**Project Report**

**Team Number-14**

**Lanka Chaitanya Krishna**

**1001675459**

**Himaja Koganti**

**1001646544**

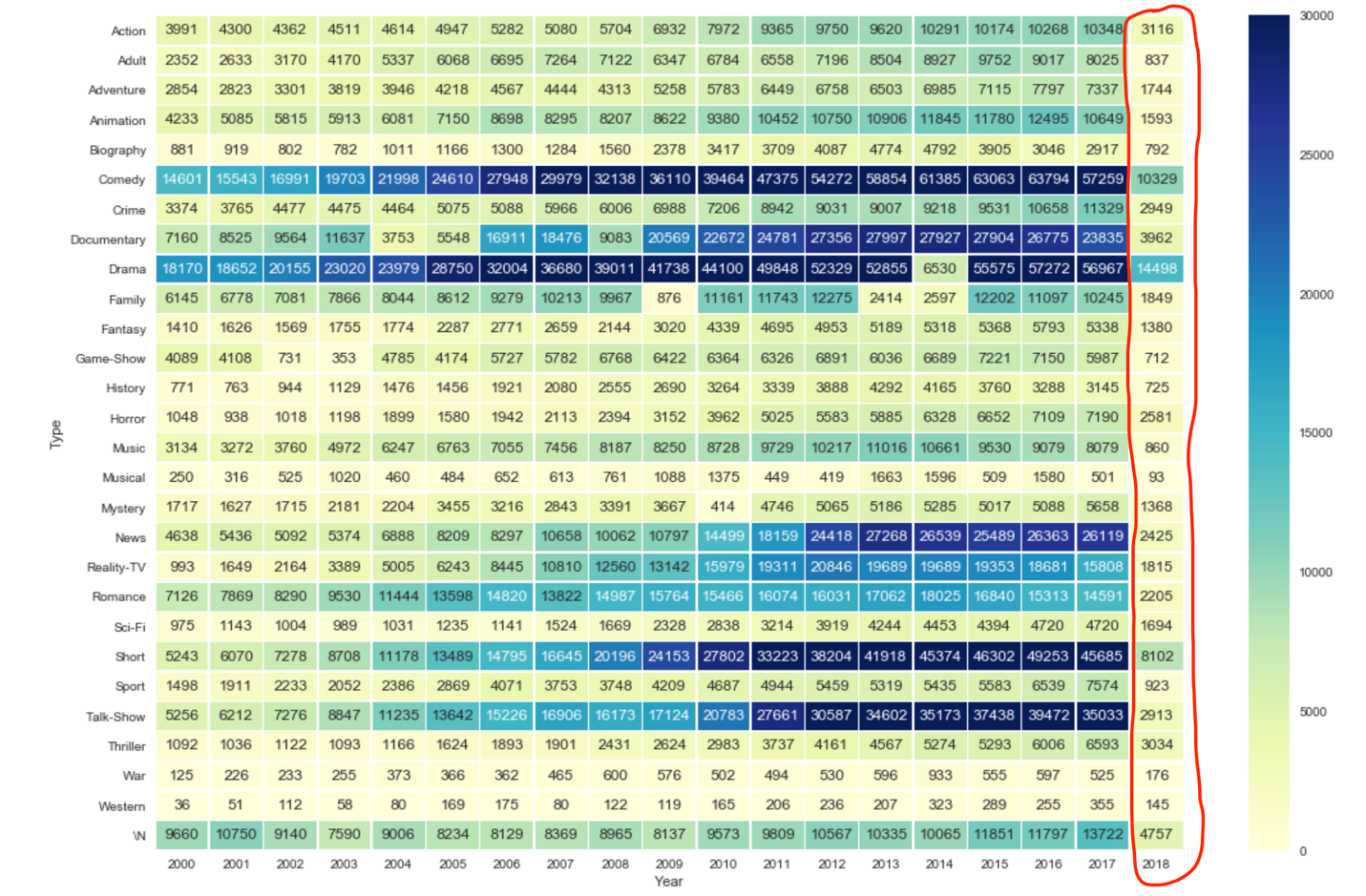
**Overall Status**

We have implemented a Map/Reduce program. The Mapper takes key values as IntWritable and each line in the text file. The mapper outputs the genre and year as key value pairs. The Reducer receives the key as genre and values as a iterable collection of years. Then the reducer counts the numbers of each type of years. The reducer outputs the year, genre as the key and the count as value.

**Analyzation and Plotting of Inferences**

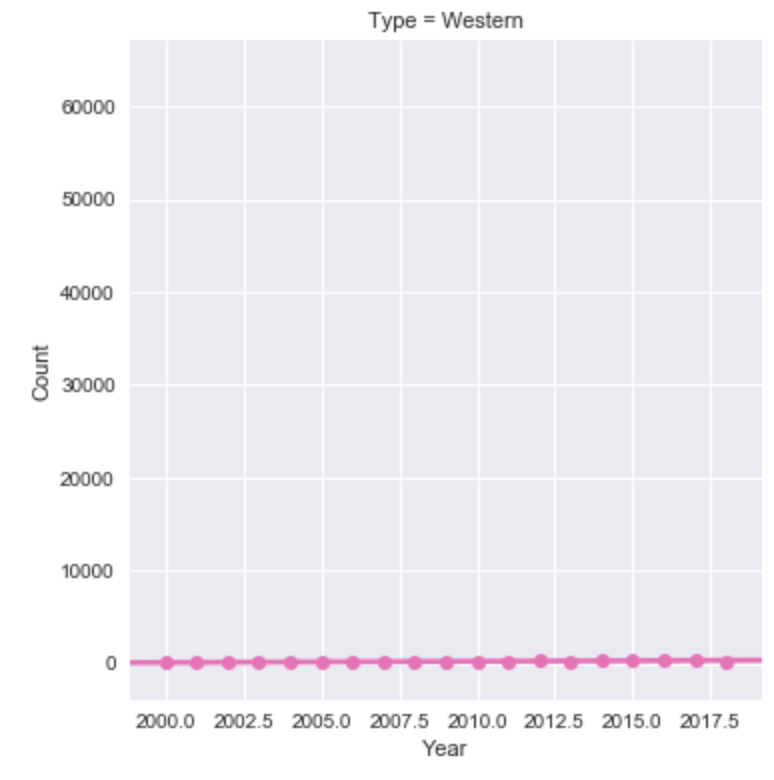
**I. Inference 1**

* The number of movies released in all genres in the year 2018 has drastic downfall when compared to the previous year, 2017.This might be because the database may have not yet been updated for 2018.

