# DATA SCIENCE, AN EMERGING FIELD IN LIBRARY AND INFORMATION SCIENCE: A YOUTUBE ANALYTICS STUDY OF TWO POPULAR DATA SCIENCE PROGRAMMING TOOLS - PYTHON AND R

PALLAVI
Research Scholar
School of Library and Information Science
Central University of Gujarat
Gandhinagar, Gujarat-382030
Email: srivastav.pal55@gmail.com

Dr. BHAKTI GALA
Assistant Professor
School of Library and Information Science
Central University of Gujarat
Gandhinagar, Gujarat-382030
Email: bhakti.gala@cug.ac.in

**ABSTRACT**

*R and Python are two popular data science programming languages being extensively used by researchers of all disciplines. This research has looked at user interaction data on YouTube videos on R and Python. The author examined the user engagement statistics from 539 R programming and 529 Python-related YouTube videos. The researchers adopted a mixed-method research approach and incorporated both quantitative and qualitative data for the analysis. For data collection, Webometric Analyst software was used to extract YouTube metadata by creating YouTube-API for analysing usage parameters and R Programming was used to collect, analyse, and visualize the comments for sentiment analysis. The study findings show an increasing trend in the publishing of videos in R programming. The maximum number of videos on both the topics was found to range from 2 to30 minutes time duration. In R programming most viewed videos are from the years 2013 and 2017 and in Python from 2019. The research further identified the positive co-relation between the view count and number of likes and comments of videos. Sentiment analysis of fetched comments on both topics was found to be positive.*

*Keywords: Data Science, Data Science Tools, R Programming, Python, YouTube Videos, YouTube Analytics, Programming language Videos, Data Visualisation, Opinion Mining, Video View Count, View Count, Like Count, Sentiment Analysis, Engagement Metrix, Licence of Videos, Webometric Analyst, Library and Information Science.*

## 1. Introduction

Continuous studies (Gupta and Chakravarty, 2021) in LIS illustrates that it has an interdisciplinary relationship with other disciplines such as information science, documentation, marketing, statistics, and so on. Over the year this discipline is changing, growing, and merging with new subjects. Gupta and Chakravarty (2021), Ma and Lund (2020) in their bibliometrics study shows the most popular areas in Indian LIS research in the last few years are machine learning, big data, text mining, sentiment analysis, social media, etc. are the fields related to data science..

Data science helps in the processing and execution of data policy making, predicting the flow of upcoming needs, and understanding users' insight, which allows stakeholders to conclude it with the help of software and programming languages. It has been seen that the rise of big data and data science has made research more impactful and data driven.

SPSS, SAS, SQL, R programming and Python are some of the software used by data librarians for the analysis of data, but R programming and Python are the two most popular programming languages used for data analysis among data scientists (Zhou and Ordonez, 2021) because they develop under the open-source licensing and offer facilities like data extraction, analysis, and visualization also. In library science R and

Python, like traditional statistical software (SPSS, SAS), not only aid in research data analysis (Carlozzi 2019), bibliometrics and keyword analysis (Wodeyar et.al 2022; Han 2020), but also help in data visualization, machine learning classification (Chang et. al (2021), extracting online interaction data from social media sites, sentiment analysis (Lund 2020) etc.

As per the future demand, librarians must learn the data science tools and its applications such as R, Python etc. and develop data scientist skills to expand data services in their libraries to better understand the needs of users in this age of microblogging and social networking. For basic to advance level of tutorials number of videos are available on YouTube that can assist academicians, librarians, researchers in learning the features and applications of R and Python.

Easy and global access to YouTube content has made it a favourable tool for teachers and students during the COVID-19 pandemic, YouTube has gained further popularity in education and has been acknowledged as a learning and problem-solving platform (Temban et al., 2021) and also known as the second most popular search engine (Gupta et. Al, 2017). Teachers and creators of various discipline like Dental education (Dascalu et al., 2021) Chemistry (Jackson, 2017) and Music (Serdaroglu, 2020) Library and Information Science (LIS) adopted this technology for teaching and learning purposes (Singh & Mahawar, 2020). They are posting tutorials and lectures on different topics and demonstrating application of many software tools and packages on the platform. Viewers can easily search and retrieve R and Python videos on YouTube.

We can see the popularity of search terms 'R programming' and 'Python' worldwide from 2008 (as data is available from 2008 on google trends) to 2022. Figure 1 depicts the Google Trends for YouTube search comparing the search trends for both the topics. It shows Python is more trending these days than R Programming.

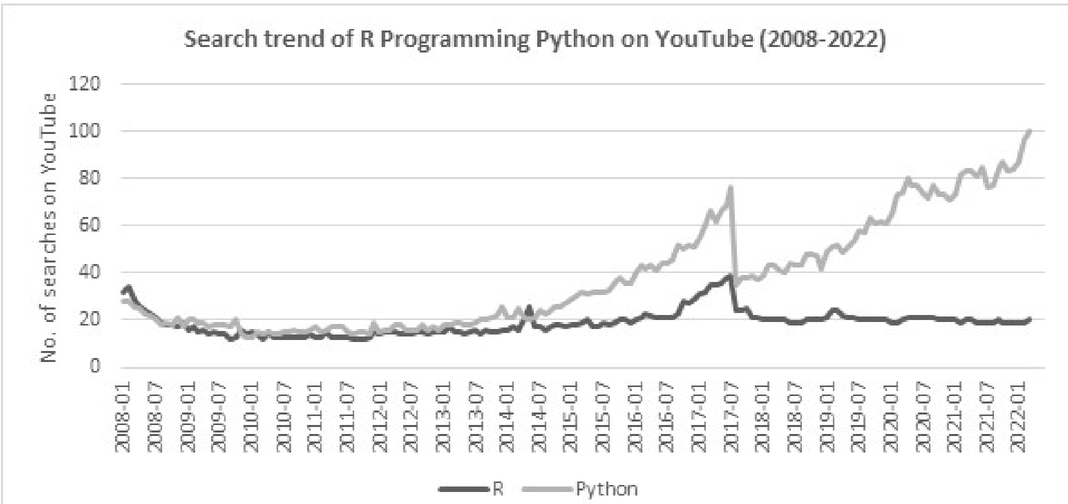



**Fig. 1. Search trend of R Programming and Python on YouTube**

Many studies (Goeser et al., 2012; Kadriu et al., 2020) have been conducted on YouTube videos for software and programming languages, but no studies have been conducted on comparison and analysis of videos available on R programming and Python. The present study is an in-depth analysis of YouTube videos on these two data analysis programming languages.

