# Study on the Evolution of Users' Interests on Social Q&A Community with WBTM: A Case Study of Quora

Heming Zhang
Information Management and Information Systems, Central China Normal University, China
hemingzhang@wustl.edu

## Abstract

This paper analyzes the users' interests and their evolution characteristics of on the social Q&A Community, so as to guide the personalized recommendation and advertising on Quora *Film and Television* topics. The data obtained by crawler were divided and cleaned according to the period, and the question and answer text within the period was endowed with user behavior data, and BTM (Biterm Topic Model) was adopted to conduct Topic mining and Topic identification. To improve the effect of mining topics, we propose a new model – Weight Biterm Topic Model (WBTM) by weighting the question answering text with user behavior data. The results show that our model has higher coherence values. Finally, the trend chart of theme evolution is drawn and analyzed by means of thermal map to make personal suggestions on recommendation and advertising of Quora *Film and Television* and other topics.

## 1. Introduction

The social Q&A platform is an emerging knowledge sharing platform developed based on Web 2.0. This platform has no clear organizational structure and allows users to ask questions or answer at any time according to their needs[1]. The core of the social question and answer platform is user participation. The role of users is vague and not clearly defined. They can be information generators or consumers of information. Moreover, with the popularization of the Internet, online Q&A communities have exploded and become important knowledge sharing platforms, such as Quora and StackOverflow in the United States, and Zhihu in China[2]. The online Q&A community provides a platform to create and share knowledge by posting and answering questions. The topics on the platform cover a wide range of topics through questions, opinions, experiences, and comments. By posting or searching questions and collecting answers, users can quickly learn and adopt knowledge related to their areas of concern, most of which are first-hand answers from domain experts[3]. As a result, people will become more and more accustomed to obtaining topics or news they are interested in from online social platforms, which makes user personalized recommendations an important online service. However, on the one hand, the accuracy of topic mining on the social question and answer platform still needs to be improved. On the other hand, there are relatively few studies on the evolution of user interest under a topic on the social question and answer platform.

At present, Quora is a relatively popular social question and answer platform in foreign countries, established in 2009. At the beginning of its establishment, it adopted the invitation registration system, attracting elites from all walks of life. One year later, it will be officially open to the public. Users can log in to Quora with their social networking site accounts to prevent the corresponding content from being searched through search engines. As an online knowledge sharing community, Quora implements the idea of IT flattening and maps people's actual social life to the Internet through online social networks. People co-author content on the Q&A website Quora and find satisfactory answers to their questions. It allows users to editing and answering questions collaboratively, which brings together a large number of questions and answers, providing a good research context for mining user interests and analyzing evolutionary laws in the platform.

Therefore, this paper selects Quora, a popular social question-and-answer platform, and hopes to identify Quora's user interests from the perspective of topics by using the topic mining model. On this basis, this paper considers the combination of auxiliary data such as the number of views in the question and answer in the topic to add weight to the text data. The weighted text data in our experiments show great optimization in the result of topic classification. Finally, this paper analyzes and processes the topic evolution trend under the topic and uses content analysis to analyze the evolution trend, which is very helpful for understanding current user discussion hotspots and further tracking hot topics in the social Q&A platform. It has very good guiding significance for user personalized recommendation, advertisement placement and public opinion supervision.

## 2. Related work

At present, researches on socialized question-and-answer platforms are widely distributed, and related researches are distributed in exploring the growth mode of knowledge quality[3], user interaction mechanism research[4], opinion leader generation[5] and detection[6], opinion extraction[7] and Interest preference mining[2] and many other aspects.

### 2.1 Topics mining related research

In terms of topic mining and interest preference acquisition of social question and answer platforms, Jiang[8] et al. used upvote number to study user interest preferences under the topic of Climate Change. Here, this paper hopes to obtain user interest under a topic and study its evolution trend through topic mining. At present, more researches are carried out based on user text information through methods of identifying and analyzing topics under certain real topics, such as the Latent Dirichlet Allocation (LDA) model proposed by Blei[9] et al. in 2003. For different situations, some scholars have improved the model and tried to apply it to various short text topic mining fields. The classic improved models include ATM[10], Twitter-LDA[11], Labeled-LDA[12] and other models.

However, the LDA model ignores the relevance of text topics between documents. For short texts on social question answering platforms such as Quora, its sparse co-occurrence mode will lead to more serious data sparse problems[13]. A simple and popular way is to aggregate tweets posted by a single user into a document before training LDA, and re-aggregate short text into pseudo-long text[14]. Zhao[11] et al. provided another solution, that is, assuming that a short text document has only one topic, this also causes the model to lose the ability to capture multiple topics in a document.

Therefore, Yan[15] et al. proposed a way to expand the hypothesis, that is, since there is correlation between short text documents, and because the traditional topic model modeling cannot well consider the correlation between short text documents when it comes to the problem of data sparseness, we can consider extending the correlation assumption between documents to the entire corpus space to establish a Biterm Topic Model (BTM). BTM increases the possibility of vocabulary belonging to other topics through assumptions, so that the effect of the BTM model in short text becomes better, and the effect of topic mining in ordinary text is also quite good. Based on the above discussion, in order to more accurately describe the topic heat of the short text document topic, this paper considers BTM as the basis, adopting the combination of the user's text information and the user's views number under the topic to add weight on the vocabulary to improve the effect of topic mining.

## 2.2 Topics evolution related search

In the topic evolution research of the social question and answer platform, Maity[16] et al. used all the content under Quora as the research background, adopted the method of crawling topic tags to obtain topic popularity and used regression analysis to predict the evolution trend; Barua sorted out the relationship between hot topics and analyzed their evolutionary trends by classifying the topics in StackOverflow through the LDA model[17]; Zou[18] et al. used LDA as a topic mining tool to analyze the NFRs (Non-functional requirements) to classify topics and obtain the occurrence rate of each topic to indicate its popularity, thereby making an evolutionary trend and conducting a qualitative analysis. The above research mainly takes all topics in the entire platform as a research perspective. This paper hopes to study the user's interest and evolution trend of social question and answer platforms from the perspective of topics in the platform. In addition, the above research did not adopt a topic mining model that is more suitable for the platform to study user interests and their evolutionary trends.

To sum up, at this stage, there are deficiencies in the selection of topic mining methods and the analysis of evolution characteristics for the evolution of user interest in social question and answer platforms. This paper will fully consider the characteristics of the social question and answer platform to choose a more appropriate method to identify the topic under the topic in the platform. On this basis, the frequency measurement method will be used to evaluate the popularity of the topic evolution in the time dimension, and then use The content analysis method reasonably analyzes its trend and puts forward some suggestions on the personalized recommendation strategy of the social question and answer platform and advertising.

# 3. Methodology

## 3.1 Data collection and data preprocess

In order to analyze the identification and evolution trend of user interest in social question and answer platforms, this paper roughly follows the general paradigm of data collection, data preprocessing, data mining and data analysis to design the research technical route (shown in **Fig. 1**).

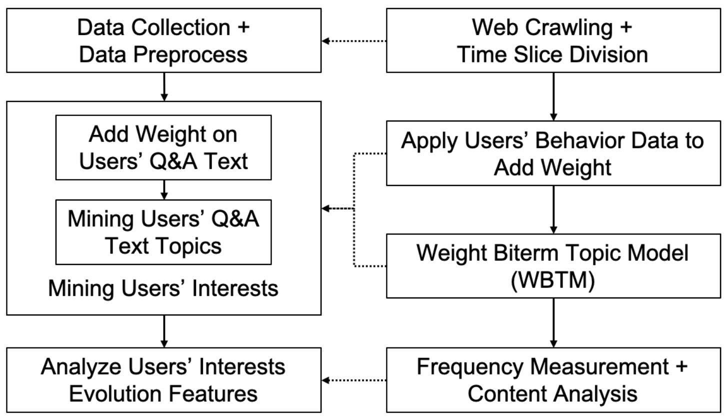


**Figure 1**. Architecture of Research on the Evolution of Users' Interests of Online Q&A Community

First of all, in terms of research content, this paper selects the social question and answer platform Quora as the research scene, but there are many topics under Quora, and selecting the right topic has a profound impact on the research of user interests or topic evolution. The topic that is more suitable for evolution analysis should have the characteristics of faster evolution and more obvious characteristics of evolution, and the topic of *Film and Television* fits these conditions.

Subsequently, with the help of Python web crawler and database manipulation technology, this paper collected and preprocessed all the Q&A texts from December 17, 2018 to April 14, 2019 under the topic of *Film and Television* in Quora. Among them, the preprocessing work is mainly to use the NLTK package in Python to remove stop words and number punctuation in English, and all lowercase. Then, it is necessary to set the length of the time slice based on the evolution law under the topic to prepare for the analysis of user interest evolution.

### 3.2 Proposed model based on original BTM

In terms of user interest topic recognition, this paper combines user behavior data to weight the question and answer text. The original BTM did not consider the influence of user behavior on the topics heat. Therefore, we propose a new model – Weight Biterm Topic Model (WBTM) by replicating text with respect to the number of upvotes to show the weight of certain answer or question (shown in **Fig. 2**).