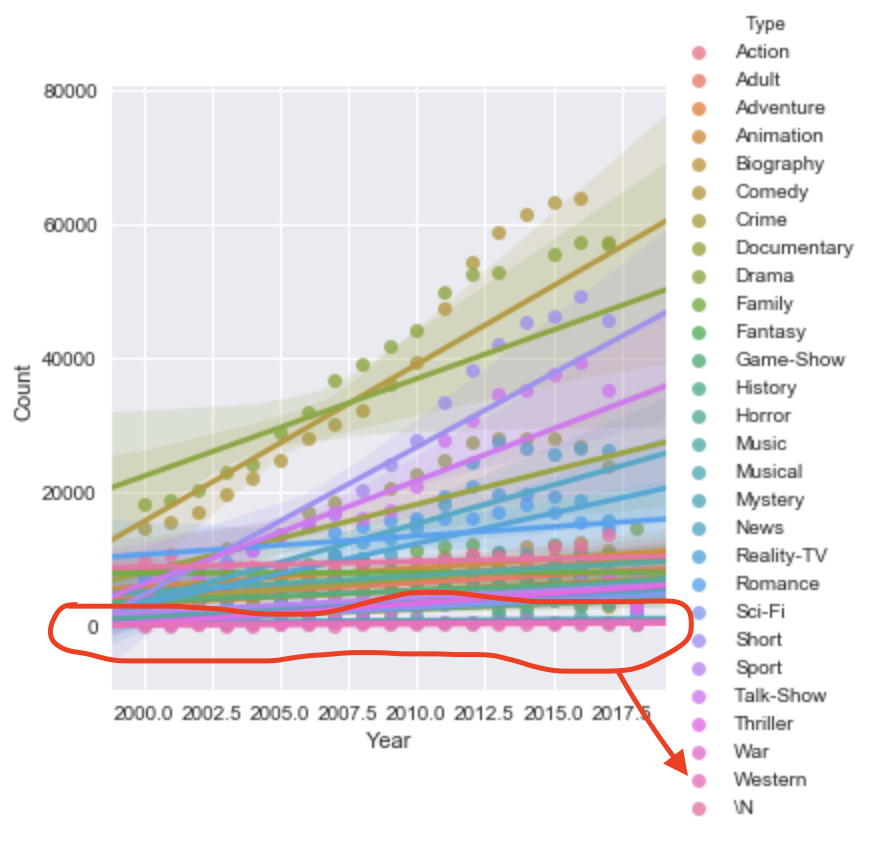


**II.Inference 2**

* The number of movies released in genre Western is almost constant when compared to the number of movies of other genres. And Western was also the genre with least moives in all the years.