## 2. Review of Literature

Educators are embracing web 2.0 and social networking tools to increase attention and engagement in learning which help to build a strong relationship between content creators and learners. (Manca & Ranieri, 2017)

The global acceptance of YouTube as a provider of access to knowledge (Yaacob and Saad, 2020), has enabled it to emerge as a new mode of communication and learning (Neumann and Herodotou, 2020). The number of material and video categories on YouTube is growing day by day (Che et al., 2015), however, to measure its usage, popularity, and satisfaction among viewers, it is required to study user-generated data on YouTube, such as engagement matrices or sentiment analysis. King (2015) and Pallavi et al. (2019) studied the social media metrics to help track viewers' interaction, evaluate the performance of YouTube videos (King, 2015; Pallavi et al. 2019). Cheng et al. (2008) emphasized that comments are less prevalent than ratings, and both are significantly less common than views, according to user behaviour and video view pattern research to understand the viewers. View count is an important aspect to understand the popularity of the videos. Barjasteh et al. (2014) evaluated the major elements of 8,000 YouTube trending videos and observed a consistent audience tendency toward popular categories (Barjasteh et al., 2014). The popularity and View count of the videos are determined by various factors. Jansen and Salminen (2017) studied the audience attribute and video attribute to understand the relation between audience attribute and View count of a video.

Different tools and methods have been used to extract and analyse metadata of YouTube content. In their study Parabhoi and Chand (2018), Mandal (2021) and many more used webometrics analyst software to extract data from YouTube related to various channels, topics, tutorials in order to analyze the growth, license, duration, and engagement matrices of the videos. Deori et al. (2021) also did the sentiment analysis to understand the viewpoint of the viewers (Deori et al., 2021). Sentiment analysis, or opinion mining is a process of text analysis to understand "people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes" (Liu, 2012). The comment feature of YouTube aims to gather feedback on the videos you've published. Madden et al. (2013) conducted content analysis of YouTube comments and classified them into ten broad categories, and 58 subcategories for their application and examination.. Al-Tamimi et al. (2017) worked on sentiment analysis of YouTube Arabic comments. Siersdorfer et al. (2010) investigated the impact of sentiment by using SentiWordNet thesaurus. Siersdorfer (2010) and Kabir et al. (2020) have done sentiment analysis of reviews of earphones available on Amazon. This study also analyses the YouTube video's publication, view, etc. with the sentiment of viewers on R programming and Python related videos.

## 3. Objectives of the Study

1. To explore the publication pattern of uploaded videos on YouTube on R Programming and Python.
2. To analyse the various parameters of the videos i.e., engagement matrices, duration, license of R Programming and Python.
3. To investigate the reasons affecting video views.
4. To study the publication patterns of the most viewed videos.
5. To examine the sentiment of comments received on R Programming and Python-related videos.

## 4. Scope and Limitation

This study aims to analyse the viewer's behaviour and sentiments expressed in form of comments on the videos available on YouTube on R Programming and Python programming languages. The limitation of the study is that it has included videos extracted under the keyword "R programming software" and "Python software" and published through 31st January, 2022. The analysis is restricted to the publication of the video, their engagements parameters, and comments.

## 5. Methodology

The purpose of this study was to investigate the online interaction behaviour and sentiment polarity of viewers of videos on R programming and Python programming language. To address the research objectives, we employed a variety of research approaches, including web extraction, data mining, quantitative, and qualitative methodologies. Data collection and analysis of this study is divided into two parts:

In the first part, the characteristics and usage parameters of the gathered video were extracted and analysed. For that data of 1074 videos (R programming 539 and Python 529) were extracted on 31st January, 2022. Webometric analyst (http://lexiurl.wlv.ac.uk/) was used to scrape data for the analysis of the usage parameters. It is efficient to retrieve a list of videos that match the searched keyterm, channel ID, or tag. Here we have run a search on webometric analyst with the terms "R programming software" and "Python software,". This software used YouTube API 3v which is an application programming interface (API) that enable users to access and retrieve data from YouTube channels and videos. All publicly available data available on YouTube was thus captured and further imported and analysed using Excel. The metadata elements collected from the webometric analyst to study various characteristics and usage parameters of videos on YouTube for the study were: Video ID, channel ID, Video Title, Publication Date, Duration of Video, Views Count, Likes Count, Comments Count and License.

In the second part, the comments of the retrieved videos were extracted for sentiment analysis. Total 86399 comments were fetched by using R programming (with the help of R studio). by creating a query with the video IDs (collected by the webometric analyst software in first part) separated with comma, followed by data cleaning like removal of semi comas, blank spacing, converting emojis (picture character) into text, etc. 36264 comments from Python (22245 comments) and R programming (14018 comments) related videos have been taken for final analysis. The metadata elements collected for sentiment analysis were Video ID, Text of comment, comment writer's name, and Reply count. This data was then cleaned by converting emojis into text; translating the Hindi language comments into English, and removing comments which were neither in English or Hindi language and therefore incomprehensible. The analysis was done on R studio by using various packages vosonSML, tuber, httpuv, purrr, ROAuth, syuzhet, etc. after installing these packages and setting up OAuth (is an open authorization, use to get limited access of information from a website or social media sites without revealing them the credentials). A query was created with YouTube video IDs to pull the comments from videos (query and other steps mentioned in figure 2. Following the execution of these codes, R Studio extracted comment data from YouTube IDs and R programme further automatically classifying them into eight separate emotions (anger, anticipation, disgust, fear, joy, sadness, surprise, and trust,) as well as two main (positive and negative) and one optional sentiment (neutral) by using NRC lexicon and NRC dictionary. For visualization of data, we used the barplot function of R studio (see figure 3).

