**Data Analysis of Movies and TV Shows using IMDB and Netflix Dataset.**

By

**STANLEY KINGSTON**   **19BEC1254**

**GAURAV PRASANNA**  **19BEC1315**

**VIGNESH S 19BEC1391**

A project report submitted to

**Dr SATHIYA NARAYANAN S**

**SCHOOL OF ELECTRONICS ENGINEERING**

in partial fulfilment of the requirements for the course of

**CSE3505 – Foundations of Data Analytics**

in

**B.TECH. - Electronics and Communication**

****

**Vandalur – Kelambakkam Road**

**Chennai – 600127**

**DECEMBER 2021**

**BONAFIDE CERTIFICATE**

Certified that this project report e­­­­ntitled “**Data Analysis of Movies and TV Shows using IMDB and Netflix Dataset.”** is a bonafide work of **Stanley Kingston-19BEC1254, Gaurav Prasanna-19BEC1315 and Vignesh S-19BEC1391** who carried out the project work under my supervision and guidance for **CSE3505 - Foundations of Data Analytics.**

**Dr SATHIYA NARAYANAN S**

Assistant Professor (Senior Grade 1)

School of Electronics Engineering (SENSE),

Vellore Institute of Technology (VIT Chennai)

Chennai – 600 127.

**ABSTRACT**

In this project, we have done exploratory data analytics on IMDB, Netflix and OTT platform datasets. A basic overview of how this works and its criteria, including its main assumptions and inferences are discussed. These tasks are done using various data analytics and visualizations software such as RStudio, Excel and Tableau. Various pre-processing techniques and libraries have been used to implement this project in RStudio. This will enable users and audience to have better understanding on which OTT platform to invest in for better content after looking at the visuals or the plots provided.

**ACKNOWLEDGEMENT**

Give honest acknowledgement thanking those who have supported you.

**Student 1 Name and Signature Student 2 Name and Signature**

**TABLE OF CONTENTS**

|  |  |  |  |
| --- | --- | --- | --- |
| **SR. NO.** |  | **TITLE** | **PAGE NO.** |
|  |  | ABSTRACT | 3 |
|  |  | ACKNOWLEDGEMENT | 4 |
|  |  |  |  |
| 1 |  | INTRODUCTION | 6 |
|  |  |  |  |