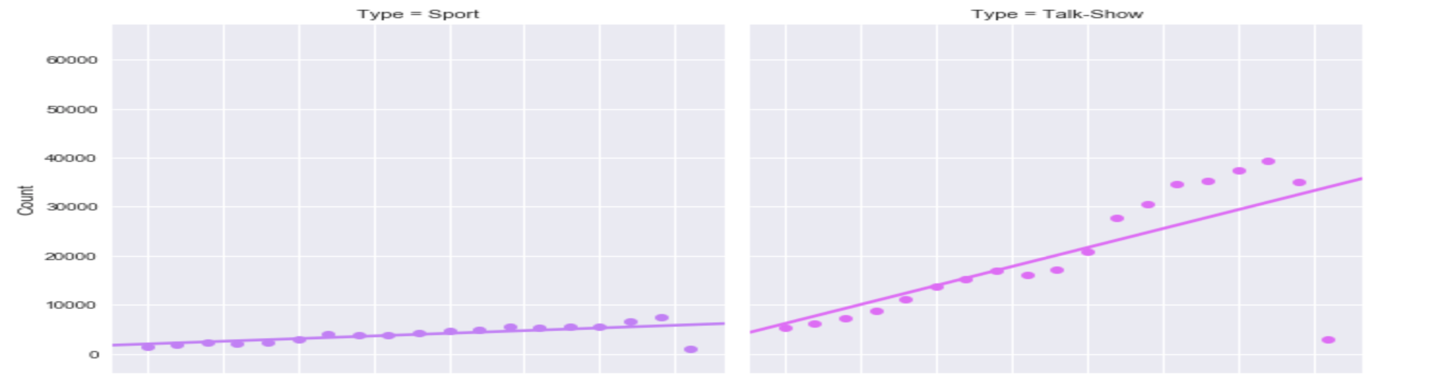
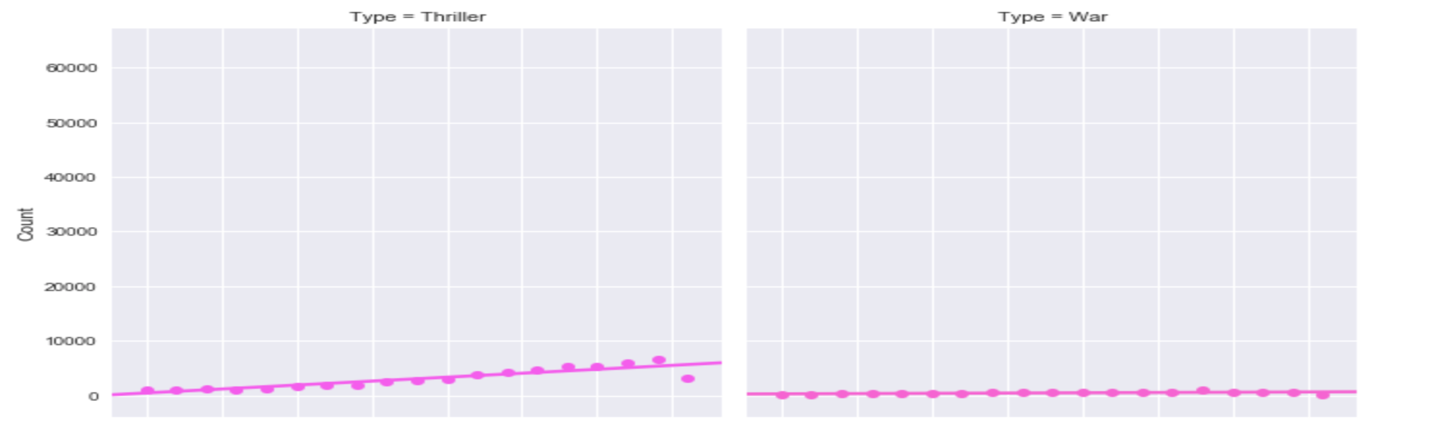
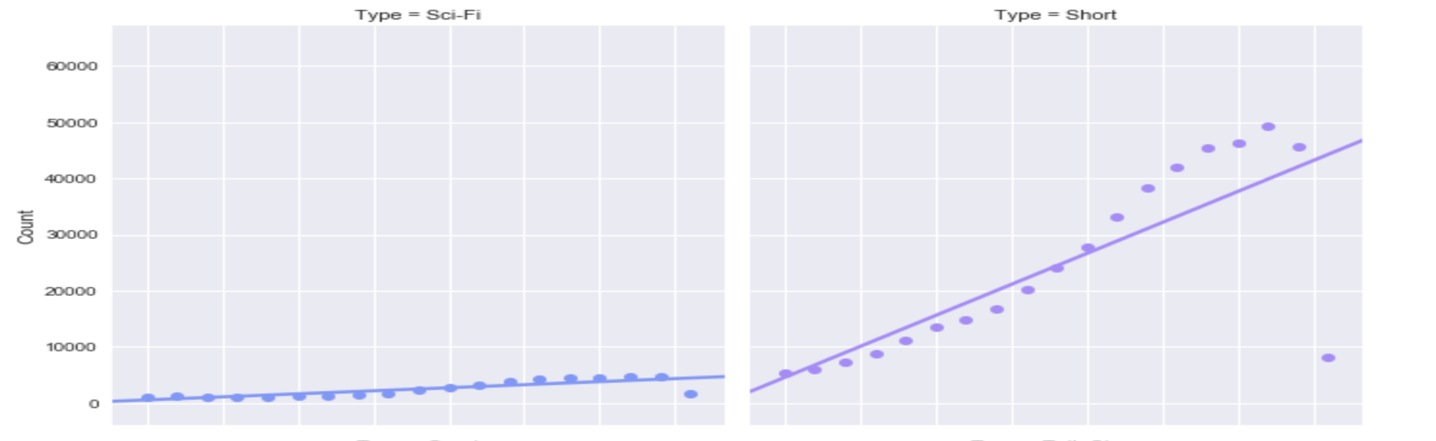
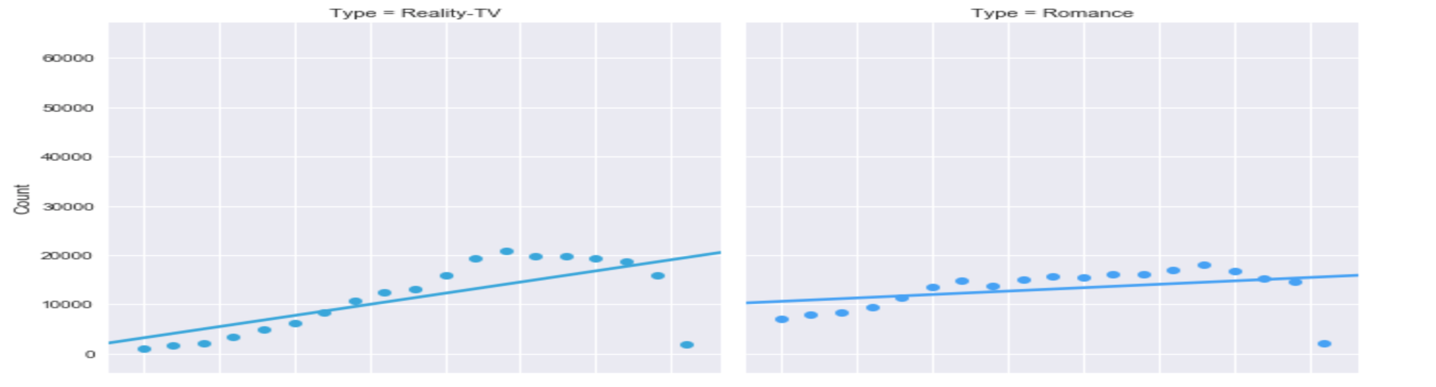
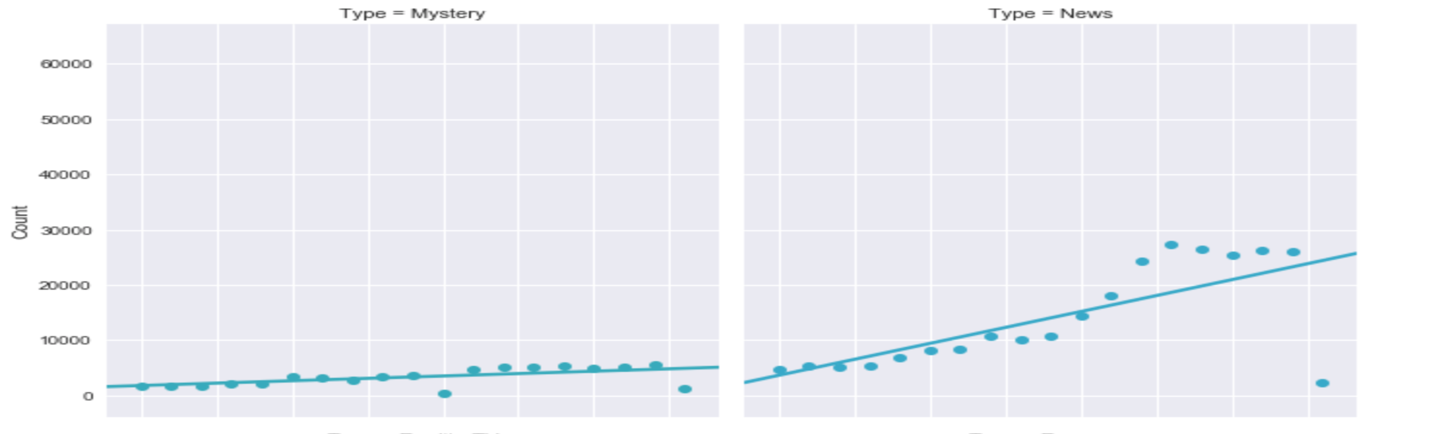