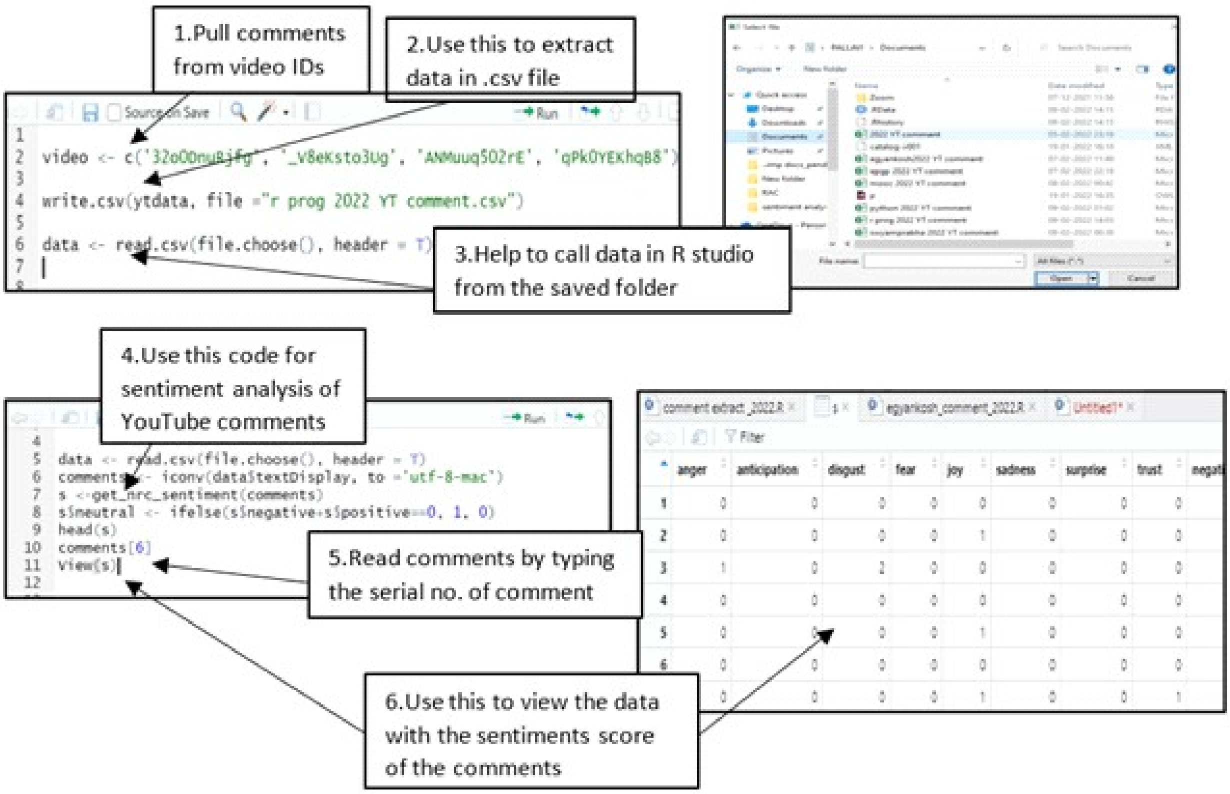



Fig.2. Coding in R Studio for YouTube Comment Scrapping and Analysing

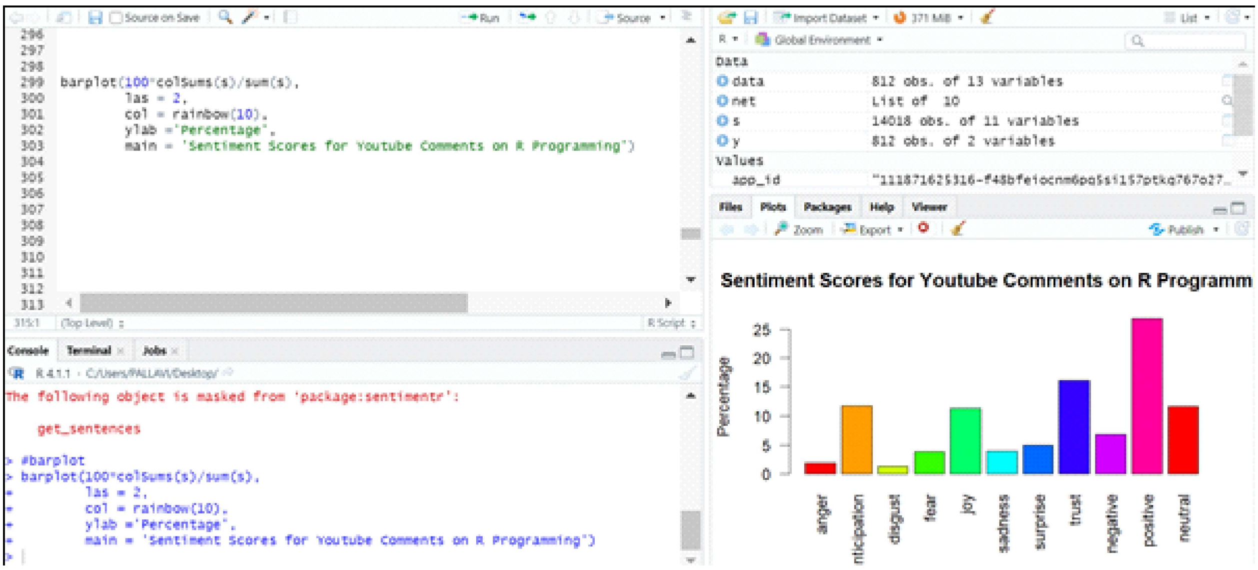



Fig. 3. Coding in R Studio for Visualization of Sentiments

# 6. Analysis and Results

## 6.1. Publication Pattern of Uploaded Videos

The growth in the publication pattern of the videos on YouTube on R programming and Python is seen in Figure 4. While the first video on R programming was released in 2009, the first video on Python was released in 2011. The results demonstrate that uploads of videos gradually increased for both the topics indicating a sustained interest in the use of these data science tools.