**CHAPTER 1**

**INTRODUCTION**

**1.1 BACKGROUND AND MOTIVATION**

**1.2 PROBLEM STATEMENT AND OBJECTIVES**

**CHAPTER 2**

**<Describe your main work>**

**2.1 PROPOSED/IMPLEMENTED METHOD**

**2.2 ADVANTAGES**

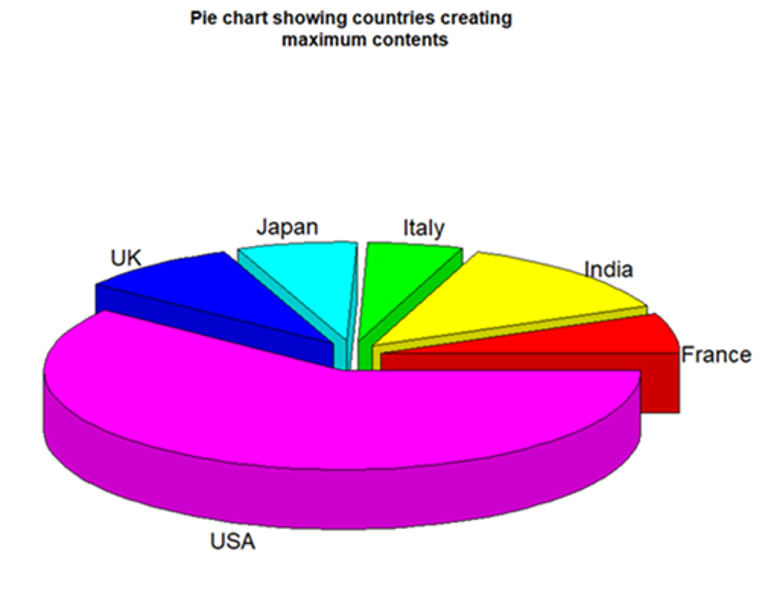
**2.3 CHALLENGES FACED**

**2.4 CODE**

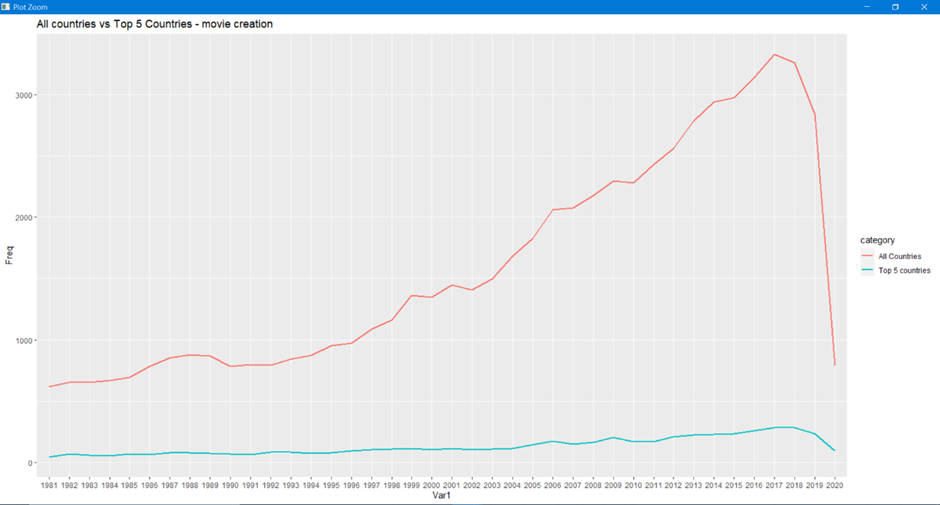
