

Part 1 - Decade Level Analysis: Once the textual data is processed, I simply wanted to look at the bag of words between the oldest (1915-1924) and the latest decade (2005 – 2014).

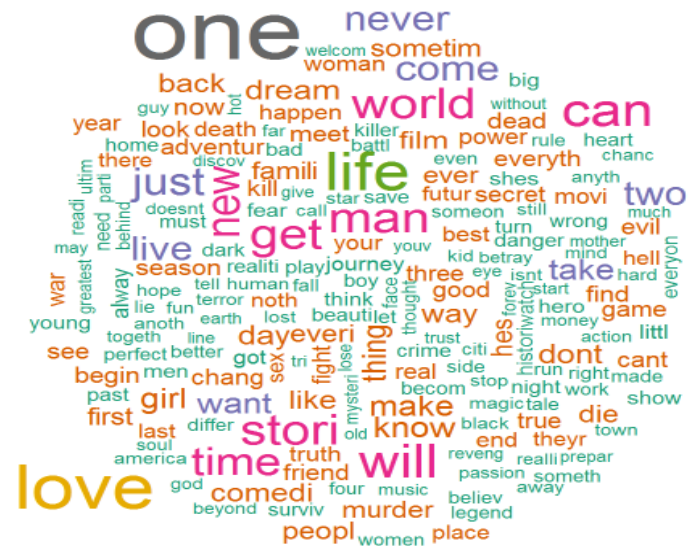


Fig.2. Popular movie tags b/w 2005- 2015

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7
[1,]	"support"	"season"	"one"	"film"	"sometim"	"hes"	"origin"
[2,]	"bride"	"und"	"love"	"world"	"zombi"	"video"	"poster"
[3,]	"tale"	"der"	"life"	"terror"	"documentari"	"adventur"	"cap"
[4,]	"contin"	"amor"	"will"	"see"	"short"	"now"	"print"
[5,]	"local"	"des"	"can"	"ever"	"inspir"	"australia"	"thrill"
[6,]	"method"	"sie"	"world"	"stori"	"communiti"	"game"	"love"
[7,]	"respons"	"una"	"man"	"motion"	"consequ"	"action"	"adventur"
[8,]	"cost"	"serial"	"stori"	"pictur"	"loss"	"cop"	"screen"
[9,]	"thirti"	"pussi"	"get"	"girl"	"cultur"	"ultim"	"action"
[10,]	"debt"	"von"	"time"	"man"	"independ"	"theatric"	"drama"

Topic 1 - has words like tale, due to which I interpreted this topic as Drama.

Topic 2 - has words like serial and season due to which I interpreted this as cult/seasonal movies. For instance these might be the movies which release during holiday seasons.

Topics 3 – This topic has basically words like love and stories due to which I interpreted this topic as romance.

Topic 4 - has words like terror/film. I interpreted this topic as horror/scary movies.

Topic 5 - has words like zombie/documentary. Based on this I interpreted this topic as Sci-Fi

Topic 6 – Based on words like theatrical I names this topic as theatrical.

Topic 7 – This topics has words like thrill/action and adventure. Hence, these movies are mostly exciting. So names this topic as Thriller/action.

Common Trend shown by the broad genres over decades/years (Interpreting Fig.3 and Fig.4):

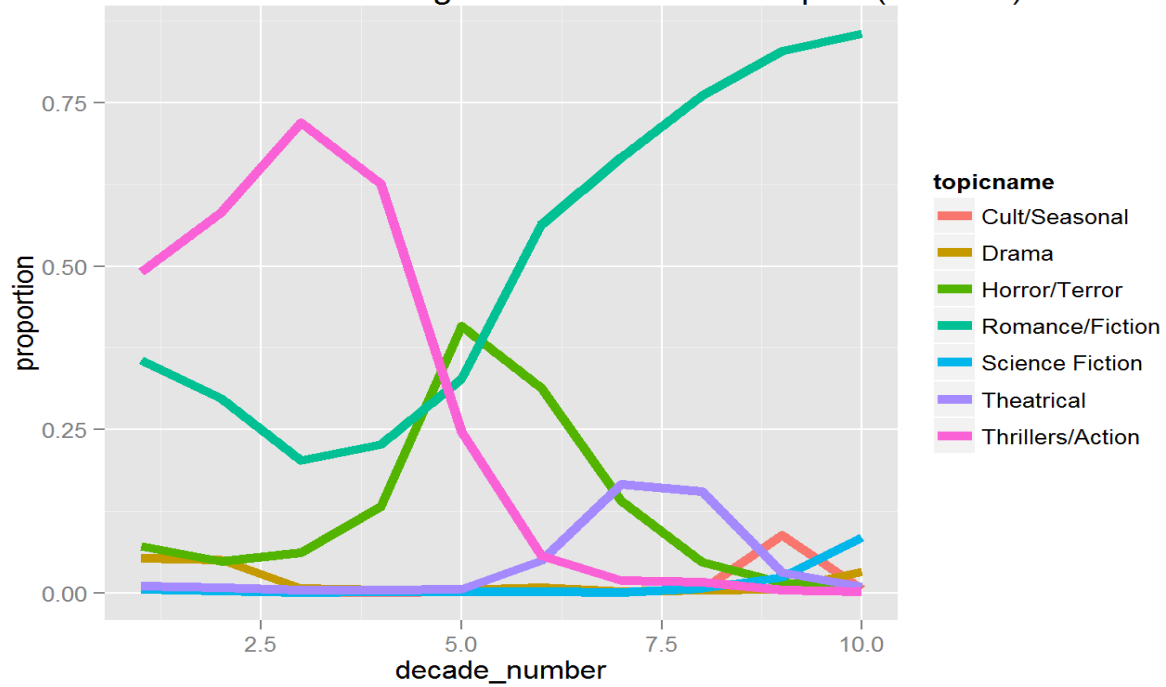
- 1) Fiction/Romance genre is undoubtedly the most popular category. There is a gradual increase in the number of such movies.
- 2) We see that in the initial years there were more number of action (war kind) movies. The number gradually increased for the next 2 decades until 1955. The might be because of the impact of World War I and II, that people are more interested in action/war type movies.
- 3) One interesting observation is that in the 2004 - 2014 decade, we see an abrupt increase in the number of scientific fiction movies which is confirmed with movies like Iron Man etc.
- 4) Also, initially there were very few terror movies. The number increased between the 1950's - 1970's. However, towards the end the number decreased.
- 5) Movies related to drama category are very few. This might be true as the number of drama movies produced are less when compared to other movies.