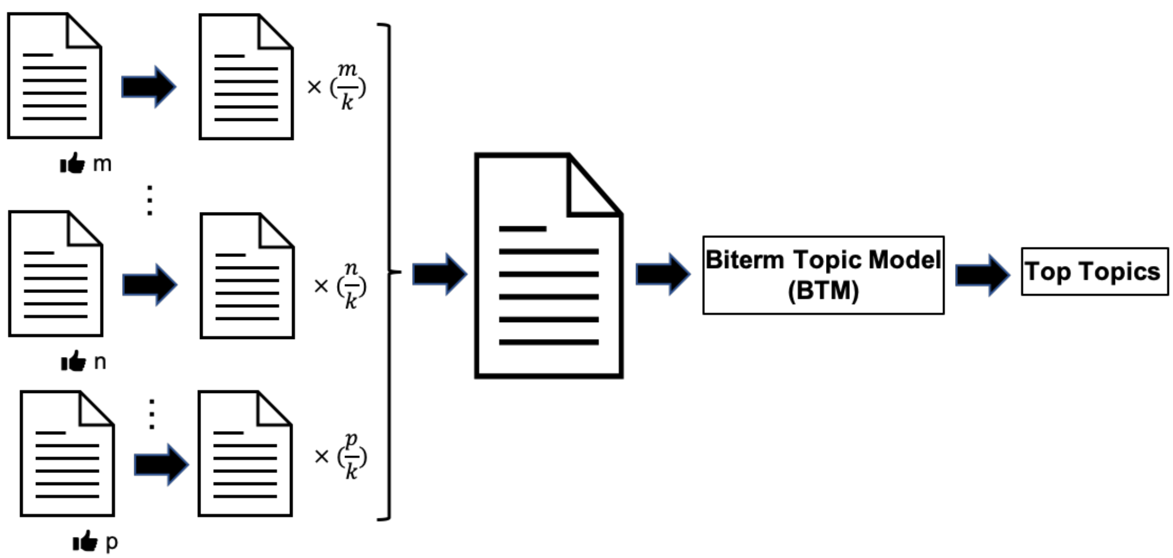


**Figure 2**. Architecture of Adding Weight on Text Collected from Online Q&A Community.

As shown in the figure, $k$ is the threshold of replication, and the text collected from Quora will be replicated once when this text upvoted $k$ times. After combining those text into a whole document with weight on certain text, BTM will be applied to classify them into different topics. Comparing our experiment results on coherence values of topics, the WBTM show great improvement on this respect.

Besides, through experiments, the weighting parameters and the number of divided topics that BTM need to set are determined. The data under each time slice is based on the theme of the *Film and Television* drama. The categories are subject to identification and induction.

Finally, the frequency measurement method was used to identify the user interest topics in each time slice. The topic to which the document belongs is determined according to the number of topics corresponding to the maximum topic probability distribution in each document, and the topic popularity is expressed by counting the occurrence frequency of each topic in all documents. Then, according to the changing trend of the theme in the seventeen weeks, this paper will use heatmap to draw the evolution trend chart, and on this basis, use content analysis to study the characteristics of the theme evolution, and summarize the topic evolution trend of the social question and answer platform , and some suggestions on Quora's personalized recommendations on this topic.

# 4. Experiment Setting

## 4.1 Data preparation

This paper uses Python as the crawler language. After logging into Quora with the help of webdriver in selenium, select the 'All Questions' option under *Film and Television* to enter the list of questions under the topic. Most of the questions are arranged in chronological order from the present to the past. This paper uses Beautifulsoup in bs4 to analyze the links of all questions in the webpage from December 17, 2018 to April 14, 2019 (17 weeks in total) and save them to the MySQL database. Next, by visiting each question link in MySQL, this paper uses the method of simulating login and analyzing the webpage to extract (1) the link of the question; (2) the title of the question; (3) the total number of answers to the question; (4) the last time of following this question; (5) the text information of the answer to this question (pictures, videos and other information are not considered); (6) the time of the answer; (7) the number of views of the answer is saved in the MySQL database.

In this paper, a total of 3849 questions and 6212 answers data were collected through crawlers. After that, a statistical analysis of the number of question answers (**Fig. 3**) showed that most questions only have 0 to 1 answer, and a few questions can get more than 10 Answer. This also shows that there will be a lot of personalized ideas and questions under this theme. But judging from the distribution of questions with 6 to 20 answers, the downward trend is not very obvious, which also shows that there are many common problems under the topic.

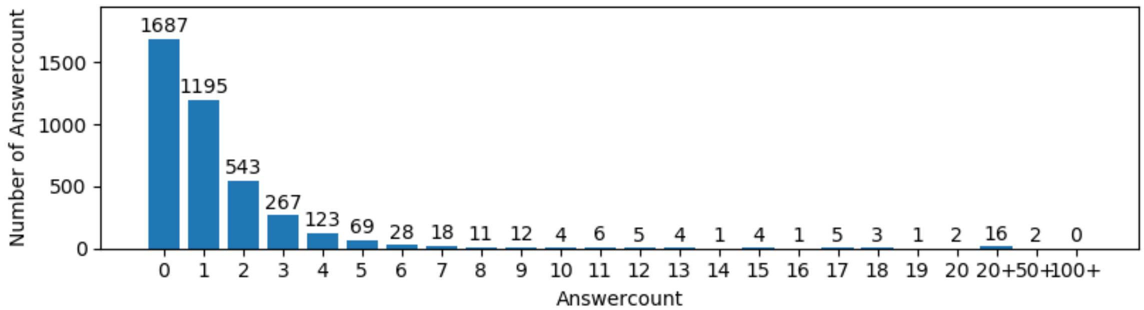


**Figure 3**. The Number of Answer Distribution Under Topic *Film and Television* (Dec 17, 2018 – Apr 14, 2019)

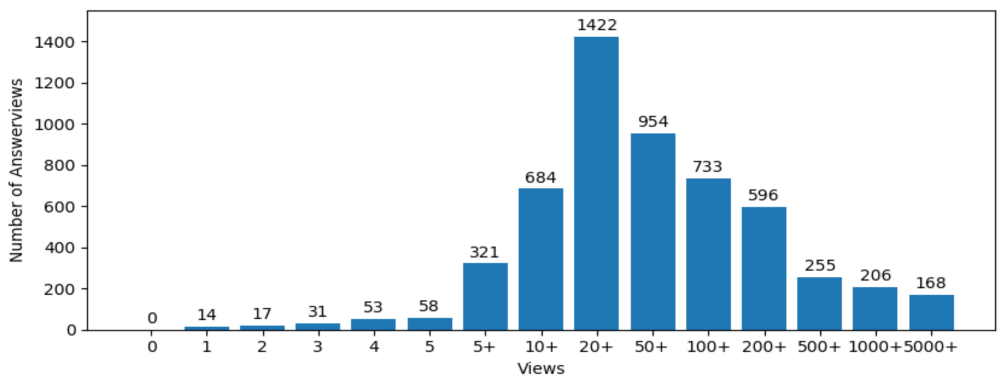


**Figure 4**. The Number of Views in Each Q&A's Distribution Under Topic *Film and Television* (Dec 17, 2018 – Apr 14, 2019)

According to the distribution number of views under each Q&A in the seventeen weeks are in **Figure 4**, after removing the unanswered questions, it can be found that the answer has been increasing from the time the answer is viewed once to the answer is viewed more than 20 times. And the number of answers that have been viewed more than 5 times is growing rapidly. In addition, from the distribution of the histogram, it is roughly in line with the right-skewed distribution, and most of the questions will be viewed 20 to 50 times. In addition, this paper also found that the number of answers that have been viewed more than 1,000 times and 5,000 times is still considerable. This phenomenon indicates that some questions and answers under the topic have attracted more users' attention, and the answers have also been praised by many users, indicating that the topic *Film and Television* has strong reader resonance. Therefore, studying the user's theme preference and theme evolution trend under this topic will have strong help and guiding significance for the improvement of user recommendation effect.

Since the update cycle of British and American dramas and TV shows is one week, this paper sets the time section to one week to analyze the evolution of user interests. Therefore, from December 17, 2018 to April 14, 2019, it is divided into seventeen weeks. The distribution of the number of weekly questions is shown in Figure 5, which shows that the number of questions in a week is 210. Floating up and down, the overall data distribution is relatively stable.

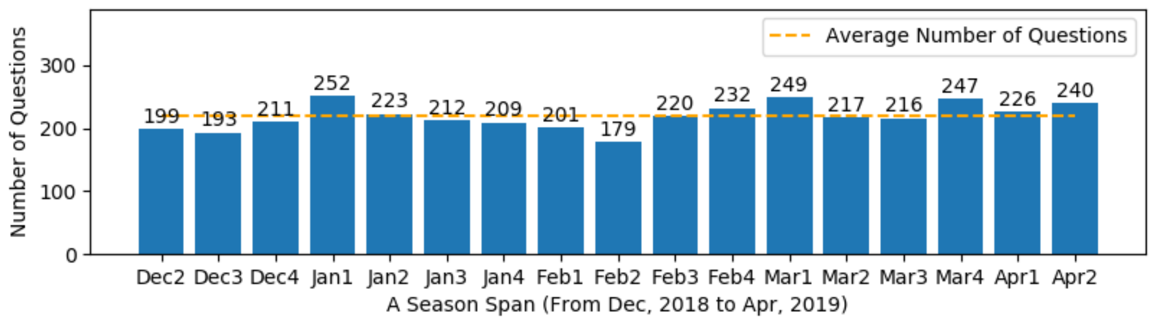


**Figure 5.** The Q&A Distribution in Different Time Section (Dec 17, 2018 – Apr 14, 2019)

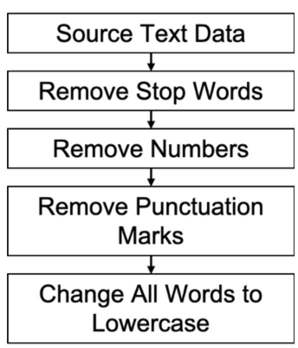


**Figure 6.** Data Cleaning Process

Next, the data cleaning process mainly uses the **nltk** package in python to clear the stop words, numbers and punctuation marks in English by using the question and answer data divided by week, and change them to lowercase (as shown in Figure 6) , and save the organized documents for use in topic mining of next procedure.

