

**WiDS ‘22 - ‘23 Final Documentation**

### **2 - Data analytics using Python**

**Aditya Prakash**

|  |  |  |
| --- | --- | --- |
| **Team Member Name** | **Roll Number** | **Email-Id** |
| Merugu Rakesh | 204100006 | 204100006@iitb.ac.in |

**Introduction to Problem Statement**

|  |
| --- |
| Following tasks on Netflix data set,  ● Understanding what content is available in different countries  ● Identifying similar content by matching text-based features  ● Network analysis of Actors / Directors and find interesting insights  ● Does Netflix has more focus on TV Shows than movies in recent years. |

**Existing Resources**

|  |
| --- |
| NA |

**Proposed Solution**

|  |
| --- |
| Read the data frame and get the info. Clean the data before analysis. Sorting the data. Plot (Matplot lib) the graphs between different parameters and draw conclusions. |

**Methodology & Progress (Mention the work done week-wise)**

|  |
| --- |
| Week-3  Introduction to numphy library  Basic operations on the given set using numphy  Week-4  Introduction to seaborn library  Applying it to the data set for plotting along with marplot lib. |

**Results**

|  |
| --- |
| 1. Top 10 popular country for moves and tv shoes   C:\Users\Mr\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\C4E3C682.tmp  2. Top 10 popular types of Movies and TV shows  C:\Users\Mr\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\F03E29E0.tmp  3. Top 10 directors  C:\Users\Mr\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\392F8AEE.tmp  4. count Vs movies added, Blue= Movies and Orange= TV shows  C:\Users\Mr\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\3F36D154.tmp |

**Learning Value**

|  |
| --- |
| **Handling the text types data frames. Use of value count function and use of seaborn library**  **for plotting.** |

**Suggestions for others**

|  |
| --- |
| NA |

**References and Citations**

|  |
| --- |
| Python NumPy Tutorial for Beginners <https://www.youtube.com/watch?v=QUT1VHiLmmI&ab_channel=freeCodeCamp.org>   1. Seaborn Tutorial: Seaborn Full Course   <https://www.youtube.com/watch?v=6GUZXDef2U0&ab_channel=DerekBanas> Python Plotting Tutorial w/ Matplotlib & Pandas (Line Graph, Histogram, Pie Chart, Box & Whiskers) <https://www.youtube.com/watch?v=0P7QnIQDBJY&t=58s> |