**CHAPTER 3**

**<Analyse and discuss your results>**

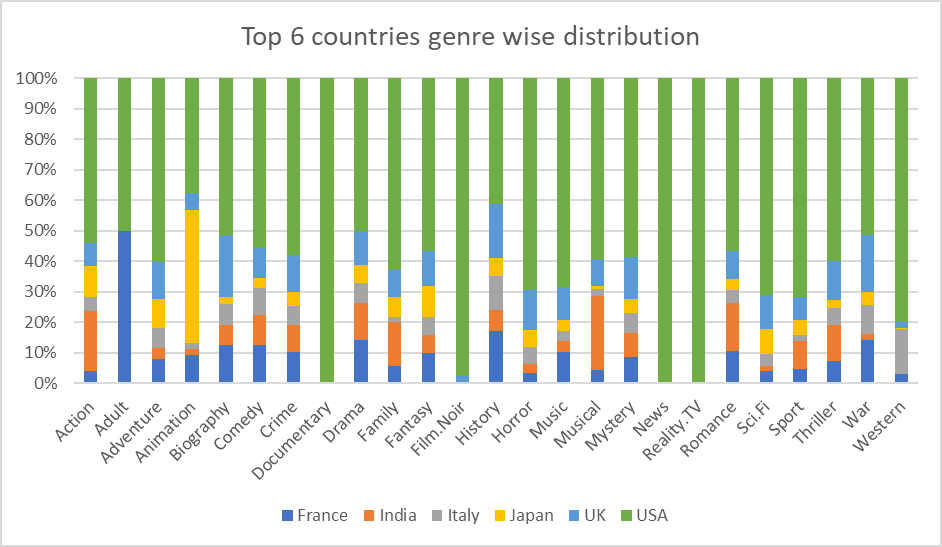
**3.1 MAIN RESULTS**



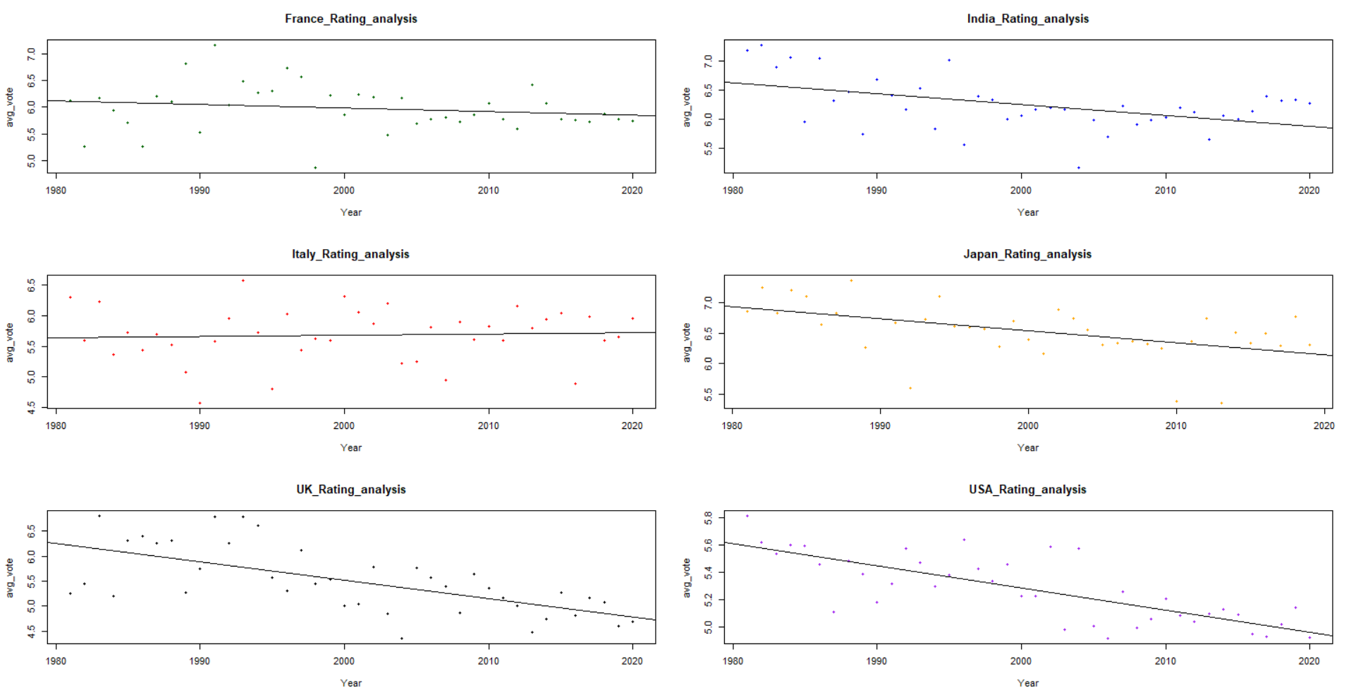
**Figure[1]:** Countries producing the most content in recent times



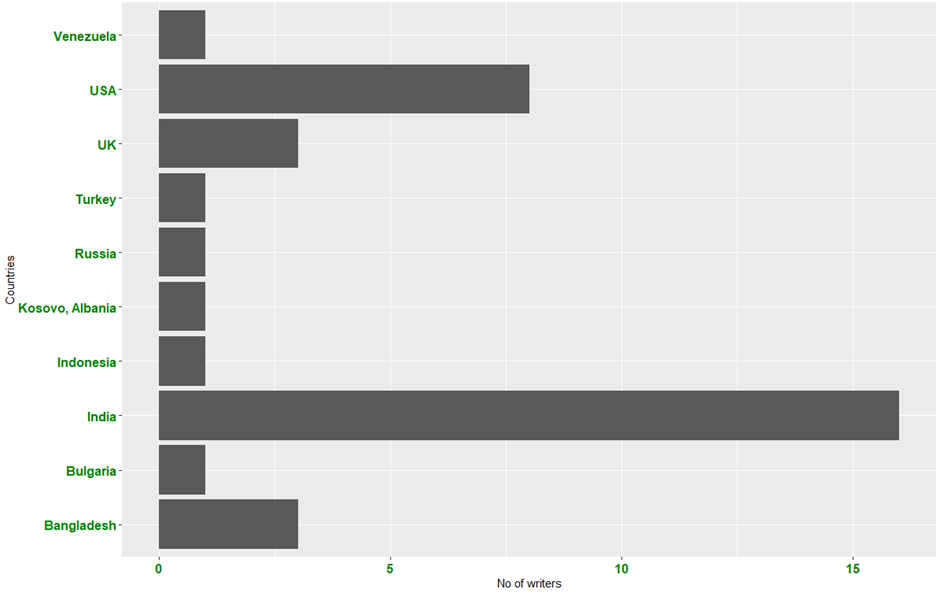
**Figure[2]:** Comparing the total movies created by the whole world vs top 6 contributors for the past 40 years.