## 4.1 Data mining experiment

When using topic mining models such as BTM to identify topic interest in text data, how to set the number of classified topics is usually a tricky problem. Getting feedback through experimental results and debugging in combination with actual conditions is a powerful way. In addition, considering that the number of views of the text content should be able to reflect the importance of the text, and many answers under the *Film and Television* topic have been viewed more than 1,000 times (shown in **Fig. 4** above). Therefore, this paper considers re-editing and evaluating the text information.

Here, the weighting method of text data in this paper is expanded by answering the number of viewed, that is, first set a threshold value, and the content of the answer under the topic is viewed more than once, and the browsing value will be copied once in the document. Regarding the weight parameters of the weighted text, it is also necessary to debug through the result feedback of the experiment, and judge whether the BTM effect of the weighted text is significantly improved.

How to select the appropriate number of topics in BTM is a key factor, and its value will have a certain impact on the coherence value of the final topic mining and the actual mining effect. Here, the number of iterations of the preliminary tentative model is 20. At the same time, in order to compare the training effect of using unweighted text data and weighted text data and to pre-tune the parameter value of the number of topics in BTM, this experiment feedbacks will be inspected through figures. Therefore, questions and answers data over 50, 100, 1000, 2000, 5000 times will be copied once in different experiment setting.

Using the question and answer data from April 8, 2019 to April 14, 2019 as the training sample, the coherence value of each topic is obtained with the help of the coherence calculation method inside BTM. After averaging the coherence value of each topic, the preliminary experimental results will be shown in Figure 7.

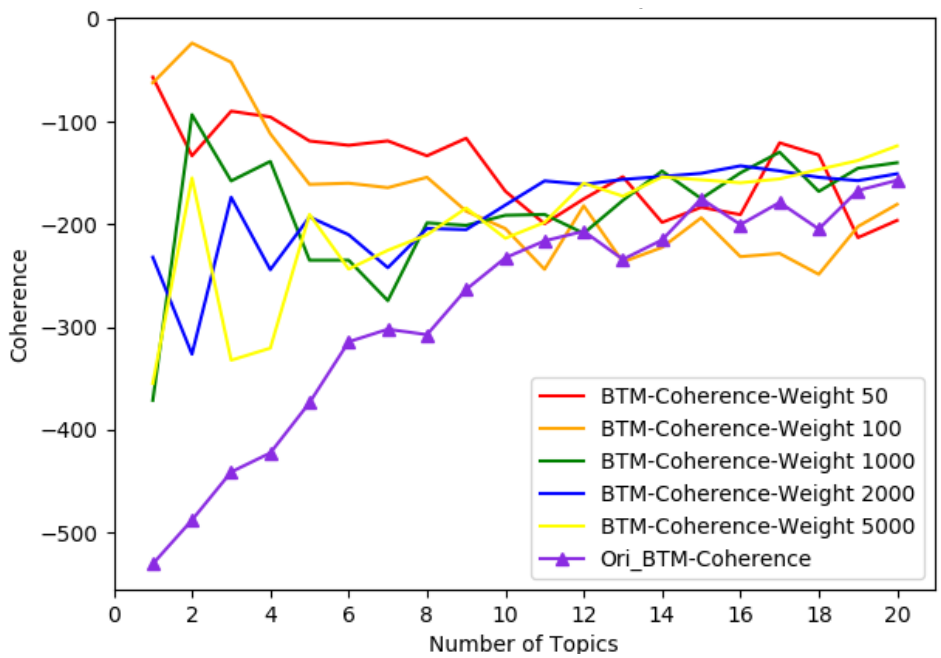


**Figure 7.** Example Comparison of Weighted and Unweighted Text Topic Mining Coherence Values Under Different Number of Topics

Through experimental comparison (as shown in **Fig. 7**), it is found that the coherence value of topic classification for unweighted text is not very high at the beginning, and it gradually increases as the number of topics increases. It can also be found here that as the weight value increases, the coherence value of the weighted text will approach the unweighted text. The reason for this is that when the weight threshold is set higher, the number of weighted texts will be less, closing to the case of unweighted text. But what is surprising is that its coherence value will be better than the value of unweighted text when the number of topics is less than 10 regardless of the weight setting.

Taking into account the reality that the number of questions for *Film and Television* topics in short text documents such as Quora will be about 200 and the number of questions and answers will be about 700 in the next week, the model here should be set within 10.

As to the threshold value of replicating text to add weight, this paper selects the number of topic categories at 6, and sets the replicating threshold value from 50 to 5000 with step size of 50, that is, from the answer being watched once and then being copied once to the answer being watched more than 5000 times and then being copied again (was shown in **Fig. 2**). From the experimental results in Figure 8, it can be seen that the coherence value of WBTM shows a downward trend after being fitted by a linear function, and the weight threshold falls faster in the interval of 50-500, but the overall effect is better than the effect of unweighted text. Here, we set the weight threshold $k$ as 100.

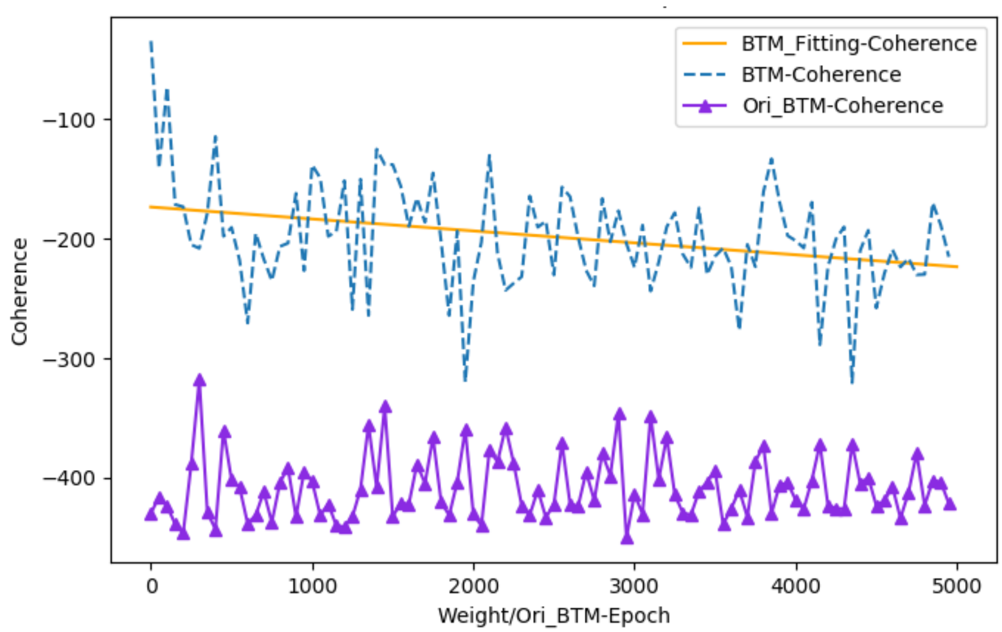


**Figure 8.** Example Comparison of Weighted and Unweighted Text Topic Mining Coherence Values Under Certain Number of Topics

## 5. Experiment results and analysis

With the help of BTM after improving the text content, this paper classifies the topics of the seventeen-week data and obtains the top 10 hot words in each topic ranked according to the probability value. Through the common sense of British and American dramas and TV program knowledge and the matching of entries in Wikipedia, this paper combines the classification method of film and television dramas, and divides the themes under Film and Television into *Feature; Affectional; War; Comedy; SciFi; Animation; Thriller; Criminal; Plot; Actor; TV Show*, totally 12 themes. After that, you can add corresponding tags to the weekly topic classification results. Here, this paper selects the topic classification results from April 8, 2019 to April 14, 2019 (shown in **Table 1**) to demonstrate the classification effect (detailed in **Appendix 1**).

**Table 1.** Result of Topic Mining (Apr 8, 2019 – Apr 14, 2019)

| Topic No. (Topic Tags) | Topic1 (Comedy) | | Topic2 (Criminal) | | Topic3 (SciFi & The Avengers) | | Topic4 (SciFi) | |
|---|---|---|---|---|---|---|---|---|
| Topics Top Words & Words Weight Value | episode | 0.04 | people | 0.045 | thanks | 0.015 | standard | 0.028 |
| | guy | 0.039 | career | 0.044 | **world** | 0.014 | **fiction** | 0.028 |
| | **family** | 0.038 | episode | 0.042 | frank | 0.014 | examine | 0.028 |
| | mantain | 0.035 | city | 0.041 | **paul** | 0.013 | structure | 0.028 |
| | quo | 0.028 | **weapon** | 0.041 | harris | 0.013 | informative | 0.028 |
| | **adult** | 0.028 | **cop** | 0.041 | invisible | 0.013 | drama | 0.028 |
| | change | 0.026 | **firing** | 0.041 | **hemsworth** | 0.013 | cast | 0.028 |
| | phenomenon | 0.026 | **shooting** | 0.041 | bana | 0.013 | **chikills** | 0.017 |
| | **simpson** | 0.026 | **criminal** | 0.041 | liam | 0.013 | **jessica** | 0.017 |
| | favourite | 0.026 | **police** | 0.041 | jordan | 0.013 | **hero** | 0.017 |
| Topics Frequency | 55 | | 182 | | 262 | | 95 | |
| Topics Proportion | 9% | | 31% | | 44% | | 16% | |

For the case where the topic belongs to multiple subjects, it is necessary to select the more popular topic words and set them as the keywords under the topic by using the vocabulary popularity of the topic under the theme. When encountering special topics such as artist affair, film and television debuts and focusing on certain classic movies, hot topics will be added after the corresponding topics, just like Topic3 in **Table 1**, but may not be in the above 12 topics Label.