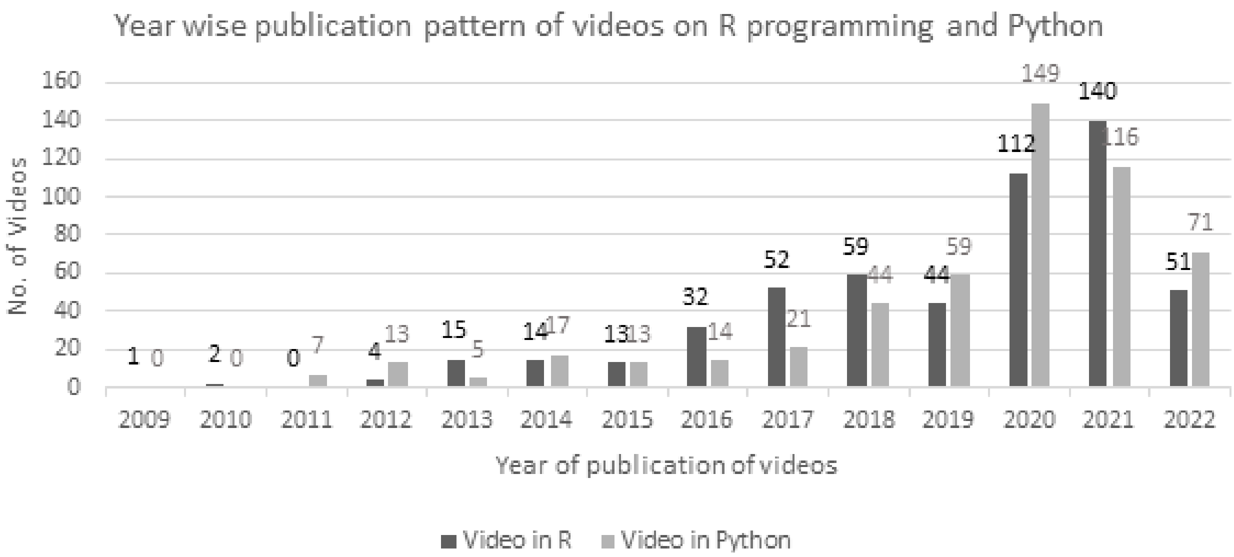



**Fig. 4. Year-wise Publication Pattern of Videos on R Programming and Python**

Thus, although the first video on R programming was launched in 2009, it took longer to apex in 2021, whereas the first video on Python which was released two years later reached its apex earlier indicating the popularity of Python. An average growth rate of 38.5 % and 44% is seen for videos on R and Python respectively, showing the popularity of Python. The results further show that the noticeable increase in popularity of videos on R and Python peaked during the years of the pandemic, revealing the high use of YouTube as a source of self-learning these tools.

## 6.2. Analysis of Various Parameters of the Videos

### 6.2.1. Analysis of Engagement Matrices

The usage parameters analysed in this study are the engagement matrices. Engagement matrices are defined as the number of times a video or channel has interacted with the user. These matrices are measured and recorded with parameters such as views, likes, dislikes, and subscribers on YouTube. These parameters are very useful in determining the overall popularity of a video or channel.

#### 6.2.1.1. View Count Analysis

View count on YouTube reflects the number of time a video has been played or viewed by a viewer. Table 1 shows that 539 videos of R programming received 2263972 views, whereas 529 videos of Python received 57945633 views. Although the minimum view count for both subjects is 0, there is variation in the maximum view count. We can observe that the mean value of the video view count is 38535 for R Programming and 109538 for Python. In both cases, the median value is smaller than the mean and the standard deviation is greater than the mean, which indicates that the data is positively skewed and not normally distributed.

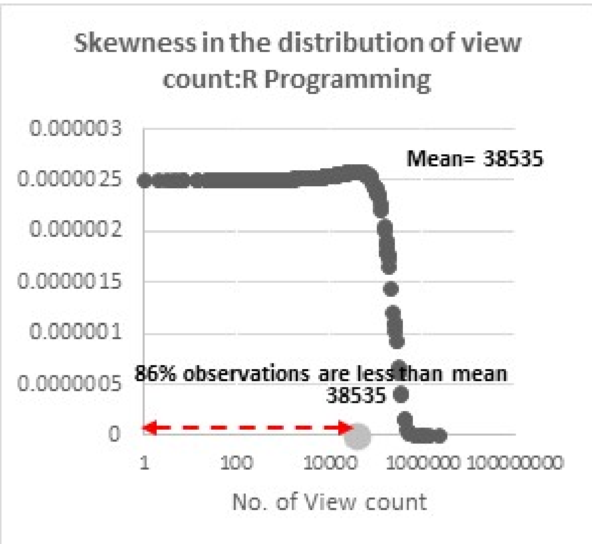
Table 1

Summary of View Count Analysis for R Programming and Python Videos

| Particulars | R Programming | Python |
|---|---|---|
| Total no. of videos | 539 | 529 |
| Total view count | 20770225 | 57945633 |
| Maximum | 2263972 | 24051387 |
| Minimum | 0 | 0 |
| Mean | 38535 | 109538 |
| Median | 4155 | 142 |
| Standard deviation | 153995 | 1109838 |

It is observed from the analysis given in figure 5 that for both the topics, viz. R programming and Python many videos have been viewed less time than its mean (average) value. Eighty six percentage of videos have been seen less than their mean value and only 14% of videos have received more view count than their mean. The data indicates that 91% of Python videos received less view count and only 9% of videos received more view count than its mean.

This analysis also shows that the maximum number of videos on both topics have fewer view count than the average value of view count received by all videos of both topics. It means that maximum number of videos receive less view count (less than average view count) and fewer videos receive highest view count (which is more than average its view count). Although, a high view count does not necessarily imply high content quality, but this parameter in addition to others described below can help LIS researchers in narrowing down their choices for identifying specific videos to enhance their learning and training.



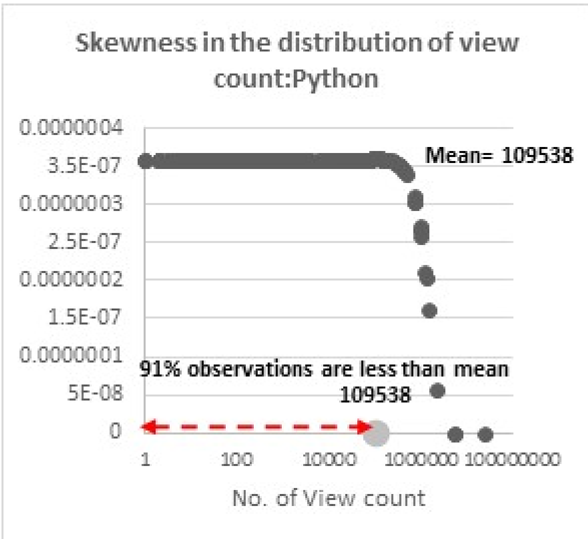
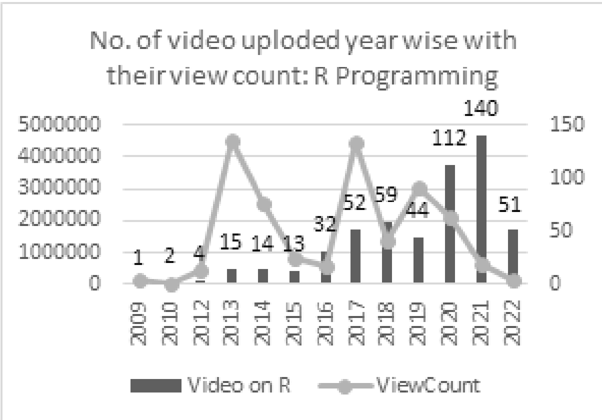