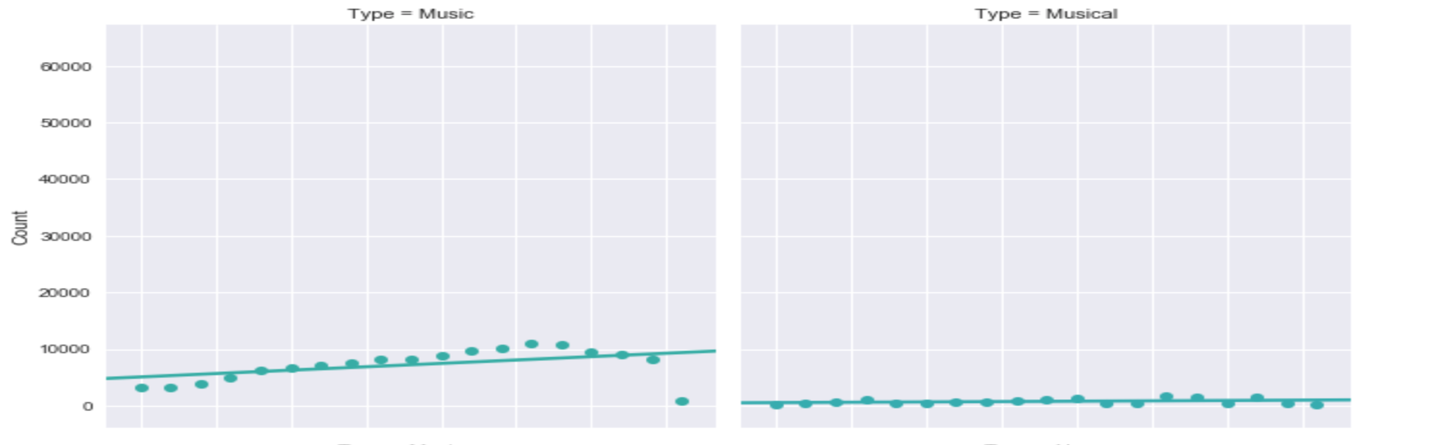
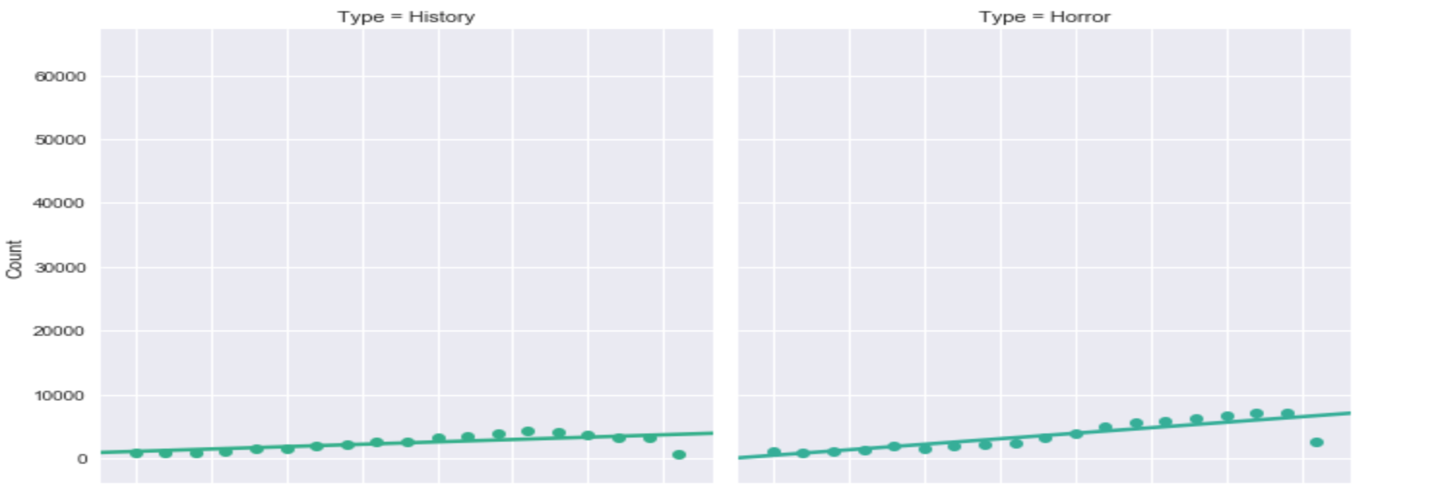
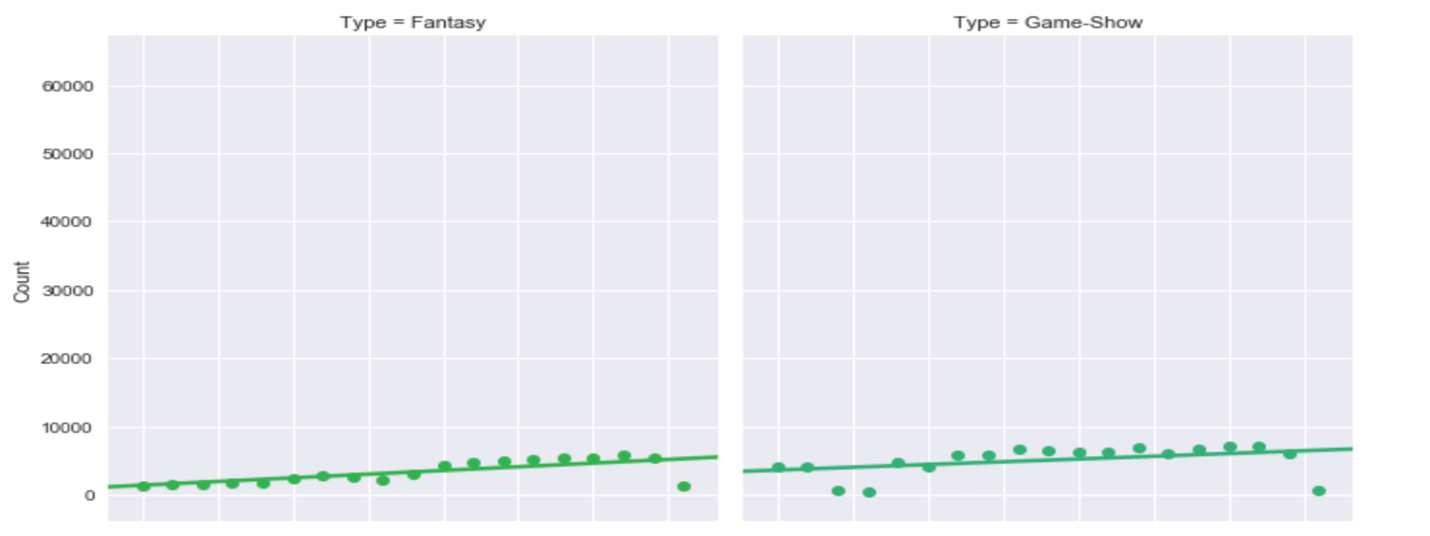
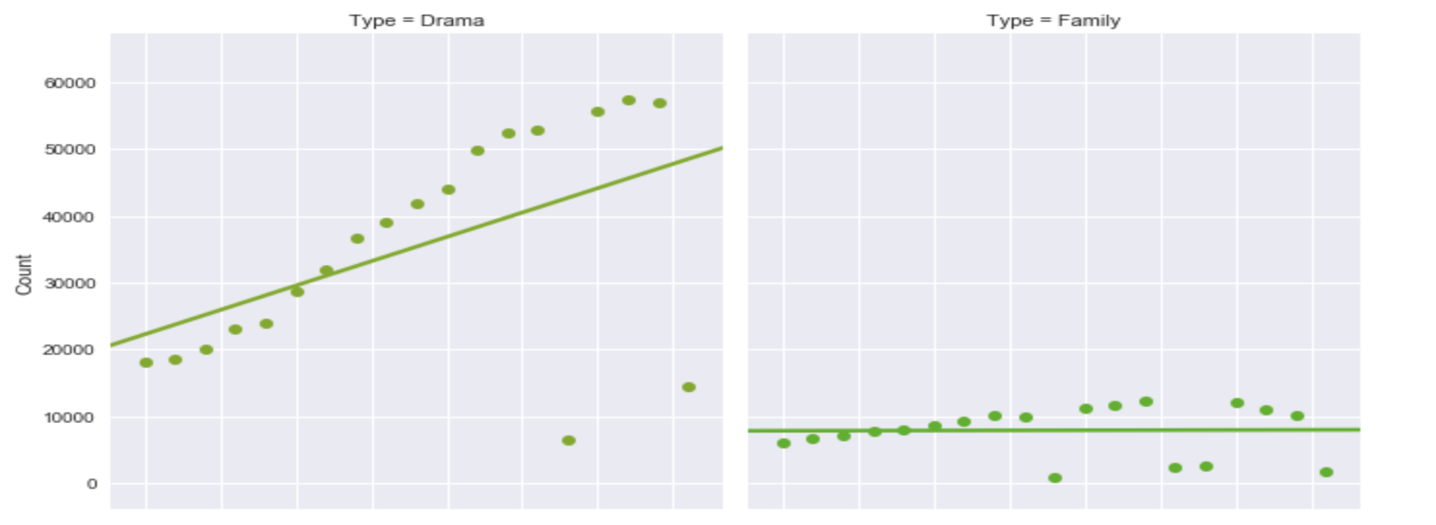
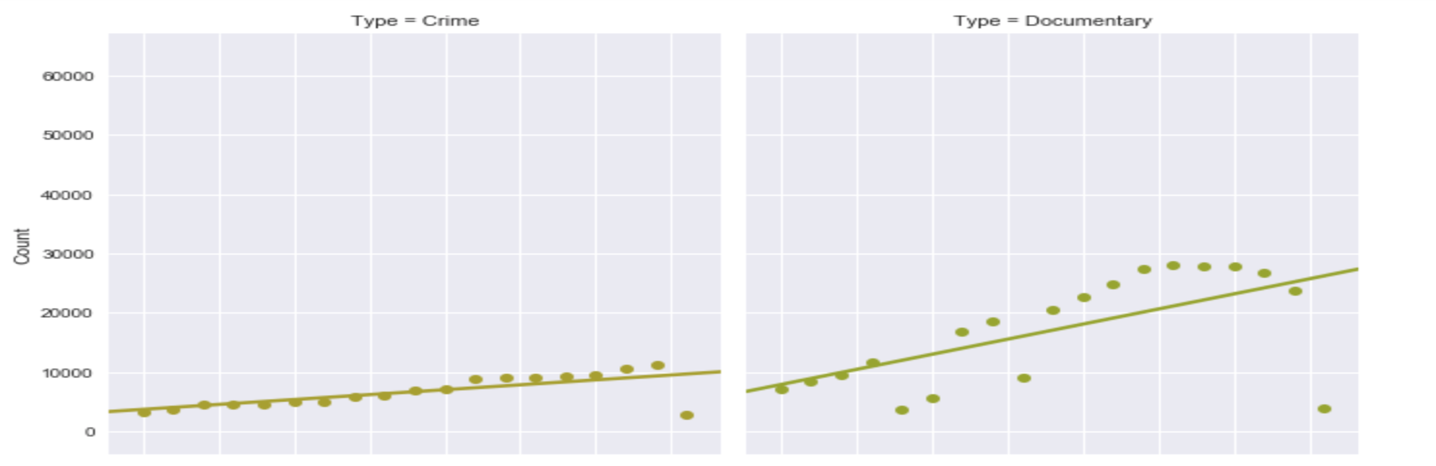