When this paper analyzes the evolution of weekly topics, it also needs to merge and simplify similar topics. In **Table 1**, in the experimental environment, WBTM separates the two SciFi topics because one topic is simply talking about some SciFi movies, and the other topic is discussing SciFi because *The Avengers* is about to be released, which set off a small climax of discussion. But in order to facilitate the drawing and theme evolution analysis, this paper chooses to merge the two topics into *SciFi & The Avengers* here. Similar processing will also exist in the labeling work of the topic classification in the previous weeks.

Finally, with the help of the frequency measurement method, the appearance frequency statistics of each topic in **Table 1** can be obtained, and the proportion of each topic in all documents can be obtained. Here, this paper uses this ratio to indicate the popularity of each topic in all discussions this week. The value will be reflected in the topic heat proportion in Table 1, and the value and the color depth in the heat map are shown in **Figure 9**. Therefore, this paper draws a heat map of the theme evolution trend from December 17, 2018 to April 14, 2019 for a total of seventeen weeks (shown in **Fig. 9**), which represents the evolution of each topic in the time dimension during each week, and each row represents the evolution of the popularity of each topic in the seventeen weeks.

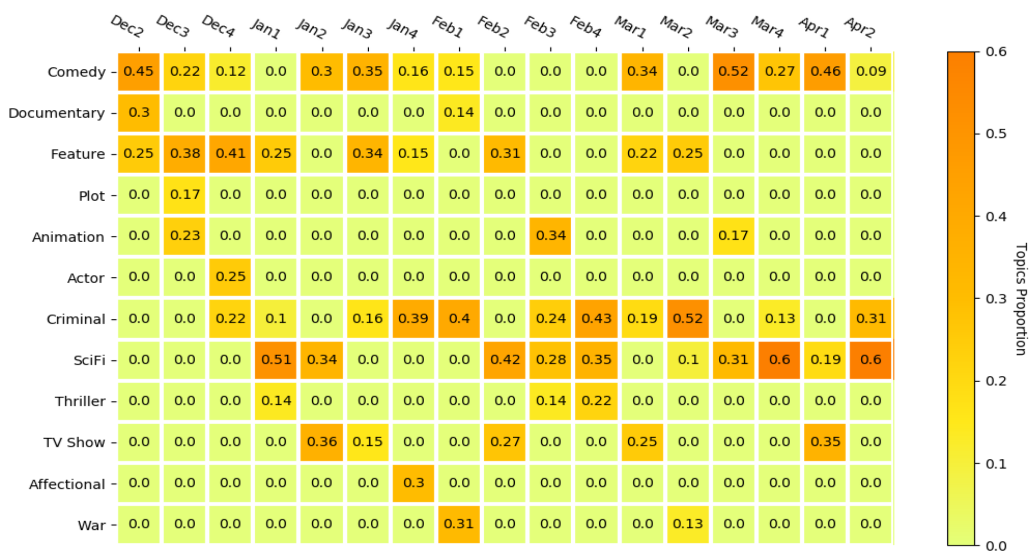


**Figure 9.** Heatmap of Topics Evolution Trend (Dec 17, 2018 – Apr 14, 2019),

Based on the analysis of the evolution of the topics in **Figure 10** from December 17, 2018 to April 14, 2019, this paper finds the following rules for the Film and Television topics in Quora:

The topic is updated quickly. Although the subject matter may still belong to one category, the specific content will change greatly. The general life cycle is about one to two weeks. This paper believes that the reason for this is British and American dramas and TV shows. The update frequency is fast, and the occurrence of events has a certain chance. For example, some news of certain artists may generate a hot topic of the week.

It takes a certain amount of time for a topic to become a hot topic, usually about a week, and then it will be replaced by a new topic; in addition, there will be a certain delay from the occurrence of a hot event in reality to becoming a hot topic, such as in *Deadpool*. After the PG13 version was released in North America on December 12, 2018, SciFi related to *Deadpool* became a hot topic of discussion two weeks later, which happened to be in the time period before the release of *Deadpool* in North America and before the release in mainland China. This paper believes that the reason for this is that it takes time for users to discuss topics. The hot topics discussed under the current topic may have become hot topics before this time period. This also shows that if you need to recommend content to users. At this time, it is necessary to combine some hot issues in real life to predict in advance.

Among the 17-week evolution themes, *Comedy, SciFi and Criminal* topics appeared more frequently, and they were the mainstream of discussion under this topic. Therefore, the system can pay attention to Film and Television topics. The users of the corresponding topic promote relevant content. Because there are not many sitcoms currently being updated, most Comedy content may be biased towards classic comedies and TV shows. In this way, the content recommendation for the *Comedy* topic is very simple. If the content is selected from the classic comedy content of the past, the probability of being liked will increase accordingly. The topic of *Criminal* will often discuss issues such as film and television situations and plots. The topics are more likely to evolve into topics such as *Thriller* and *Character* (movie characters) in the next week. *SciFi (science fiction)* topics will discuss some new topics other than classic sci-fi movies, because there are a lot of movies and TV works currently being updated, when recommending user content, you need to consider the current hot works being updated and the new products that will be shown to make recommendations.

Here, although this paper uses the *Film and Television* topic in Quora as an entry point to study the evolution of user interest, the experience here can be extended to similar topics, such as *Movies, Television Series* and other film and television drama topics. The personalized recommendation and advertising of drama topics have a good reference.

# 6. Conclusion

For the research on the evolution of user interest in social question-and-answer platforms, this paper uses Quora as a study case and uses topic of *Film and Television* as an entry point. Through Python web crawler technology, database manipulation technology and data cleaning toolkit, the time section is set to one week according to the update cycle of British and American dramas and TV shows. Based on the paper's understanding and discussion of the LDA model and BTM, this paper finally chooses WBTM as the topic mining method and weights the text data based on the number of user's answers viewed as a parameter. The experimental results prove that the effect of weighted text is better than that of weighted text. It is better to determine the number of WBTM topics at 3-6. The parameter of weighted text is set to be copied once when the answer is viewed more than 100 times. This is also an important improvement and innovation of this paper. Afterwards, through the topic mining of the weekly empowered text data for seventeen weeks, this paper finally obtained the weekly user interest topics and hot vocabulary, and the topics were summarized and classified.

Through the top words under each topic, the discussion content of the topic can be evaluated and classified. After that, based on the occurrence frequency of the topic in each question and answer document in each week, the popularity of the topic can be identified, and the corresponding value can be obtained. The heat of the topics discussed each week is indicated by the color shades through the heat map, and then a topic evolution trend chart can be obtained by chronologically from the past to the future according to the horizontal axis.

Finally, this paper combines the theme evolution trend from December 17, 2018 to April 14, 2019, and qualitatively analyzes the evolution trend of Film and Television topic and believes that the evolution of Film and Television topic is relatively fast. In addition, through a specific analysis of the 17-week data, this paper also believes that there will be a certain delay in the generation of topic hotspots from reality to the platform, and the main topics of the discussion are *Comedy*, *SciFi* and *Criminal* subject matter. As a result, the system can promote relevant content to users who follow the corresponding topic under the Film and Television topic. The content recommendation for *Comedy* topics can select the pushed content from the classic comedy content of the past, and its probability of being liked will increase accordingly. The topic of *Criminal* will often discuss issues such as film and television situations and plots. The topics are more likely to evolve into topics such as *Thriller* and *Character* (movie characters) in the next week. When recommending user content on *SciFi* topics, you can consider currently updating hot works and upcoming new products for recommendation.

In summary, this paper conducts topic mining on user interests under the topic of *Film and Television* in Quora and optimizes the topic recognition effect of BTM by weighting text data. In the subsequent analysis of the evolution trend of user interests, through qualitative analysis of the topic, some suggestions were put forward for Quora to make personalized recommendations and advertising on the topic. However, this paper did not perform quantitative modeling and analysis on the subsequent theme evolution trends to obtain quantifiable laws, which will also become the next important breakthrough direction in the evolution of user interests.

**Reference**

[1] F. M. Harper, D. Raban, S. Rafaeli, and J. A. Konstan, “Predictors of answer quality in online Q&A sites,” *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 865–874, 2008, doi: 10.1145/1357054.1357191.

[2] G. Liu, Y. Wei, and F. Li, “Understanding consumer preferences-eliciting topics from online Q&A community,” *Proc. Int. Conf. Electron. Bus.*, vol. 2018-Decem, pp. 690–698, 2018.

[3] G. Wang, K. Gill, M. Mohanlal, H. Zheng, and B. Y. Zhao, “Wisdom in the social crowd,” 2013, pp. 1341–1352, doi: 10.1145/2488388.2488506.