****

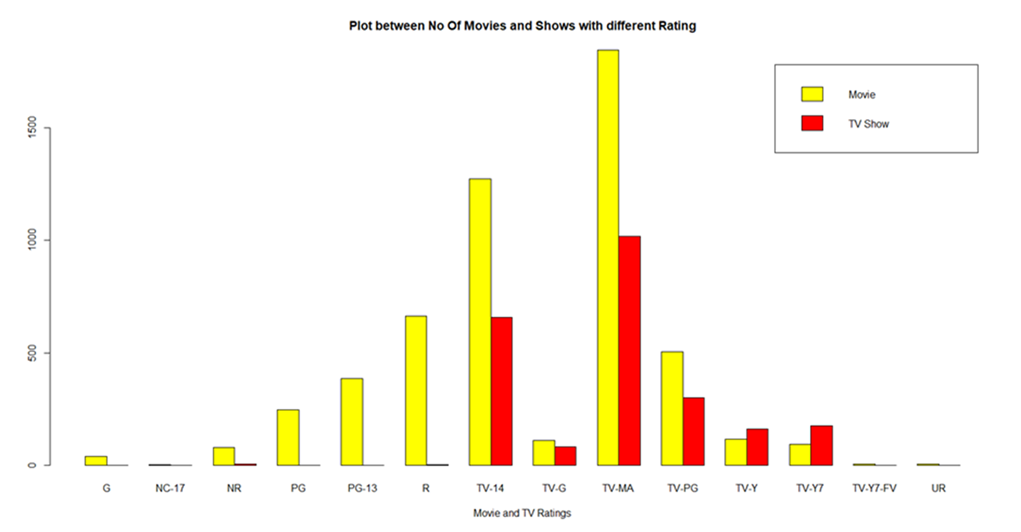
**Figure[3]:** Contribution by the top 6 countries towards each genre