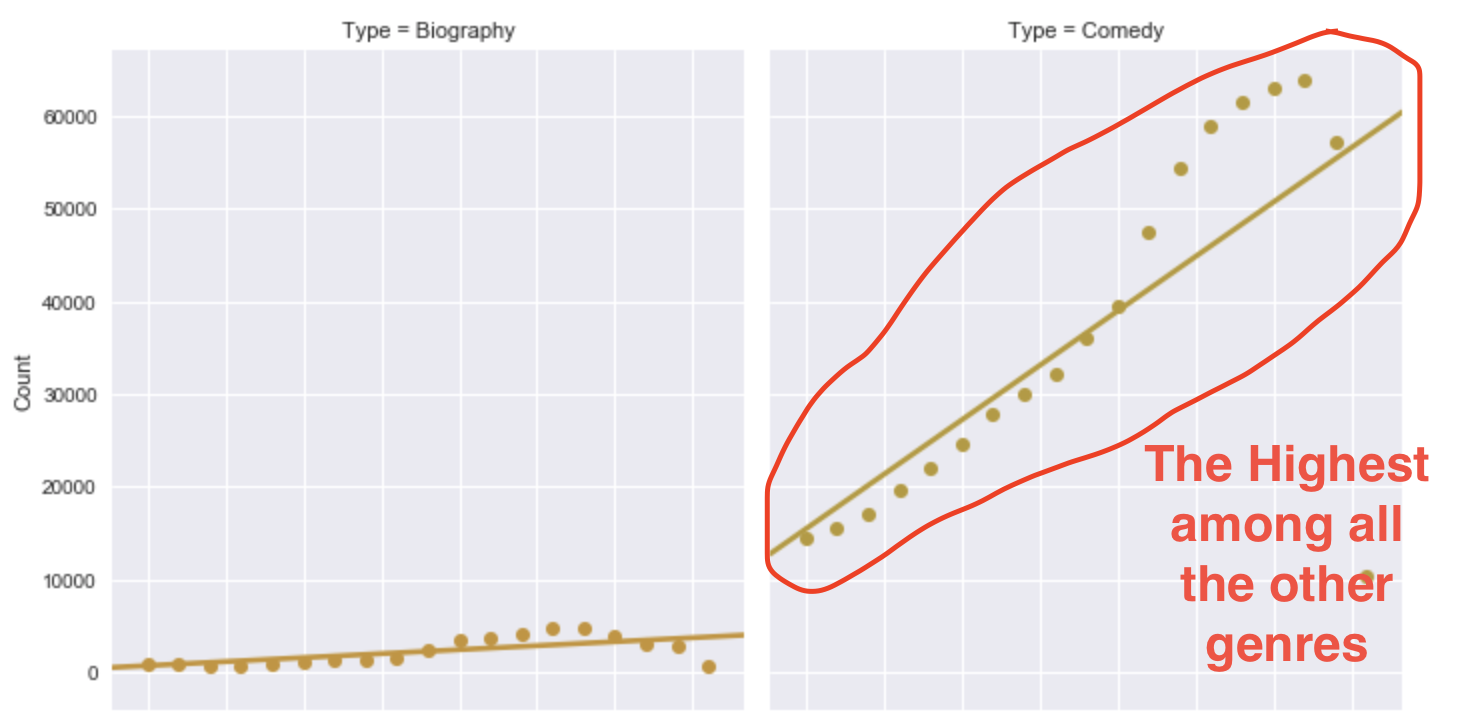
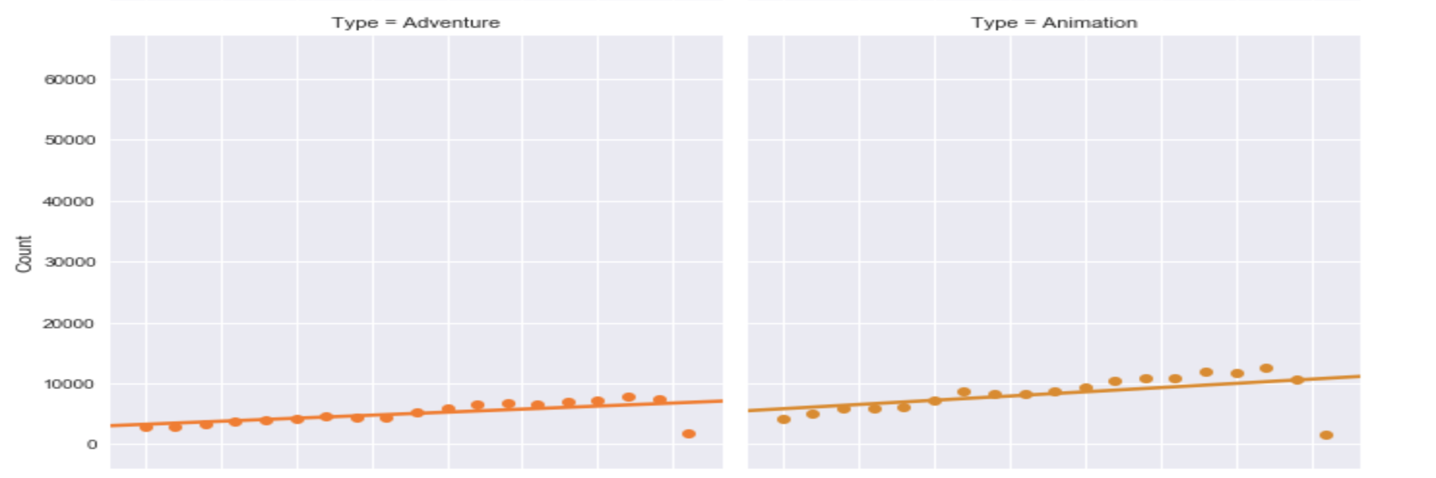
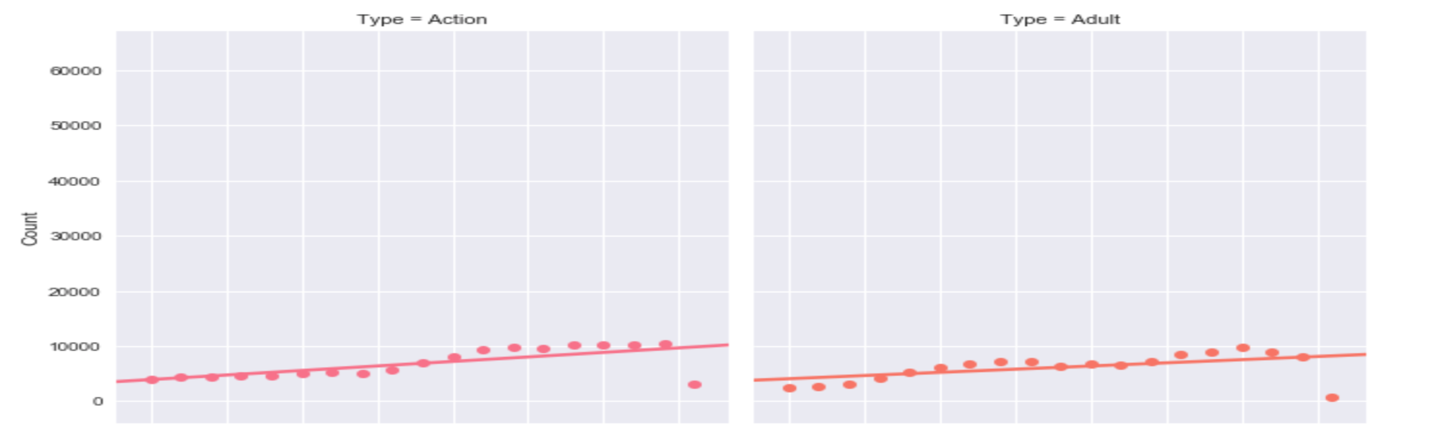
This graph proves that the count of Western movies is almost constant when compared to other genre movies.

