**NETFLIX MOVIES AND TV SHOWS CLUSTERING**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:** |
| Name – Kartika Sharma  Email-id – [kartikasharma17@gmail.com](mailto:kartikasharma17@gmail.com)  Contribution –   * EDA * Null values * Data cleaning * Data preprocessing * Topic Modelling * Movie and TV-show recommendation * K-means |
| **Please paste the GitHub Repo link.** |
| Github Link:- https://github.com/cartika/NETFLIX-MOVIES-AND-TV-SHOWS-CLUSTERING.git |
| **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)** |
| **Netflix is an application that keeps growing bigger and faster with its popularity, shows and content.** **Netflix is the world’s leading Internet television network with over 83 million members in over 190 countries enjoying more than 125 million hours of TV shows and movies per day, including original series, documentaries and feature films. Members can watch as much as they want, anytime, anywhere, on nearly any Internet-connected screen. Members can play, pause and resume watching, all without commercials or commitments. This dataset consists of tv-shows and movies available on Netflix as of 2019. The dataset is collected from flixable which is a third-party Netflix search engine.** **In 2018, they released an interesting report which shows that the number of TV shows on Netflix has nearly tripled since 2010. The streaming service’s number of movies has decreased by more than 2,000 titles since 2010, while its number of TV shows has nearly tripled. It will be interesting to explore what all other insights can be obtained from the same dataset.**  **Integrating this dataset with other external datasets such as IMDB ratings, rotten tomatoes can also provide many interesting findings.** **The dataset consists of meta details about the movies and tv shows such as the title, director, and cast of the shows / movies. Details such as the release year, the rating, duration etc. As the first step, let's load the dataset, create some new features. In this kernel, I have analysed this dataset to find top insights and findings.** |