**Figure[4]:** Rating Improvement for the top 6 contributing countries from 1980-2020

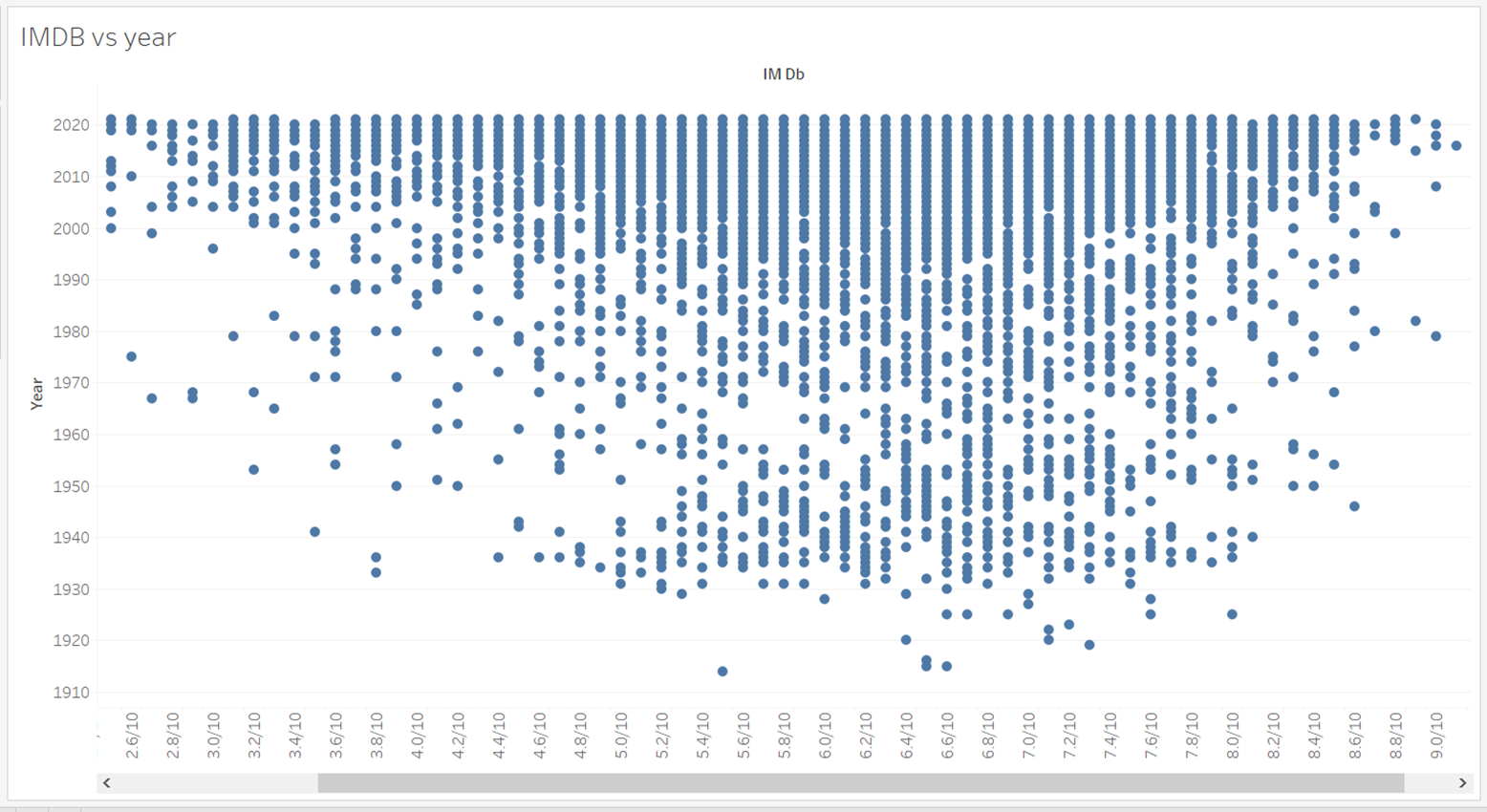
****

**Figure[5]:** Count of most voted writer across the world

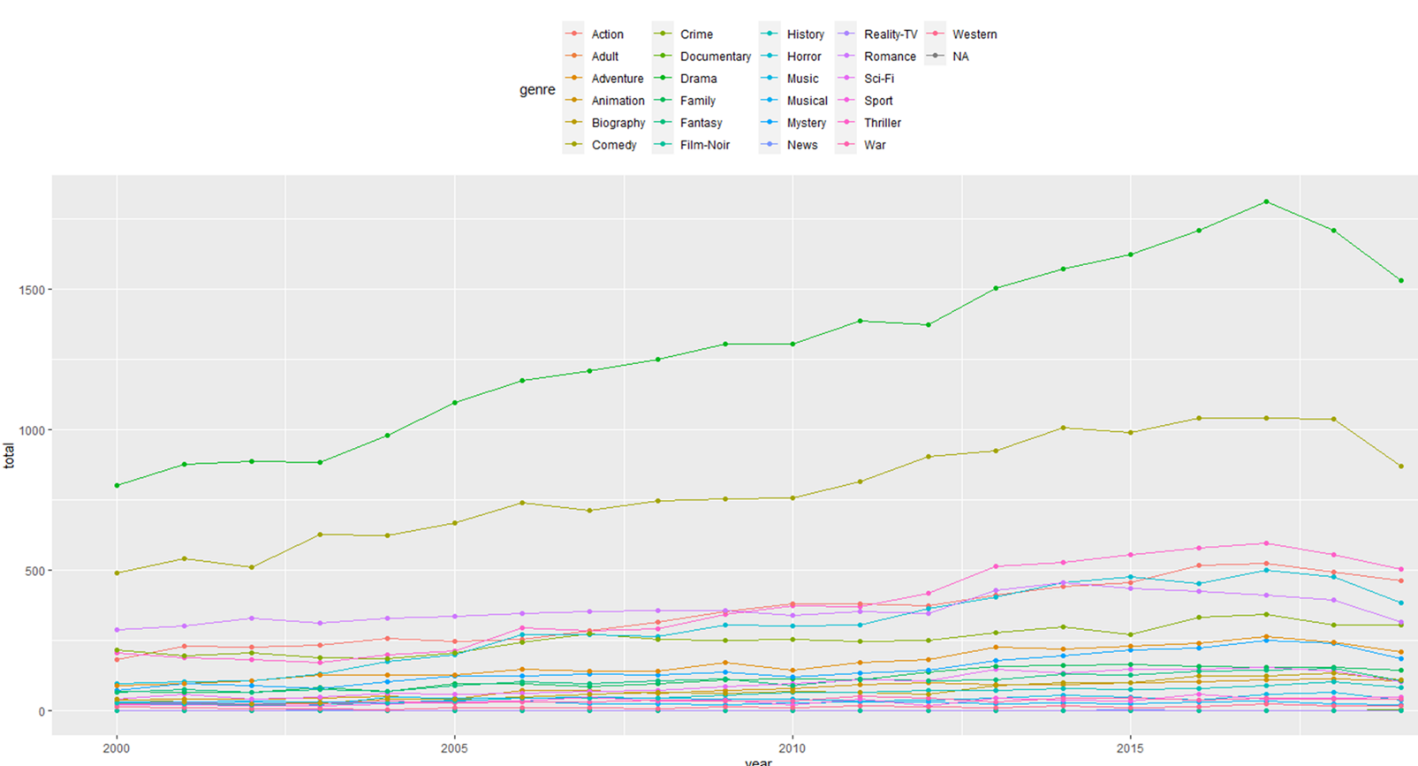


