analysis showed that movies dominate Netflix’s library and that content production increased significantly after 2015.

Conclusion

This data science analysis demonstrated how preprocessing, descriptive statistics, and visualization can be used to gain insights from real-world datasets. The Netflix dataset revealed meaningful trends related to content growth and distribution.