Fig. 5. Measurement of Skewness in the Distribution of View Count of R Programming and Python Videos

#### 6.2.1.2. Relation Between Number of Videos Uploaded in a Year with the View Count Received

The analysis given in graph (Fig. 6) shows that for R Programming, maximum videos were uploaded in the years 2021 and 2020 but videos that are most popular among viewers are from the years 2013 and 2017. Most Python videos were uploaded in the year 2020 and 2021, but the videos that gained more popularity among viewers are from the year 2019.



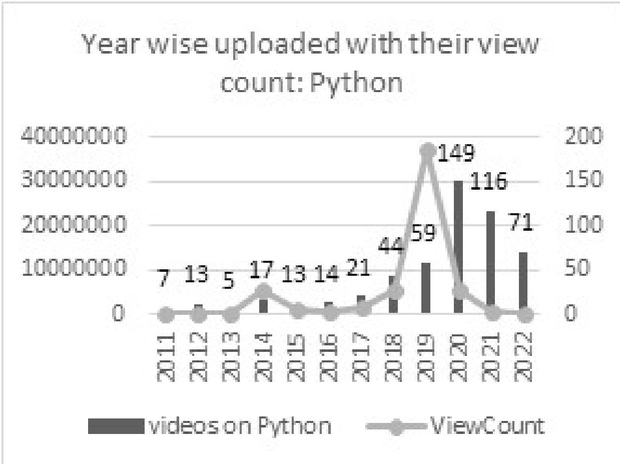




**Fig. 6. Year-wise Uploaded with their View Count of R Programming and Python**

#### 6.2.1.3. Like Count Analysis

Like count on social media is a way to measure the viewer's opinion about the video content. It is counted as the number of Likes received for a video by using frequency distribution. The results show that the total like count of R Programming videos is 229413 and Python is 1440631. The result (table 2) shows that 100 people like 326 videos from R Programming (total like count is 229413) and 404 videos for Python.

**Table 2**

**Like Count Analysis of R Programming and Python**

| Like count | Videos on R Programming | Videos on Python |
|---|---|---|
| 1-100 | 326 | 404 |
| 101-1000 | 154 | 54 |
| 1001-10000 | 33 | 47 |
| 10001-100000 | 4 | 12 |
| 100001-1000000 | 0 | 2 |

#### 6.2.2. Analysis of Duration

Duration of video indicates the length (whether the video is short or long) of the video uploaded to YouTube. The author used frequency distribution method to represent the "duration of videos" on topics. The findings of the analysis (table 3) demonstrate that 419 videos on R programming and 386 videos on Python are of 30 minutes duration. Number of 300+minute videos are very less, with only 9 and 7 videos in R Programming and Python, respectively. The data reveals that shorter videos of less than 30 minutes are the most popular for both topics.

Table 3

Duration of Video

| Duration of Video | Videos on R Programming | Videos on Python |
|---|---|---|
| 0-1 min | 4 | 39 |
| 2-30 min | 419 | 386 |
| 31-60 min | 37 | 50 |
| 61-90 min | 23 | 24 |
| 91-120 min | 16 | 8 |
| 121-150 min | 14 | 5 |
| 151-180 min | 3 | 4 |
| 181-210 min | 8 | 4 |
| 211-240 min | 4 | 1 |
| 241-270 min | 1 | 1 |
| 271-300 min | 1 | 0 |
| 300 min< | 9 | 7 |

### 6.2.3. License of Videos

A license is an agreement between a video and its creator which allows them to control its use and reuse. There are two types of licence with YouTube "standard YouTube License" (means creator grants the broadcasting rights to YouTube) and Creative Commons license, CC BY license (creator retain their copyright)" (Wikipedia, 2022). The analysis (figure 7) shows both the topics have more videos 98% of R Programming and 94% of Python under YouTube license and 2% from R Programming and 6% from Python videos are made under CC BY license.

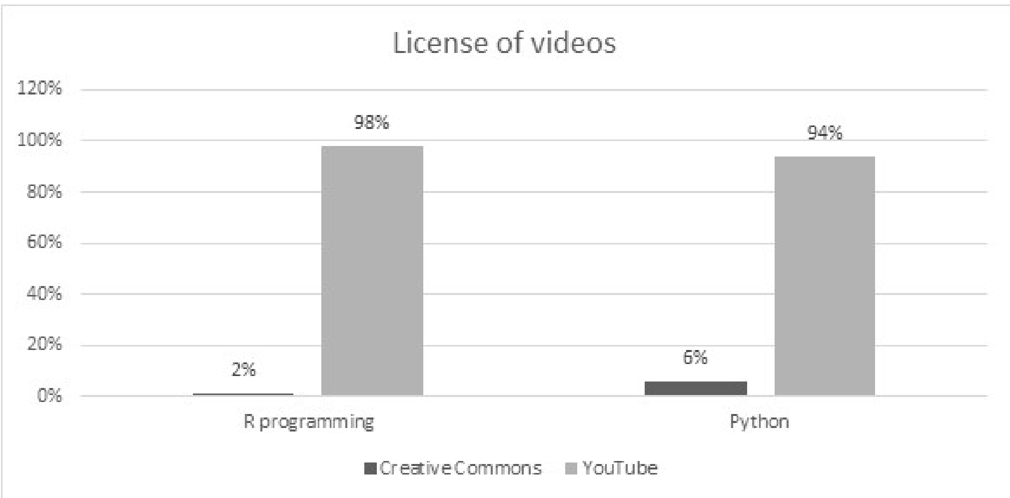



**Fig. 7. Comparison on License of Videos for R programming and Python**

## 6.3. Reasons Affecting Video Views

To find the relation between view count and publication year, duration, like, and comments which affect the view count of a video, this study used the Pearson correlation coefficient method. Choi (2019) describes how one variable (view count) is linearly connected to another in terms of direction and degree. The findings of the study demonstrate that view count has a strong correlation with like count and comment count but no significant correlation with upload year of video (Publish at) and duration of the video (Tables 4 and 5).