**Figure[6]:** Count of Movies and TV Shows rating in Netflix

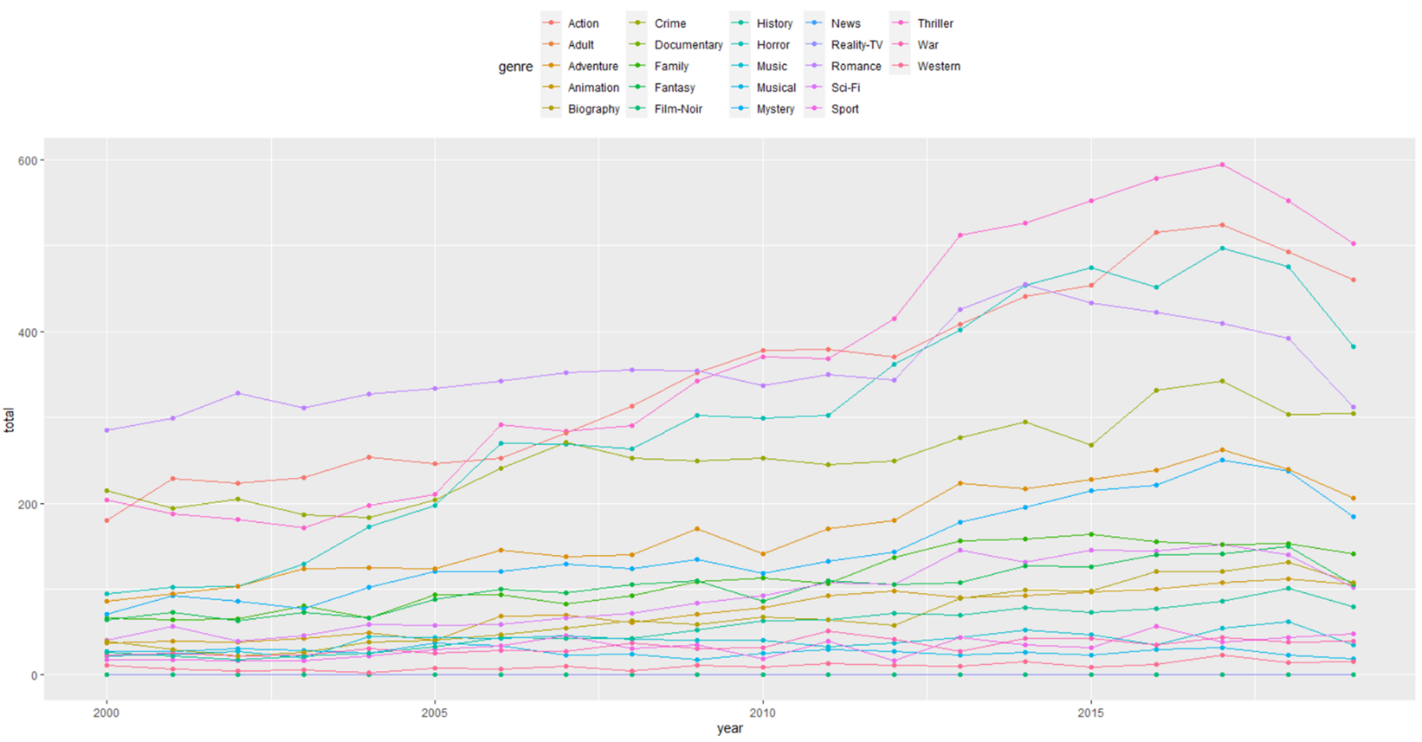
**Figure[7]**: Count of OTT platforms with content with IMDB rating greater than 8.0