[4] R. Rughiniş, A. P. Marinescu-Nenciu, Ş. Matei, and C. Rughiş, “Computer-supported collaborative questioning. Regimes of online sociality on Quora,” *Iber. Conf. Inf. Syst. Technol. Cist.*, pp. 222–227, 2014, doi: 10.1109/CISTI.2014.6876946.

[5] S. A. Paul, L. Hong, and E. H. Chi, “Who is Authoritative? Understanding Reputation Mechanisms in Quora,” no. 2010, 2012, [Online]. Available: http://arxiv.org/abs/1204.3724.

[6] S. Patil and K. Lee, “Detecting experts on Quora: by their activity, quality of answers, linguistic characteristics and temporal behaviors,” *Soc. Netw. Anal. Min.*, vol. 6, no. 1, pp. 1–11, 2016, doi: 10.1007/s13278-015-0313-x.

[7] A. Kumar, S. Praveen, N. Goel, and K. Sanwal, “Opinion extraction from quora using user-

biased sentiment analysis," in *Advances in Intelligent Systems and Computing*, 2018, vol. 672, pp. 219–228, doi: 10.1007/978-981-10-7512-4_22.

[8] H. Jiang, M. Qiang, D. Zhang, Q. Wen, B. Xia, and N. An, "Climate change communication in an online Q&A community: A case study of Quora," *Sustain.*, vol. 10, no. 5, 2018, doi: 10.3390/su10051509.

[9] D. M. Blei, A. Y. Ng, and J. B. Edu, "Latent Dirichlet Allocation Michael I. Jordan," 2003. Accessed: Sep. 05, 2020. [Online]. Available: http://www.jmlr.org/papers/v3/blei03a.

[10] M. Rosen-Zvi, T. Griffiths, M. Steyvers, and P. Smyth, "The Author-Topic Model for Authors and Documents." Accessed: Sep. 05, 2020. [Online]. Available: https://arxiv.org/abs/1207.4169.

[11] W. X. Zhao *et al.*, "Comparing twitter and traditional media using topic models," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2011, vol. 6611 LNCS, pp. 338–349, doi: 10.1007/978-3-642-20161-5_34.

[12] D. Ramage, D. Hall, R. Nallapati, and C. D. Manning, "Labeled LDA: A supervised topic model for credit attribution in multi-labeled corpora," 2009. Accessed: Sep. 05, 2020. [Online]. Available: https://www.aclweb.org/anthology/D09-1026.pdf.

[13] L. Hong and B. D. Davison, "Empirical study of topic modeling in Twitter," in *SOMA 2010 - Proceedings of the 1st Workshop on Social Media Analytics*, 2010, pp. 80–88, doi: 10.1145/1964858.1964870.

[14] J. Weng, E. P. Lim, J. Jiang, and Q. He, "TwitterRank: Finding topic-sensitive influential twitterers," in *WSDM 2010 - Proceedings of the 3rd ACM International Conference on Web Search and Data Mining*, 2010, pp. 261–270, doi: 10.1145/1718487.1718520.

[15] X. Yan, J. Guo, Y. Lan, and X. Cheng, "A biterm topic model for short texts," in *WWW 2013 - Proceedings of the 22nd International Conference on World Wide Web*, 2013, pp. 1445–1455, doi: 10.1145/2488388.2488514.

[16] S. Maity, J. Sahni, A. M.- ICWSM, and undefined 2015, "Analysis and Prediction of Question Topic Popularity in Community Q&A Sites: A Case Study of Quora."

[17] A. Barua, S. W. Thomas, A. E. Hassan, A. Barua, S. W. Thomas, and A. E. Hassan, "What are developers talking about? An analysis of topics and trends in Stack Overflow," *Springer*, vol. 19, pp. 619–654, 2014, doi: 10.1007/s10664-012-9231-y.

[18] J. Zou, L. Xu, W. Guo, M. Yan, … D. Y.-2015 I. 12th, and undefined 2015, “Which non-functional requirements do developers focus on? an empirical study on stack overflow using topic analysis,” *ieeexplore.ieee.org*, Accessed: Sep. 05, 2020. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7180114/.

# Appendix 1

Table 1 2018.12.17-2018.12.23 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| Feature & Criminal | Topic1 | girl | 0.018 | 69 | 16% |
| | | character | 0.017 | | |
| | | **murderer** | 0.017 | | |
| | | favourite | 0.016 | | |
| | | obscure | 0.013 | | |
| | | historical | 0.013 | | |
| | | event | 0.013 | | |
| | | **breaking** | 0.012 | | |
| | | **pinkman** | 0.012 | | |
| | | **jesse** | 0.012 | | |
| Feature & Adventure | Topic2 | sure | 0.059 | 39 | 9% |
| | | imposed | 0.058 | | |
| | | **silverlake** | 0.055 | | |
| | | **tonto** | 0.05 | | |
| | | actor | 0.049 | | |
| | | man | 0.047 | | |
| | | silted | 0.046 | | |
| | | workshop | 0.046 | | |
| | | character | 0.043 | | |
| | | real | 0.043 | | |
| Documentary | Topic3 | movie | 0.048 | 126 | 29% |
| | | **throne** | 0.043 | | |
| | | great | 0.041 | | |
| | | delegate | 0.03 | | |
| | | **polish** | 0.03 | | |
| | | important | 0.03 | | |
| | | wikipedia | 0.03 | | |
| | | **henry** | 0.03 | | |
| | | explain | 0.03 | | |
| | | france | 0.03 | | |
| Comedy | Topic4 | piece | 0.025 | 75 | 17% |
| | | **parent** | 0.022 | | |
| | | **ben** | 0.022 | | |
| | | plan | 0.022 | | |
| | | **greg** | 0.022 | | |
| | | **byrnes** | 0.022 | | |
| | | propose | 0.022 | | |
| | | wikipedia | 0.022 | | |
| | | **focker** | 0.022 | | |

| | | | | | |
|---|---|---|---|---|---|
| | | learn | 0.022 | | |
| Comedy | Topic5 | time | 0.017 | 84 | 22% |
| | | sound | 0.015 | | |
| | | age | 0.015 | | |
| | | home | 0.012 | | |
| | | experience | 0.01 | | |
| | | **guy** | 0.009 | | |
| | | **prarie** | 0.009 | | |
| | | **foley** | 0.009 | | |
| | | **horse** | 0.009 | | |
| | | **companion** | 0.009 | | |

Table 2  2018.12.24-2018.12.30 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| Comedy | Topic1 | time | 0.015 | 84 | 22% |
| | | drama | 0.012 | | |
| | | **comedy** | 0.012 | | |
| | | big | 0.011 | | |
| | | book | 0.009 | | |
| | | **sherlock** | 0.009 | | |
| | | day | 0.008 | | |
| | | suggest | 0.008 | | |
| | | think | 0.008 | | |
| | | scenario | 0.008 | | |
| Plot | Topic2 | **film** | 0.032 | 67 | 17% |
| | | **stroytelling** | 0.016 | | |
| | | challenge | 0.014 | | |
| | | **visual** | 0.014 | | |
| | | director | 0.014 | | |
| | | speaking | 0.014 | | |
| | | device | 0.014 | | |
| | | **fun** | 0.014 | | |
| | | recent | 0.014 | | |
| | | creative | 0.014 | | |
| Animation & Character | Topic3 | **actor** | 0.014 | 87 | 23% |
| | | **character** | 0.014 | | |
| | | thing | 0.013 | | |
| | | called | 0.012 | | |
| | | person | 0.01 | | |
| | | mind | 0.008 | | |
| | | **cartoon** | 0.008 | | |
| | | girl | 0.008 | | |

| | | | | | |
|---|---|---|---|---|---|
| | | great | 0.008 | | |
| | | real | 0.007 | | |
| Feature & Criminal | Topic4 | episode | 0.018 | 148 | 38% |
| | | **bad** | 0.017 | | |
| | | face | 0.016 | | |
| | | moment | 0.012 | | |
| | | scene | 0.011 | | |
| | | died | 0.011 | | |
| | | **role** | 0.01 | | |
| | | **beraking** | 0.009 | | |
| | | romance | 0.009 | | |
| | | expecting | 0.009 | | |

Table 3 2018.12.30-2019.01.06 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| Criminal & Actor | Topic1 | time | 0.015 | 137 | 21% |
| | | drama | 0.012 | | |
| | | **comedy** | 0.012 | | |
| | | big | 0.011 | | |
| | | book | 0.009 | | |
| | | **sherlock** | 0.009 | | |
| | | day | 0.008 | | |
| | | suggest | 0.008 | | |
| | | think | 0.008 | | |
| | | scenario | 0.008 | | |
| Actor | Topic2 | featured | 0.025 | 163 | 25% |
| | | woman | 0.023 | | |
| | | thought | 0.022 | | |
| | | young | 0.02 | | |
| | | **penny** | 0.017 | | |
| | | **shirley** | 0.017 | | |
| | | **laverne** | 0.017 | | |
| | | marshall | 0.017 | | |
| | | recent | 0.017 | | |
| | | **news** | 0.017 | | |
| Feature & Character | Topic3 | rick | 0.018 | 265 | 41% |
| | | time | 0.017 | | |
| | | episode | 0.016 | | |
| | | question | 0.011 | | |
| | | actor | 0.01 | | |
| | | **jackson** | 0.01 | | |
| | | meant | 0.01 | | |

| | | | | | |
|---|---|---|---|---|---|
| | | **percy** | 0.01 | | |
| | | mean | 0.01 | | |
| | | teach | 0.01 | | |
| Comedy & Actor | Topic4 | happy | 0.032 | 78 | 12% |
| | | old | 0.032 | | |
| | | star | 0.031 | | |
| | | think | 0.027 | | |
| | | **winkler** | 0.027 | | |
| | | student | 0.026 | | |
| | | **high** | 0.026 | | |
| | | **fonzie** | 0.026 | | |
| | | wit | 0.026 | | |
| | | **henry** | 0.026 | | |