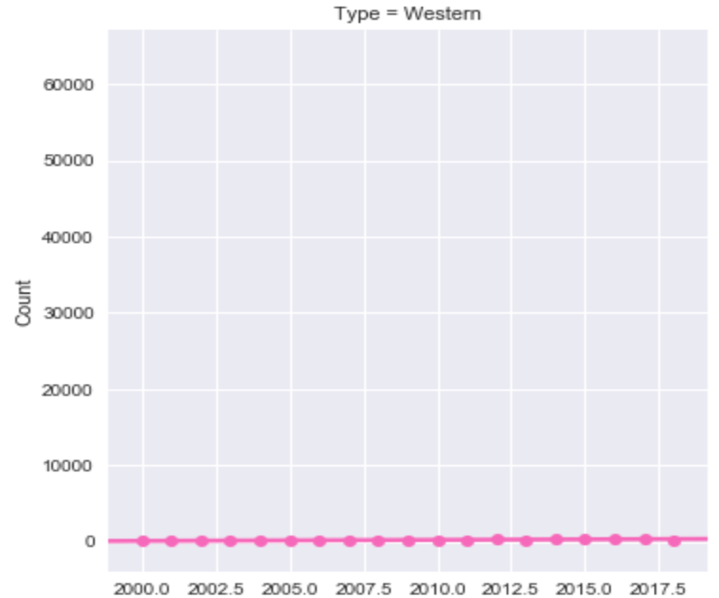


This graph proves that the Western genre movies was the least in all the years when compared to all other genres.