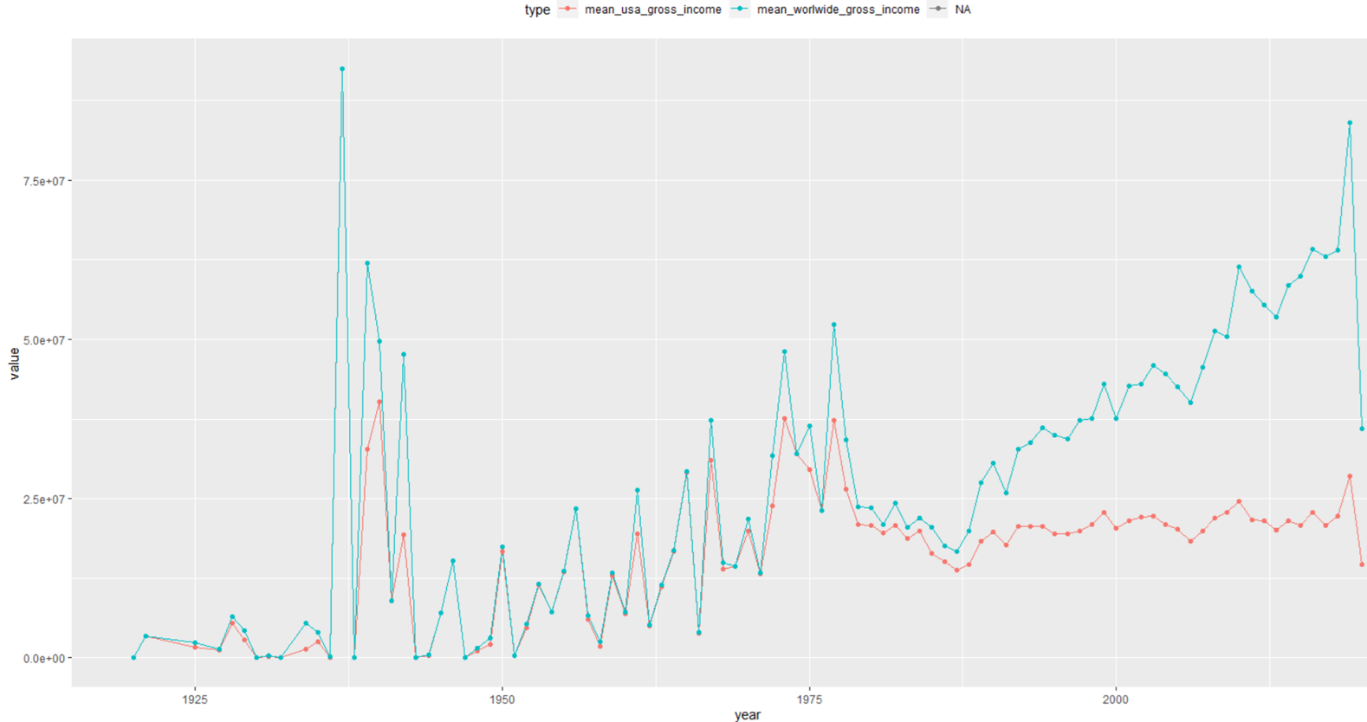
**Figure[8]**: Evolution of Movie and Film industry throughout the years