Table 4

Reasons Affecting Video View Count: Correlation Analysis for R Programming Data

| | View Count | Publish at | Duration Sec | Like Count | Comment Count |
|---|---|---|---|---|---|
| View Count | 1 | | | | |
| Publish at | 0.277985 | 1 | | | |
| Duration Sec | 0.062473 | -0.21648 | 1 | | |
| Like Count | 0.923649 | 0.124189 | 0.100918 | 1 | |
| Comment Count | 0.890362 | 0.194852 | 0.082852 | 0.859703 | 1 |

Table 5

Reasons Affecting Video View Count: Correlation Analysis for Python Data

| | View Count | Publish at | Duration Sec | Like Count | Comment Count |
|---|---|---|---|---|---|
| View Count | 1 | | | | |
| Publish at | 0.03794 | 1 | | | |
| Duration Sec | 0.268265 | -0.0125 | 1 | | |
| Like Count | 0.987523 | 0.017292 | 0.26092 | 1 | |
| Comment Count | 0.970524 | 0.025271 | 0.252962 | 0.983519 | 1 |

## 6.4. Year-wise Publication of Most Viewed Videos (top 200)

In this segment, we examined the most viewed (top 200) videos from both topics regarding their year of publication and duration. The top 200 most viewed videos (Figure 7) show that although most of the videos are uploaded in the year 2018, the most viewed videos are from 2013, 2017, and 2019 indicating that there is volatility in its View count in R programming. The following Python graph in figure 6 shows that the maximum videos are from uploaded in the year 2020 however, the most viewed videos are from 2019.

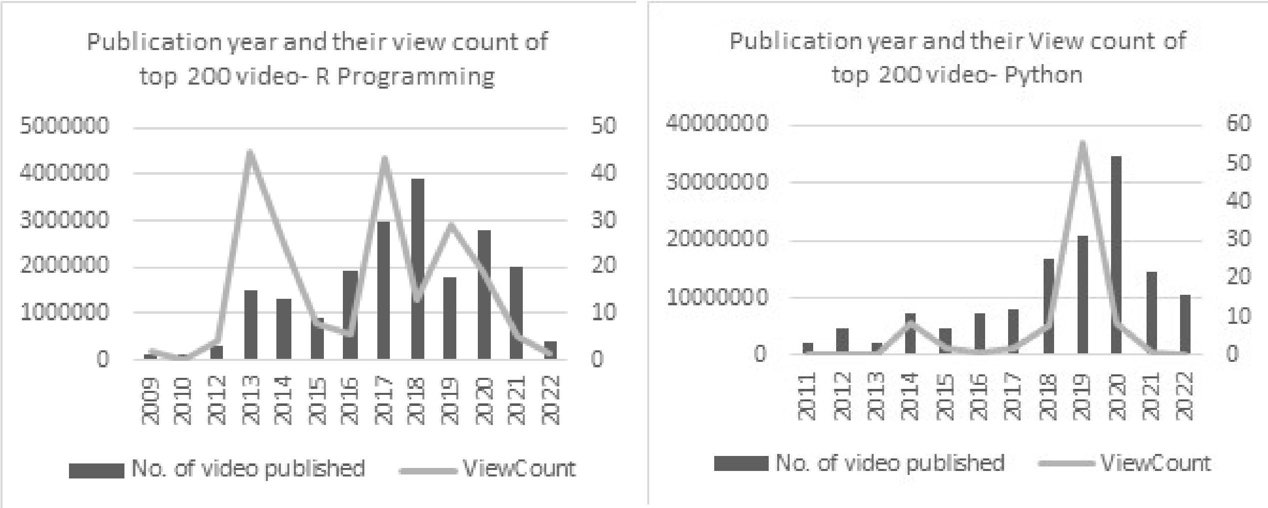



**Fig. 8. Publication year and their View count of top 200 video- R Programming and Python**

According to the data given in figure 8, majority of the top 200 most viewed videos are 30 minutes in duration, followed by 60 and 90 minutes respectively.