Part 2 - Year Level Analysis: Even for my year level analysis I have similar topics. So I used the same approach as above to assign topic names. From Fig.4 we see that Trend shown by movies over years is very close to the trend show by movies over decades. Here are a few observations:

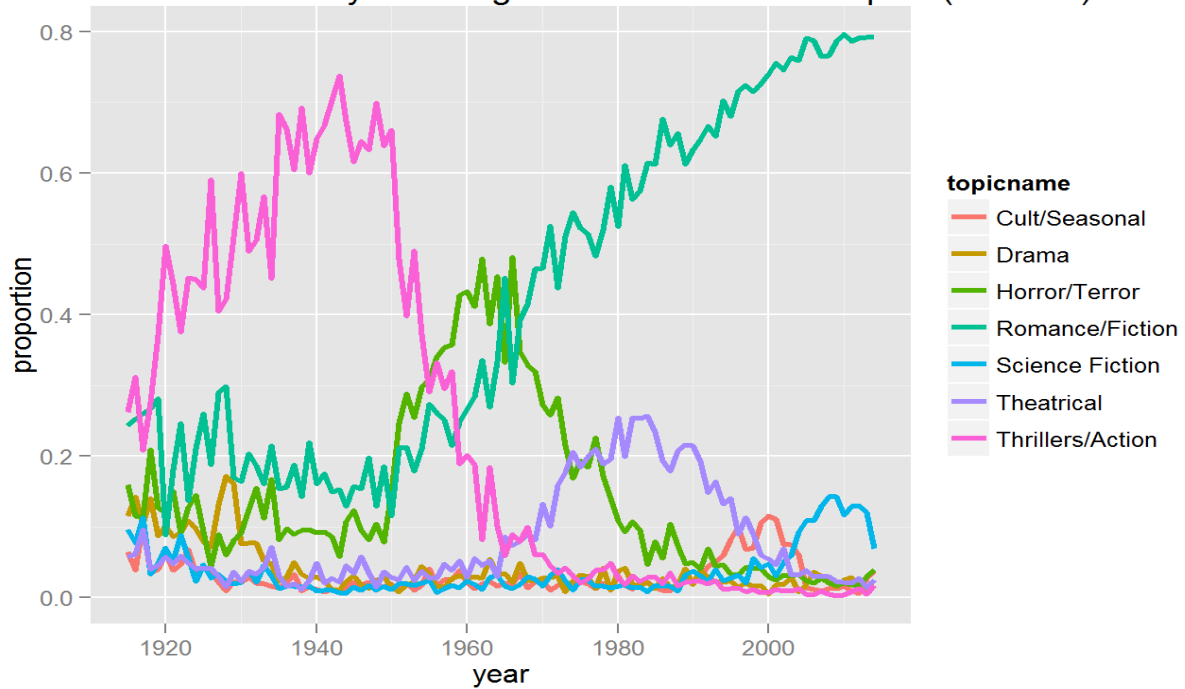
Executive Summary and conclusions

- Romance/Fiction movies is the most dominant category across all genres. This number has been gradually increasing.
- Initially there were not many genres in the type of movies. However, since 1950, we see more diversity in the type of movies produced.
- More number of thriller/action movies were produced in the past years.
- We see an increase in the number of scientific fiction movies in the last 10 years.

Genres Across Decades using Gibbs LDA - Clear Topics (Advised)

*Fig.3. Trend over 10 decades*

e Genres Across each year using Gibbs LDA - Clear Topics (Advised)

*Fig.4. Trend over 100 years*