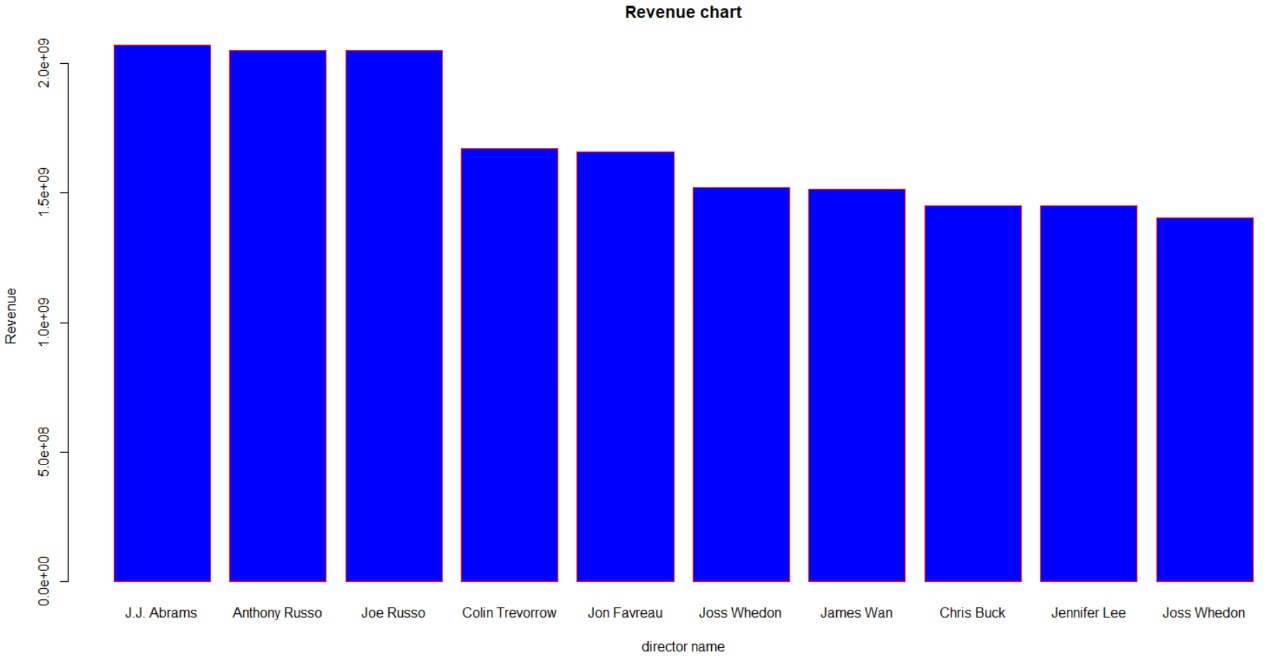
**Figure[9]:**

****

**Figure[10]:**



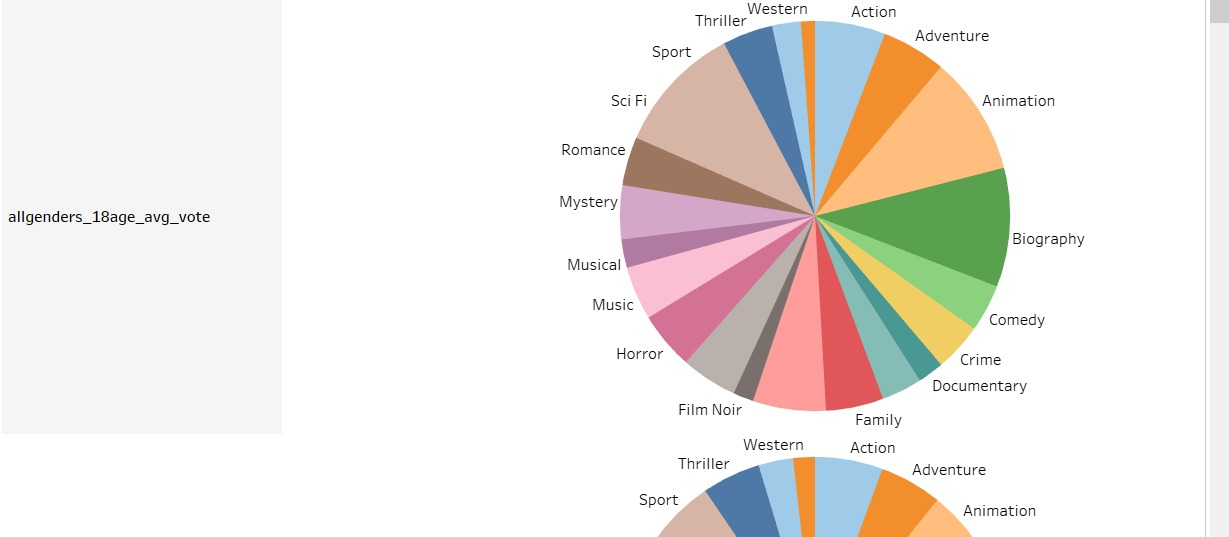
**Figure[11]:**

****

**Figure[12]:**

****

**Figure[13]:**

****

**Figure[14]:**

****

**Figure[15]:**

**3.2 INFERENCES**

**CHAPTER 4**

**CONCLUSION AND RECOMMENDATION FOR FUTURE WORK**

**4.1 CONCLUSION**

**4.2 RECOMMNDATIONS FOR FUTURE WORK**

**REFERENCES**

List all sources you referred and make sure you cite it in appropriate place(s) in the text.