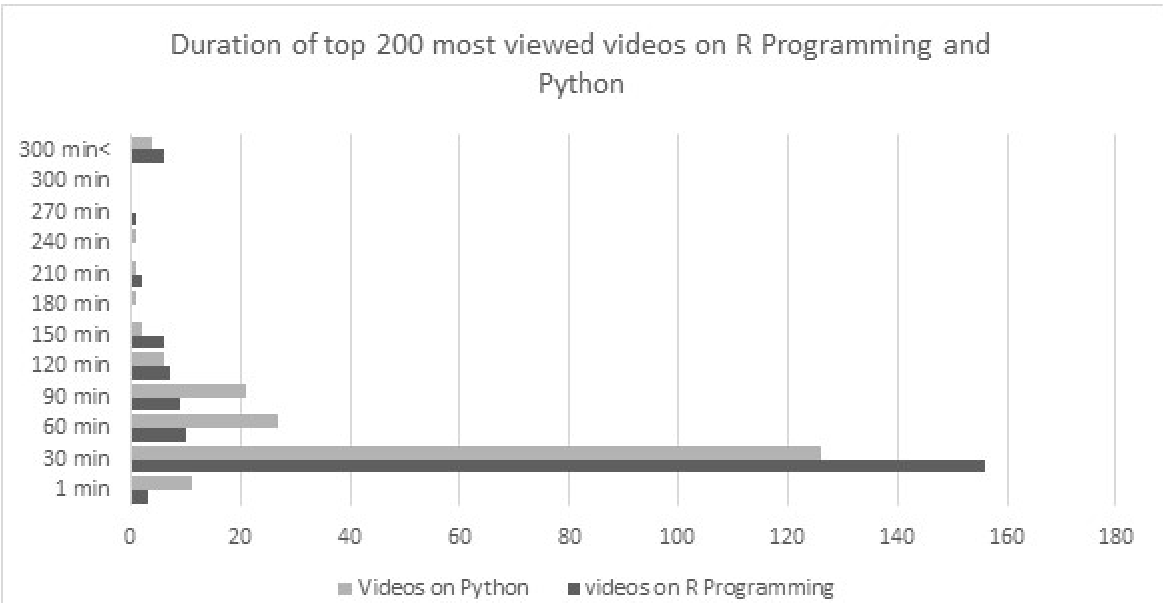



**Fig. 9. Duration of top 200 most viewed videos on R Programming and Python**

## 6.5. Sentiment Analysis: Comments Received for Python and R programming Videos

Social media empowers the public to express their opinions on a topic, product, content, service, and so on in the form of reviews or comments, which can be used for future improvement. In this study total number of comments extracted from R and Python was 35078 and 51321 respectively, but for analysis 14018 and 22246 were taken from the respective topics which were in English language.

Analysis of the comments extracted from R Programming related videos, demonstrates (figure 10) that most of the viewers are happy and satisfied with the content because maximum comments (25%) are showing their positive attitude towards the videos, while 7% are negative and 13% are from the neutral category.

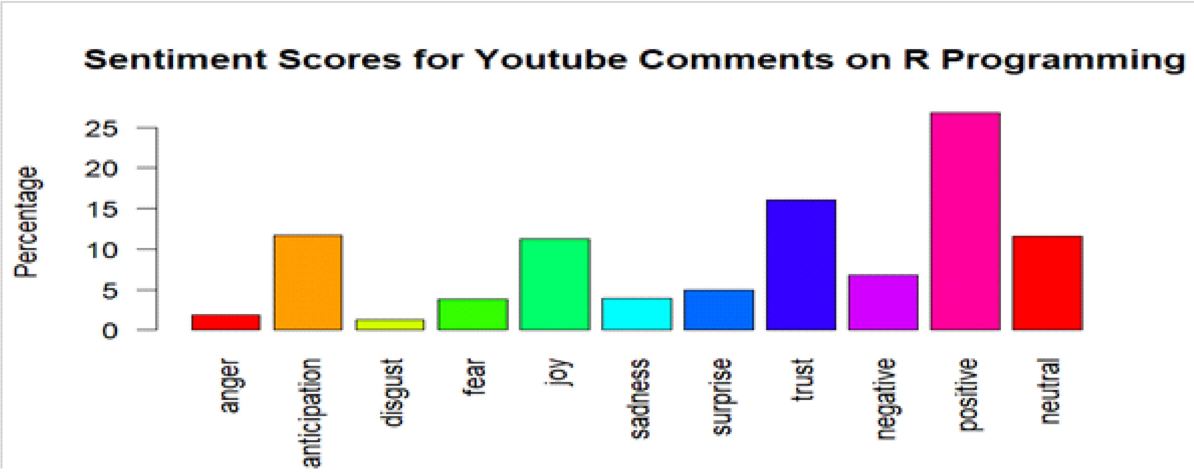



Fig. 10. Sentiment Analysis of YouTube Comments on R Programming

The analysis given in figure 11 demonstrates that Python videos have received a positive (24.8%) attitude from viewers they are happy and satisfied with the content whereas 8% are negative and 14% are from the neutral category.

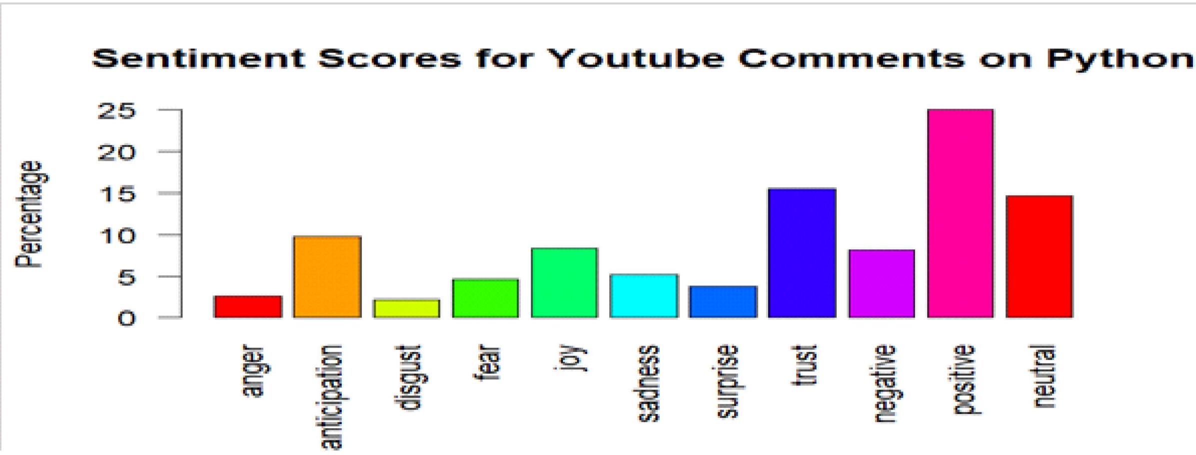



Fig. 11. Sentiment analysis of YouTube comments on Python

## 7. Findings

R Programming is an older concept than Python as seen from the data received by Google trends for YouTube search and data extracted from the YouTube platform. There is more content available on R programming than python but view count and like count of Python are higher than R programming.

In R Programming, only 7.79% and for Python 44.61 % of videos have been seen from zero to hundred times (less time viewed) respectively. In this category majority of videos were recently uploaded on YouTube in 2022 for R programming but in Python the videos uploaded are from the year 2020. It is also observed that for both topics more videos (86% in R programming and 91% in Python) got less view count than their mean value and less videos (only 14% in R programming and 9% in Python) got more view count than its mean value. Videos that have been seen more than one lakh times (got good View count) are quite few 7.3 percent in R Programming and 9 percent in Python. R programming videos from 2013 and 2017 and in Python from 2019 videos received maximum View count.

Most videos for both topics have been liked between 0 to 100 times. It was also found that less videos on R programming (0.7%) and for Python (2.6 %) have been liked more than ten thousand times. The duration of most of the uploaded videos for both topics is of 30 minutes.