**III.Inference 3**

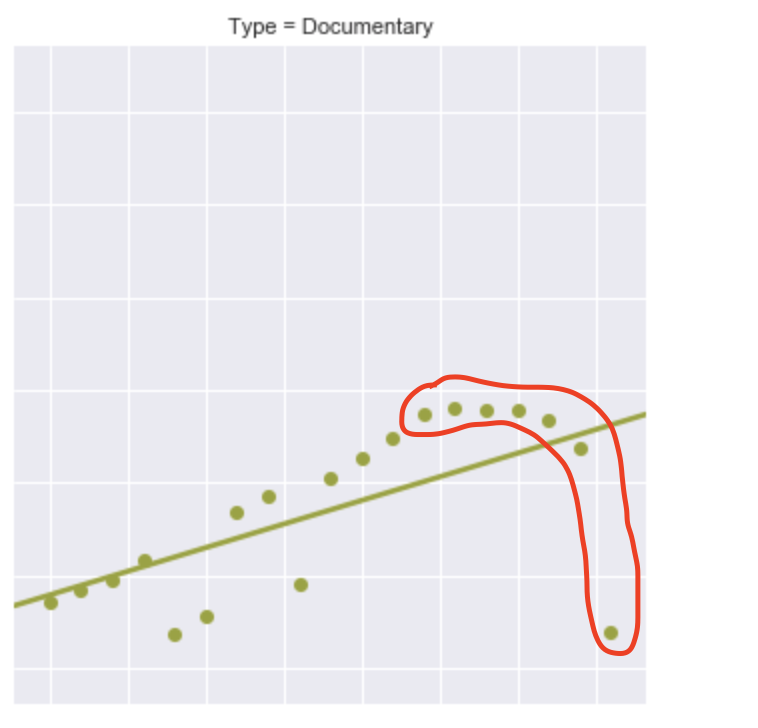
* The rate of increase of the movies belonging to the genre Comedy is the highest among all other genres till 2016.





**IV.Inference 4**

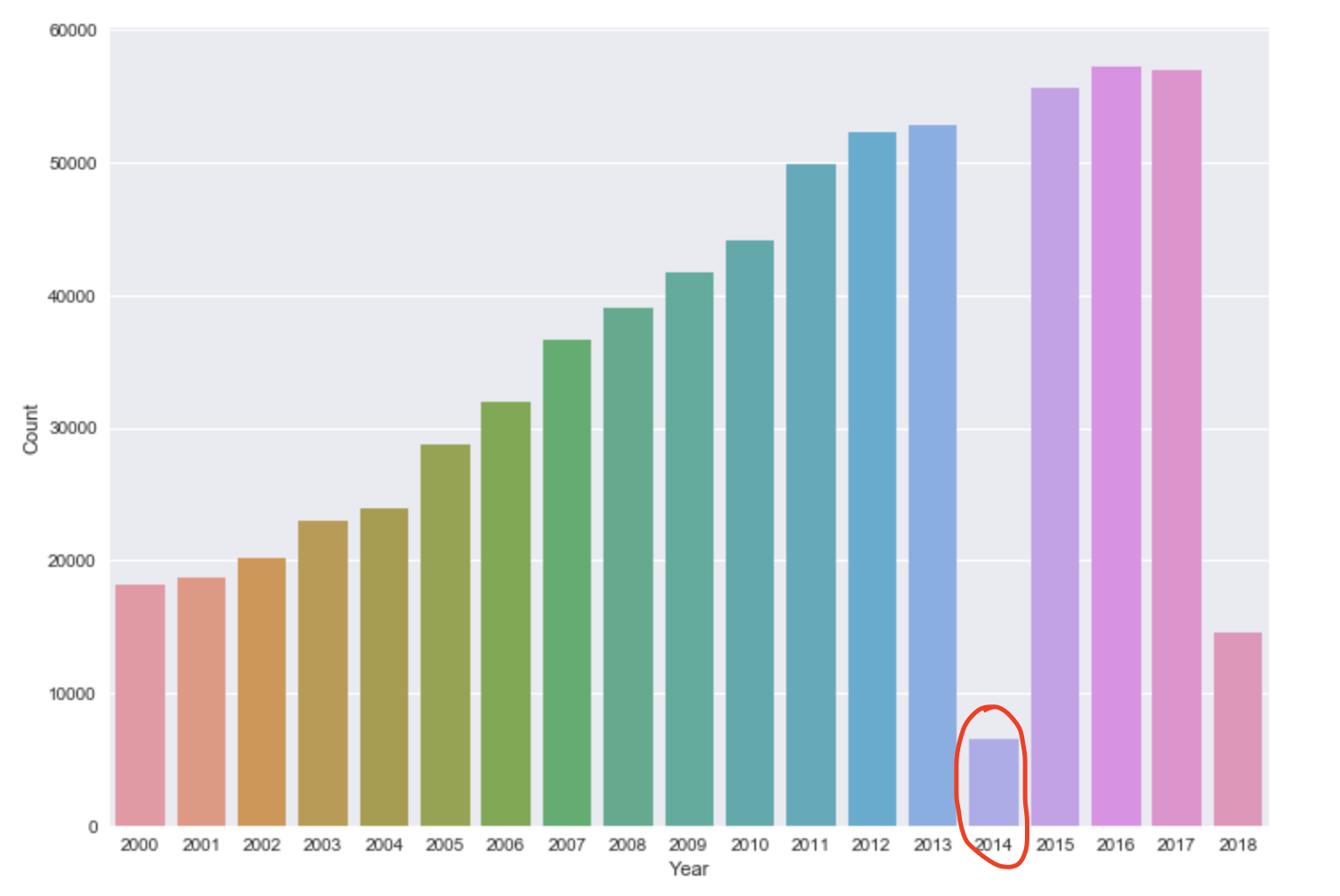
* Documentary is continuously decreasing from the year 2016.If it decreases at this rate, there is high chance that this genre may be discontinued.