Table 4 2019.01.07-2019.01.13 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| Thriller | Topic1 | home | 0.034 | 115 | 14% |
| | | american | 0.032 | | |
| | | **horror** | 0.032 | | |
| | | long | 0.032 | | |
| | | toddler | 0.032 | | |
| | | mom | 0.032 | | |
| | | **bloody** | 0.032 | | |
| | | hidden | 0.032 | | |
| | | belly | 0.032 | | |
| | | **corpse** | 0.032 | | |
| Feature & Cult | Topic2 | human | 0.048 | 201 | 25% |
| | | **godfrey** | 0.048 | | |
| | | **reggio** | 0.048 | | |
| | | experience | 0.047 | | |
| | | produced | 0.047 | | |
| | | concentrate | 0.047 | | |
| | | philip | 0.047 | | |
| | | glass | 0.047 | | |
| | | musical | 0.047 | | |
| | | **koyanisqaasti** | 0.047 | | |
| SciFi & Deadpool | Topic3 | fred | 0.07 | 193 | 24% |
| | | **recreate** | 0.068 | | |
| | | savage | 0.067 | | |
| | | bedroom | 0.064 | | |
| | | hero | 0.063 | | |
| | | **deadpool** | 0.062 | | |

| | | | | | |
|---|---|---|---|---|---|
| | | **version** | 0.06 | | |
| | | kidnap | 0.048 | | |
| | | movie | 0.044 | | |
| | | princess | 0.043 | | |
| SciFi & Starwar | Topic4 | best | 0.039 | 217 | 27% |
| | | movie | 0.033 | | |
| | | **adlibbed** | 0.028 | | |
| | | **carbonite** | 0.028 | | |
| | | **empire** | 0.027 | | |
| | | **strike** | 0.027 | | |
| | | leia | 0.027 | | |
| | | choose | 0.027 | | |
| | | han | 0.027 | | |
| | | frozen | 0.027 | | |
| Criminal & Thriller | Topic5 | scene | 0.055 | 79 | 10% |
| | | upset | 0.04 | | |
| | | needle | 0.04 | | |
| | | **murdered** | 0.04 | | |
| | | **jesse** | 0.04 | | |
| | | **andrea** | 0.04 | | |
| | | **breaking** | 0.04 | | |
| | | **bad** | 0.04 | | |
| | | brock | 0.04 | | |
| | | inside | 0.04 | | |

Table 5 2019.01.14-2019.01.20 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| SciFi & Deadpool | Topic1 | movie | 0.025 | 324 | 34% |
| | | say | 0.016 | | |
| | | scene | 0.016 | | |
| | | **pg** | 0.016 | | |
| | | **version** | 0.015 | | |
| | | **deadpool** | 0.014 | | |
| | | princess | 0.014 | | |
| | | break | 0.014 | | |
| | | bedroom | 0.014 | | |
| | | savage | 0.014 | | |
| TV Show & Comedy | Topic2 | sex | 0.015 | 290 | 30% |
| | | written | 0.013 | | |
| | | character | 0.012 | | |
| | | main | 0.011 | | |
| | | **stern** | 0.01 | | |

| | | | | | |
|---|---|---|---|---|---|
| | | **prank** | 0.01 | | |
| | | potter | 0.01 | | |
| | | wizard | 0.01 | | |
| | | porkies | 0.01 | | |
| | | convince | 0.01 | | |
| TV Show | Topic3 | kind | 0.011 | 343 | 36% |
| | | read | 0.01 | | |
| | | say | 0.01 | | |
| | | **bengi** | 0.01 | | |
| | | **kya** | 0.01 | | |
| | | **fav** | 0.01 | | |
| | | **karna** | 0.01 | | |
| | | **ka** | 0.01 | | |
| | | **bhi** | 0.01 | | |
| | | **mai** | 0.01 | | |

Table 6 2019.01.21-2019.01.27 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| Comedy & Character | Topic1 | green | 0.047 | 72 | 16% |
| | | **chandler** | 0.041 | | |
| | | **schwimmer** | 0.041 | | |
| | | **cast** | 0.04 | | |
| | | **joey** | 0.039 | | |
| | | **leblanc** | 0.036 | | |
| | | **aniston** | 0.035 | | |
| | | **kudrow** | 0.035 | | |
| | | **rachel** | 0.035 | | |
| | | **best** | 0.035 | | |
| Scene & Criminal | Topic2 | movie | 0.016 | 71 | 16% |
| | | **story** | 0.013 | | |
| | | **dark** | 0.011 | | |
| | | **criminal** | 0.01 | | |
| | | people | 0.01 | | |
| | | **scene** | 0.01 | | |
| | | believe | 0.01 | | |
| | | english | 0.009 | | |
| | | thing | 0.009 | | |
| | | time | 0.009 | | |
| TV Show | Topic3 | lot | 0.016 | 68 | 15% |
| | | used | 0.016 | | |
| | | actor | 0.013 | | |
| | | sinatra | 0.012 | | |

| | | | | | |
|---|---|---|---|---|---|
| | | ncis | 0.012 | | |
| | | **superbowl** | 0.012 | | |
| | | **football** | 0.012 | | |
| | | loop | 0.012 | | |
| | | downtown | 0.012 | | |
| | | parking | 0.012 | | |
| Feature | Topic4 | movie | 0.017 | 153 | 40% |
| | | series | 0.016 | | |
| | | actor | 0.014 | | |
| | | people | 0.012 | | |
| | | **tom** | 0.01 | | |
| | | funny | 0.01 | | |
| | | **bombadil** | 0.01 | | |
| | | certain | 0.009 | | |
| | | juvenil | 0.009 | | |
| | | giant | 0.009 | | |
| Comedy | Topic5 | member | 0.031 | 91 | 20% |
| | | fear | 0.02 | | |
| | | near | 0.019 | | |
| | | day | 0.016 | | |
| | | **sitcom** | 0.016 | | |
| | | way | 0.016 | | |
| | | run | 0.016 | | |
| | | cast | 0.016 | | |
| | | loyal | 0.016 | | |
| | | left | 0.016 | | |

Table 7 2019.01.28-2019.02.03 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| Affectional | Topic1 | time | 0.047 | 205 | 30% |
| | | **ancient** | 0.046 | | |
| | | **love** | 0.045 | | |
| | | culture | 0.045 | | |
| | | language | 0.045 | | |
| | | **platonic** | 0.045 | | |
| | | understanding | 0.045 | | |
| | | code | 0.045 | | |
| | | space | 0.045 | | |
| | | exist | 0.045 | | |
| Feature | | watch | 0.016 | 268 | 39% |
| | | follow | 0.015 | | |
| | | like | 0.013 | | |

| | | | | | |
|---|---|---|---|---|---|
| & Criminal | Topic2 | set | 0.012 | | |
| | | **breaking** | 0.01 | | |
| | | plenty | 0.009 | | |
| | | coffee | 0.009 | | |
| | | listen | 0.009 | | |
| | | lion | 0.009 | | |
| | | **bad** | 0.009 | | |
| Feature | Topic3 | ancient | 0.019 | 105 | 15% |
| | | life | 0.016 | | |
| | | series | 0.015 | | |
| | | **game** | 0.013 | | |
| | | replaced | 0.012 | | |
| | | **throne** | 0.011 | | |
| | | nostalgia | 0.011 | | |
| | | capture | 0.009 | | |
| | | failure | 0.009 | | |
| | | cristo | 0.009 | | |
| Comedy & Actor | Topic4 | acting | 0.06 | 113 | 16% |
| | | role | 0.06 | | |
| | | **tarrak** | 0.06 | | |
| | | **metha** | 0.06 | | |
| | | **ka** | 0.06 | | |
| | | **ooltah** | 0.06 | | |
| | | **chashmah** | 0.06 | | |
| | | skill | 0.06 | | |
| | | experssion | 0.06 | | |
| | | great | 0.06 | | |

Table 8 2019.02.04-2019.02.10 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| Comedy | Topic1 | work | 0.014 | 101 | 15% |
| | | right | 0.014 | | |
| | | entertainment | 0.014 | | |
| | | bad | 0.013 | | |
| | | mood | 0.013 | | |
| | | fantasy | 0.012 | | |
| | | happen | 0.009 | | |
| | | **comedy** | 0.009 | | |
| | | advert | 0.009 | | |
| | | **broadcast** | 0.009 | | |
| | | death | 0.029 | 260 | 38% |
| | | **villian** | 0.026 | | |

| | | | | | |
|---|---|---|---|---|---|
| Criminal | Topic2 | line | 0.026 | | |
| | | slaughtere | 0.025 | | |
| | | potty | 0.025 | | |
| | | **trigger** | 0.025 | | |
| | | **cop** | 0.025 | | |
| | | action | 0.025 | | |
| | | caused | 0.025 | | |
| | | demand | 0.025 | | |
| War | Topic3 | **artillery** | 0.052 | 202 | 30% |
| | | intense | 0.052 | | |
| | | soviet | 0.052 | | |
| | | **american** | 0.052 | | |
| | | massive | 0.052 | | |
| | | say | 0.052 | | |
| | | affair | 0.052 | | |
| | | number | 0.052 | | |
| | | **gun** | 0.052 | | |
| | | **russian** | 0.052 | | |
| Documentary & Comedy | Topic4 | **sitcom** | 0.083 | 95 | 14% |
| | | **documentary** | 0.077 | | |
| | | mood | 0.067 | | |
| | | **picturization** | 0.066 | | |
| | | style | 0.066 | | |
| | | unlike | 0.065 | | |
| | | make | 0.064 | | |
| | | **trust** | 0.062 | | |
| | | **precise** | 0.059 | | |
| | | **jiffy** | 0.056 | | |