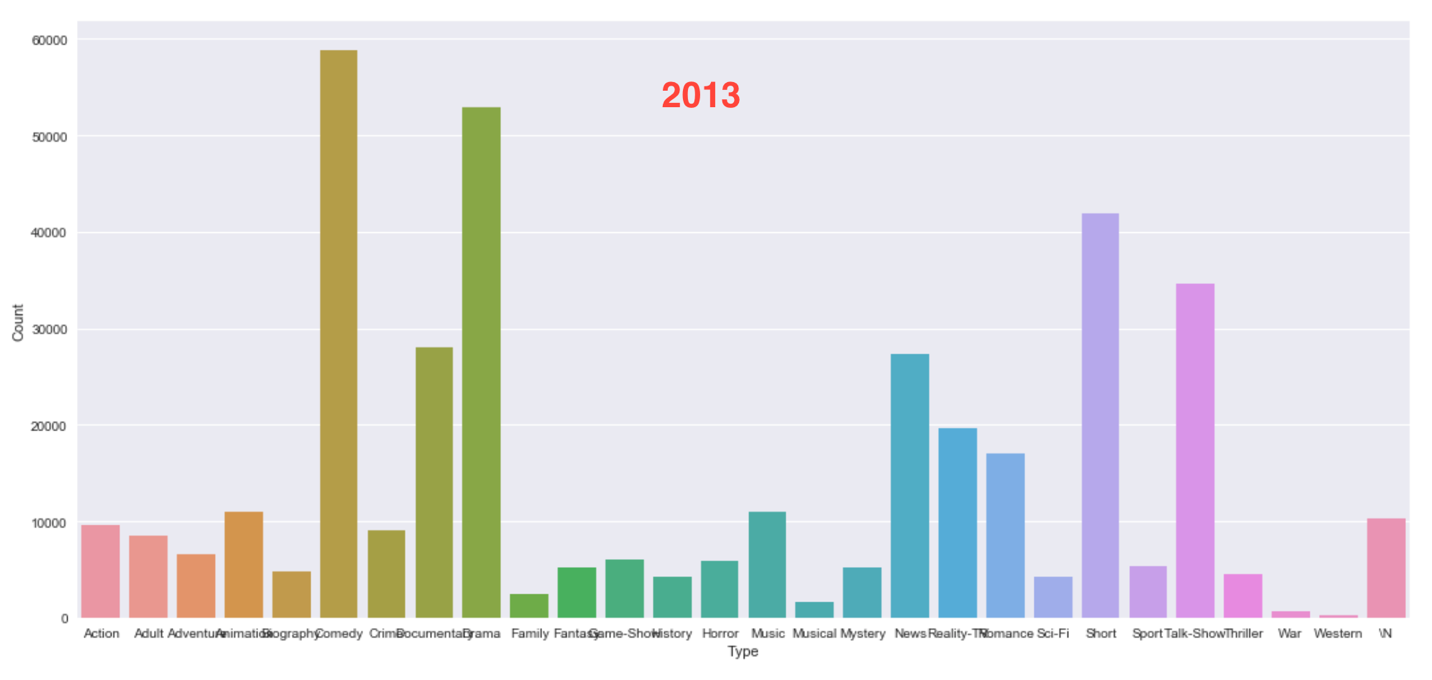
The Graph is plotted on year and count.

**V.Inference 5**

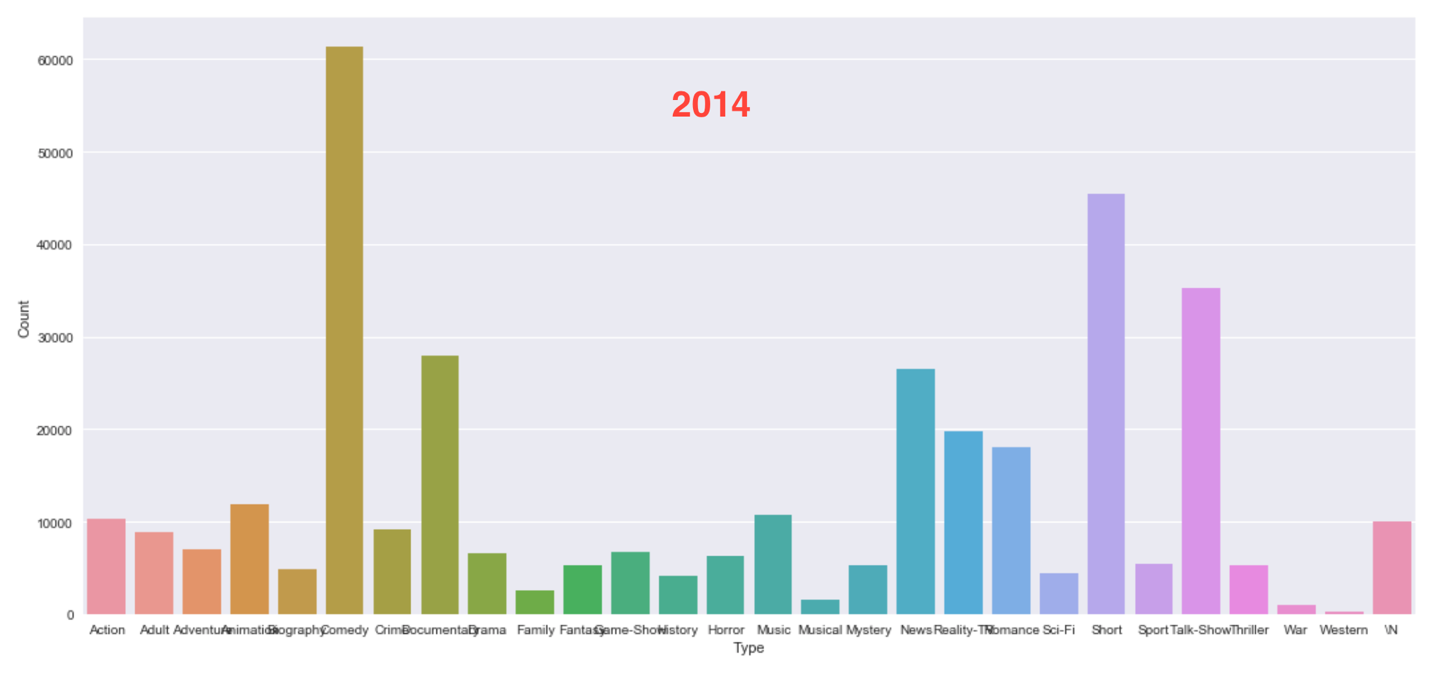
* In 2014, Drama experienced a drastic decline. It was the top second contender in the previous year but lost its count in the year 2014.



This graph is plotted on Year and count. It shows the rapid fall of Drama in 2014.



This graph is plotted on Year and Count. It shows the number of movies of each genre in the year 2013.Drama is the second Contender in the list.



This graph is plotted on Year and Count. It shows the number of movies of each genre in the year 2014.Drama is reduced drastically.

**File Descriptions**

The file we have created is Wordcount.java which has the mapper and the reducer code. We just named it so because we started by modifying the WordCount.java file and did not change the name.

**Division of Labor**

We almost participated equally to complete this project and there is no division of labor in this project. We both wrote the mapper and reducer.

**Logical Errors**

* Improper input to the mapper and reducer. We corrected it after getting the wrong output.
* Improper counting of years in the reducer. This lead to zeros in the output file.
* Got number format exception because of few outliers in the data. Used try catch to exempt them.