Table 9 2019.02.11-2019.02.17 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| Comedy | Topic1 | **shatner** | 0.029 | 41 | 7% |
| | | **skywalker** | 0.029 | | |
| | | audrey | 0.029 | | |
| | | hepburn | 0.029 | | |
| | | **william** | 0.029 | | |
| | | dorothy | 0.029 | | |
| | | **indiana** | 0.029 | | |
| | | **jones** | 0.029 | | |
| | | luke | 0.029 | | |
| | | lucy | 0.029 | | |

| | | | | | |
|---|---|---|---|---|---|
| Criminal | Topic2 | **batman** | 0.013 | 196 | 35% |
| | | completely | 0.013 | | |
| | | film | 0.013 | | |
| | | birdman | 0.012 | | |
| | | typical | 0.012 | | |
| | | freeze | 0.012 | | |
| | | **terminator** | 0.012 | | |
| | | arnold | 0.012 | | |
| | | **superhero** | 0.012 | | |
| | | **lantern** | 0.012 | | |
| War | Topic3 | cute | 0.057 | 154 | 27% |
| | | banoo | 0.054 | | |
| | | dulhann | 0.053 | | |
| | | divyanka | 0.05 | | |
| | | dabmn | 0.049 | | |
| | | rose | 0.047 | | |
| | | acting | 0.045 | | |
| | | **mohabbatein** | 0.044 | | |
| | | household | 0.042 | | |
| | | favourite | 0.042 | | |
| Documentary & Comedy | Topic4 | thing | 0.022 | 173 | 31% |
| | | country | 0.014 | | |
| | | **orphan** | 0.014 | | |
| | | **chumlum** | 0.014 | | |
| | | ron | 0.014 | | |
| | | rice | 0.014 | | |
| | | week | 0.013 | | |
| | | **aesthetically** | 0.013 | | |
| | | beautiful | 0.013 | | |
| | | example | 0.013 | | |

Table 10 2019.02.18-2019.02.24 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| SciFi | Topic1 | carrie | 0.034 | 159 | 28% |
| | | actor | 0.03 | | |
| | | **kirk** | 0.029 | | |
| | | neil | 0.029 | | |
| | | patrick | 0.029 | | |
| | | fisher | 0.029 | | |
| | | **skywalker** | 0.029 | | |
| | | luke | 0.029 | | |
| | | friend | 0.029 | | |

| | | | | | |
|---|---|---|---|---|---|
| | | reef | 0.029 | | |
| Thriller & Criminal | Topic2 | mother | 0.048 | 82 | 14% |
| | | **lake** | 0.046 | | |
| | | birth | 0.046 | | |
| | | new | 0.046 | | |
| | | jersey | 0.046 | | |
| | | **caystal** | 0.046 | | |
| | | **voorhees** | 0.046 | | |
| | | allow | 0.046 | | |
| | | introduce | 0.046 | | |
| | | **jason** | 0.046 | | |
| Animation | Topic3 | best | 0.017 | 195 | 34% |
| | | subtitle | 0.016 | | |
| | | enjoy | 0.016 | | |
| | | netfilx | 0.012 | | |
| | | **anime** | 0.012 | | |
| | | attention | 0.012 | | |
| | | understand | 0.012 | | |
| | | wat | 0.012 | | |
| | | actual | 0.012 | | |
| | | tongue | 0.012 | | |
| Criminal | Topic4 | movie | 0.027 | 139 | 24% |
| | | mind | 0.026 | | |
| | | deep | 0.026 | | |
| | | **matthau** | 0.022 | | |
| | | crook | 0.021 | | |
| | | audience | 0.021 | | |
| | | death | 0.02 | | |
| | | bank | 0.02 | | |
| | | **robbery** | 0.02 | | |
| | | resulting | 0.02 | | |

Table 11 2019.02.25-2019.03.03 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| SciFi & Adventure | Topic1 | series | 0.058 | 195 | 25% |
| | | replaced | 0.051 | | |
| | | **stargate** | 0.05 | | |
| | | half | 0.05 | | |
| | | **farscape** | 0.05 | | |
| | | tv | 0.05 | | |
| | | cast | 0.05 | | |

| | | | | | |
|---|---|---|---|---|---|
| | | **sg** | 0.05 | | |
| | | death | 0.05 | | |
| | | scene | 0.05 | | |
| Thriller | Topic2 | character | 0.057 | 170 | 22% |
| | | rewatch | 0.05 | | |
| | | slayer | 0.05 | | |
| | | **vampire** | 0.05 | | |
| | | **firefly** | 0.05 | | |
| | | whedon | 0.05 | | |
| | | think | 0.05 | | |
| | | school | 0.05 | | |
| | | creator | 0.05 | | |
| | | joss | 0.05 | | |
| Criminal | Topic3 | work | 0.026 | 73 | 9% |
| | | jake | 0.026 | | |
| | | **prove** | 0.026 | | |
| | | **detective** | 0.026 | | |
| | | talented | 0.026 | | |
| | | cop | 0.026 | | |
| | | peralta | 0.026 | | |
| | | ray | 0.026 | | |
| | | holt | 0.026 | | |
| | | brooklyn | 0.026 | | |
| Criminal & Thriller | Topic4 | actor | 0.018 | 266 | 34% |
| | | role | 0.017 | | |
| | | **netfilx** | 0.016 | | |
| | | **crime** | 0.011 | | |
| | | **stalker** | 0.011 | | |
| | | new | 0.011 | | |
| | | dirty | 0.011 | | |
| | | perfect | 0.01 | | |
| | | success | 0.01 | | |
| | | **thriller** | 0.01 | | |
| SciFi & Captain Marvel | Topic5 | **man** | 0.024 | 75 | 10% |
| | | started | 0.019 | | |
| | | imperium | 0.016 | | |
| | | **captain** | 0.015 | | |
| | | **jr** | 0.014 | | |
| | | **iron** | 0.013 | | |
| | | **marvel** | 0.012 | | |
| | | peak | 0.012 | | |
| | | screen | 0.012 | | |

| | | | | | |
|---|---|---|---|---|---|
| | | success | 0.012 | | |

Table 12 2019.03.04-2019.03.10 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| SciFi | Topic1 | time | 0.021 | 316 | 34% |
| | | question | 0.017 | | |
| | | clean | 0.016 | | |
| | | april | 0.016 | | |
| | | **sabachthani** | 0.016 | | |
| | | **anatomy** | 0.016 | | |
| | | **jesus** | 0.016 | | |
| | | died | 0.016 | | |
| | | cross | 0.016 | | |
| | | **eloi** | 0.016 | | |
| Thriller & Criminal | Topic2 | affair | 0.027 | 232 | 25% |
| | | urinate | 0.025 | | |
| | | club | 0.016 | | |
| | | **carson** | 0.012 | | |
| | | wife | 0.011 | | |
| | | **johnny** | 0.011 | | |
| | | **friar** | 0.011 | | |
| | | roast | 0.011 | | |
| | | camera | 0.011 | | |
| | | tonight | 0.011 | | |
| Animation | Topic3 | looking | 0.02 | 177 | 19% |
| | | try | 0.019 | | |
| | | **case** | 0.019 | | |
| | | **drama** | 0.018 | | |
| | | lost | 0.017 | | |
| | | feel | 0.014 | | |
| | | day | 0.012 | | |
| | | **hwang** | 0.011 | | |
| | | serial | 0.011 | | |
| | | uncover | 0.011 | | |
| Criminal | Topic4 | california | 0.021 | 207 | 22% |
| | | way | 0.013 | | |
| | | fun | 0.011 | | |
| | | beautiful | 0.011 | | |
| | | violence | 0.011 | | |
| | | uk | 0.011 | | |
| | | **battle** | 0.011 | | |

| | | | | | |
|---|---|---|---|---|---|
| | | **throne** | 0.011 | | |
| | | **war** | 0.011 | | |
| | | boring | 0.011 | | |

Table 13 2019.03.11-2019.03.17 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| Criminal & Sherlock Holmes | Topic1 | **holmes** | 0.028 | 181 | 29% |
| | | school | 0.026 | | |
| | | produced | 0.026 | | |
| | | fan | 0.026 | | |
| | | **law** | 0.025 | | |
| | | world | 0.024 | | |
| | | **benedict** | 0.024 | | |
| | | longer | 0.024 | | |
| | | wife | 0.024 | | |
| | | **cumberbach** | 0.024 | | |
| Criminal & Actor | Topic2 | **series** | 0.044 | 143 | 23% |
| | | watch | 0.022 | | |
| | | favourite | 0.021 | | |
| | | **holmes** | 0.02 | | |
| | | **solomin** | 0.019 | | |
| | | fan | 0.019 | | |
| | | music | 0.019 | | |
| | | **watson** | 0.019 | | |
| | | **vasily** | 0.019 | | |
| | | youtube | 0.019 | | |
| Feature & Actor | Topic3 | movie | 0.047 | 159 | 25% |
| | | best | 0.029 | | |
| | | **role** | 0.029 | | |
| | | version | 0.028 | | |
| | | talkie | 0.028 | | |
| | | **eagels** | 0.028 | | |
| | | paramount | 0.028 | | |
| | | **letter** | 0.028 | | |
| | | nominated | 0.028 | | |
| | | **jeanne** | 0.028 | | |
| SciFi | Topic4 | ca | 0.05 | 61 | 10% |
| | | urge | 0.05 | | |
| | | people | 0.05 | | |
| | | meet | 0.05 | | |
| | | **convention** | 0.05 | | |

| | | | | | |
|---|---|---|---|---|---|
| | | **attendee** | 0.05 | | |
| | | **firefly** | 0.05 | | |
| | | stuff | 0.05 | | |
| | | **resist** | 0.05 | | |
| | | really | 0.05 | | |
| Feature & War | Topic5 | world | 0.018 | 83 | 13% |
| | | series | 0.015 | | |
| | | like | 0.013 | | |
| | | actress | 0.012 | | |
| | | **sky** | 0.012 | | |
| | | scene | 0.01 | | |
| | | **normandy** | 0.008 | | |
| | | **brother** | 0.008 | | |
| | | **company** | 0.008 | | |
| | | **hbo** | 0.008 | | |

Table 14 2019.03.18-2019.03.24 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| Animation | Topic1 | good | 0.045 | 70 | 17% |
| | | instantly | 0.043 | | |
| | | **maid** | 0.043 | | |
| | | **datin** | 0.042 | | |
| | | **programmer** | 0.042 | | |
| | | **geek** | 0.042 | | |
| | | intorvert | 0.042 | | |
| | | **miss** | 0.042 | | |
| | | **kobayashi** | 0.042 | | |
| | | **dragon** | 0.042 | | |
| Comedy & War | Topic2 | character | 0.067 | 81 | 19% |
| | | popular | 0.05 | | |
| | | actor | 0.031 | | |
| | | homepage | 0.031 | | |
| | | **holocust** | 0.029 | | |
| | | toner | 0.029 | | |
| | | video | 0.025 | | |
| | | pop | 0.022 | | |
| | | **vudu** | 0.022 | | |
| | | base | 0.022 | | |
| | | watch | 0.12 | 130 | 31% |

| | | | | | |
|---|---|---|---|---|---|
| Documentary & SciFi | Topic3 | time | 0.1 | | |
| | | **speed** | 0.096 | | |
| | | aspect | 0.093 | | |
| | | mankind | 0.093 | | |
| | | **future** | 0.09 | | |
| | | commercial | 0.016 | | |
| | | **hulu** | 0.016 | | |
| | | **documentary** | 0.008 | | |
| | | **ai** | 0.008 | | |
| Comedy & SciFi | Topic4 | work | 0.045 | 141 | 33% |
| | | decribed | 0.045 | | |
| | | **comedian** | 0.045 | | |
| | | screen | 0.045 | | |
| | | battle | 0.045 | | |
| | | **monster** | 0.045 | | |
| | | **seller** | 0.045 | | |
| | | **peter** | 0.045 | | |
| | | depression | 0.045 | | |
| | | manic | 0.045 | | |

Table 15 2019.03.25-2019.03.31 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| Criminal & Character | Topic1 | experience | 0.056 | 96 | 13% |
| | | tv | 0.052 | | |
| | | **depicted** | 0.052 | | |
| | | bad | 0.052 | | |
| | | **prison** | 0.051 | | |
| | | **pain** | 0.051 | | |
| | | **violence** | 0.051 | | |
| | | separated | 0.051 | | |
| | | heatache | 0.051 | | |
| | | witnessed | 0.01 | | |
| Comedy & Actor | Topic2 | **hazzard** | 0.045 | 201 | 27% |
| | | episode | 0.045 | | |
| | | **john** | 0.045 | | |
| | | ran | 0.045 | | |
| | | salary | 0.045 | | |
| | | **wopat** | 0.045 | | |
| | | **tom** | 0.045 | | |
| | | dispute | 0.045 | | |
| | | pretty | 0.045 | | |
| | | thing | 0.045 | | |
| | | information | 0.021 | 216 | 29% |

| | | | | | |
|---|---|---|---|---|---|
| SciFi & Actor | Topic3 | happens | 0.016 | | |
| | | character | 0.014 | | |
| | | watch | 0.014 | | |
| | | **fictional** | 0.012 | | |
| | | nature | 0.012 | | |
| | | **khardshian** | 0.012 | | |
| | | use | 0.012 | | |
| | | project | 0.012 | | |
| | | **software** | 0.012 | | |
| SciFi & Star Trek | Topic4 | **gravity** | 0.026 | 226 | 31% |
| | | **star** | 0.022 | | |
| | | movie | 0.019 | | |
| | | **trek** | 0.018 | | |
| | | science | 0.018 | | |
| | | **artificial** | 0.016 | | |
| | | **production** | 0.014 | | |
| | | commercial | 0.013 | | |
| | | fiction | 0.011 | | |
| | | cast | 0.01 | | |

Table 16 2019.04.01-2019.04.07 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| TV Show & Comedy | Topic1 | thing | 0.031 | 102 | 19% |
| | | **comedy** | 0.027 | | |
| | | distant | 0.026 | | |
| | | smile | 0.021 | | |
| | | simple | 0.021 | | |
| | | **tera** | 0.021 | | |
| | | **monkey** | 0.021 | | |
| | | **sharma** | 0.021 | | |
| | | **mimicry** | 0.021 | | |
| | | **kapil** | 0.021 | | |
| SciFi & Thriller | Topic2 | movie | 0.03 | 104 | 19% |
| | | like | 0.018 | | |
| | | think | 0.018 | | |
| | | **stranger** | 0.018 | | |
| | | want | 0.017 | | |
| | | real | 0.014 | | |
| | | **thing** | 0.014 | | |
| | | body | 0.011 | | |

| | | | | | |
|---|---|---|---|---|---|
| | | supes | 0.011 | | |
| | | **demodogs** | 0.011 | | |
| Comedy & Affectional | Topic3 | movie | 0.026 | 85 | 16% |
| | | actor | 0.025 | | |
| | | **mcluhan** | 0.017 | | |
| | | **woody** | 0.016 | | |
| | | line | 0.016 | | |
| | | theory | 0.016 | | |
| | | **annie** | 0.016 | | |
| | | **alvy** | 0.016 | | |
| | | **singer** | 0.016 | | |
| | | **hall** | 0.016 | | |
| TV Show & Host | Topic4 | question | 0.017 | 85 | 16% |
| | | missed | 0.015 | | |
| | | example | 0.014 | | |
| | | **jeopardy** | 0.014 | | |
| | | **diagonsed** | 0.013 | | |
| | | probabaly | 0.013 | | |
| | | routine | 0.013 | | |
| | | **host** | 0.013 | | |
| | | **alex** | 0.013 | | |
| | | **stage** | 0.013 | | |
| Comedy & Thriller | Topic5 | movie | 0.037 | 166 | 30% |
| | | thing | 0.021 | | |
| | | star | 0.02 | | |
| | | **miss** | 0.019 | | |
| | | based | 0.019 | | |
| | | **rodriguez** | 0.017 | | |
| | | nanjano | 0.017 | | |
| | | **bala** | 0.017 | | |
| | | **gina** | 0.017 | | |
| | | **comedy** | 0.016 | | |

Table 17 2019.04.08-2019.04.14 Results of Topics Mining

| Topic Tags | Topic No. | Top Words | Words Weight | Topics Frequency | Topics Proportion |
|---|---|---|---|---|---|
| Comedy | Topic1 | episode | 0.04 | 55 | 9% |
| | | guy | 0.039 | | |
| | | **family** | 0.038 | | |
| | | maintain | 0.035 | | |
| | | quo | 0.028 | | |
| | | **adultoriented** | 0.028 | | |

| | | | | | |
|---|---|---|---|---|---|
| | | change | 0.026 | | |
| | | phenomenon | 0.026 | | |
| | | **simpson** | 0.026 | | |
| | | favourite | 0.026 | | |
| Criminal | Topic2 | people | 0.045 | 182 | 31% |
| | | carrer | 0.044 | | |
| | | episode | 0.042 | | |
| | | city | 0.041 | | |
| | | **weapon** | 0.041 | | |
| | | **cop** | 0.041 | | |
| | | **firing** | 0.041 | | |
| | | **shooting** | 0.041 | | |
| | | **criminal** | 0.041 | | |
| | | **police** | 0.041 | | |
| SciFi & The Avengers | Topic3 | thanks | 0.015 | 262 | 44% |
| | | **world** | 0.014 | | |
| | | frank | 0.014 | | |
| | | **paul** | 0.013 | | |
| | | harris | 0.013 | | |
| | | invisible | 0.013 | | |
| | | **hemsworth** | 0.013 | | |
| | | bana | 0.013 | | |
| | | liam | 0.013 | | |
| | | jordan | 0.013 | | |
| SciFi & Fantastic Four | Topic4 | standard | 0.028 | 95 | 16% |
| | | **fiction** | 0.028 | | |
| | | examine | 0.028 | | |
| | | structure | 0.028 | | |
| | | informative | 0.028 | | |
| | | drama | 0.028 | | |
| | | cast | 0.028 | | |
| | | **chikills** | 0.017 | | |
| | | **jessica** | 0.017 | | |
| | | **hero** | 0.017 | | |