The study also clears that View count has a strong correlation with like count and comment count but no significant correlation with upload year of video (Publish at) and duration.

Sentiment polarity of extracted comments of the videos shows that most of the viewers appreciated and liked the videos for both topics, and were positive towards the uploaded content.

## 8. Conclusion

This work highlights and presents a detailed description of various characteristics of YouTube videos on R programming and Python. The researchers examined 539 videos on R programming and 529 videos on Python to analyze the publication distribution of videos, the average duration of the video, engagement matrices (i.e., views, likes, comments) and sentiment analysis of comments on videos which aids in the development of a video performance indicator by assessing the interest and perception of viewers of that video.

According to the analysis, R programming has more videos compared to Python, whereas Python has a higher view count and like count. Videos with a Creative Commons license are increasingly prevalent in Python. After sentiment analysis, it is discovered that 25% of comments for R programming and Python are positive. The study also demonstrated that whether a video is old or new, it will be viewed by people provided the content is intriguing. View count is co-related with like count and comment count but it has no relation with age and duration.

It is observed that creator's time, effort, and money are invested in the creation of these videos, so this type of research allows them to quantify the ROI, allows them to analyze the trend and anticipate the popularity of the uploaded material. As more research is undertaken by LIS researchers using R and Python programming languages, this study provides valuable contribution to emerging research domains.

## References


1. **Al-Tamimi, A. K. et al.** (2017). Arabic sentiment analysis of YouTube comments. 2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies, AEECT 2017, 2018-Janua (January 2018), 1–6. https://doi.org/10.1109/AEECT.2017.8257766

2. **Barjasteh, I. et al.** (2014). Trending Videos/ : Measurement and Analysis. *ArXiv E-Prints,* September, n.p.

3. **Carlozzi, M.** (2019). Factor Analysis For Librarians in R. Code4Lib Journal, 46. https://journal.code4lib.org/articles/14821

4. **Che, X. et al.** (2015). A Survey of Current YouTube Video Characteristics. *IEEE Multimedia,* 22, 2, 56–63. https://doi.org/10.1109/MMUL.2015.34

5. **Chang, H., et al.** (2021). Supervised Machine Learning and Deep Learning Classification Techniques to Identify Scholarly and Research Content. In 2021 Systems and Information Engineering Design Symposium (SIEDS). 1-6.

6. **Cheng, X. et al.** (2008). Statistics and social network of YouTube videos. IEEE International Workshop on Quality of Service, IWQoS, 229–238. https://doi.org/10.1109/IWQOS.2008.32

7. **Choi, Y.** (2019). Finding "just right" books for children: analyzing sentiments in online book reviews. *Electronic Library,* 37, 3, 563–576. https://doi.org/10.1108/EL-01-2019-0018.

8. **Dascalu, C. G. et al.** (2021). Blended learning - the efficiency of video resources

and youtube in the modern dental education. *Revista De Cercetare Si Interventie Sociala,* 72, 288–310. https://doi.org/https://doi.org/10.33788/rcis.72.18.

9. **Deori, M. et al.** (2021). Analysis of YouTube video contents on Koha and DSpace, and sentiment analysis of viewers' comments. *Library Hi Tech,* October. https://doi.org/10.1108/LHT-12-2020-0323

10. **Goeser, P. T. et al.** (2012). MatLab Marina/ : Web-based tutorials for teaching programming concepts using MATLAB. In ASEE Southern Section Conference.

11. **Gupta, H. et al.** (2017). Multimedia tool as a predictor for social media advertising- a YouTube way. *Multimedia Tools and Applications,* 76, 18, 18557-18568.

12. **Gupta, N. and Chakravarty, R.** (2021). Research visualization of Indian LIS research using VOSviewer and Bibliometrix. *Library Hi Tech News,* 38, 8, 6–8. https://doi.org/10.1108/LHTN-10-2021-0076

13. **Han, X.** (2020). Evolution of research topics in LIS between 1996 and 2019: An analysis based on latent Dirichlet allocation topic model. *Scientometrics,* 125, 3, 2561-2595.

14. **Jackson, D. M.** (2017). Establishing an Instructor YouTube Channel as an Open Educational Resource (OER) supplementing general and organic chemistry Courses. In M. A. Christiansen; & J. M. Weber2 (Eds.), Teaching and the Internet: The Application of Web Apps, Networking, and Online Tech for Chemistry Education (pp. 115–135). American Chemical Society. https://doi.org/DOI: 10.1021/bk-2017-1270

15. **Jansen, B. J. and Salminen, J.** (2017). Viewed by too many or vewed too little/ : Using information dissemination for audience segmentation. 80th Annual Meeting of the Association for Information Science & Technology, Washington, DC, VA. https://doi.org/10.1002/pra2.2017.1450 5401021

16. **Kabir, A. I. et al.** (2020). Word cloud and sentiment analysis of Amazon Earphones Reviews with R programming language. *Informatica Economica,* 24, 4, 55–71. https://doi.org/10.24818/issn14531305/24.4.2020.05

17. **Kadriu, A. et al.** (2020). Investigating trends in learning programming using YouTube tutorials. *International Journal of Learning and Change,* 12, 2, 190–208. https://doi.org/10.1504/IJLC.2020.106721.

18. **King, D. L.** (2015). Analytics, Goals, and Strategy for Social Media. *Library Technology Reports,* 51, 1, 26–33.

19. **Liu, B.** (2012). Sentiment Analysis: A Fascinating Problem. In Sentiment Analysis and Opinion Mining (pp. 7–8). Morgan & Claypool Publishers, Kentfield, CA. https://doi.org/10.1142/9789813100459_0007.

20. **Lund, B. D.** (2020). Assessing library topics using sentiment analysis in R: a discussion and code sample. *Public Services Quarterly,* 16, 2, 112-123

21. **Ma, J. and Lund, B.** (2020). The evolution of LIS research topics and methods from 2006 to 2018: A content analysis. *Proceedings of the Association for Information Science and Technology,* 57, 1, 1–10. https://doi.org/10.1002/pra2.241.

22. **Madden, A. et al.** (2013). A classification scheme for content analyses of YouTube video comments. *Journal of Documentation,* 69, 5, 693–714. https://doi.org/10.1108/JD-06-2012-0078

23. **Manca, S. and Ranieri, M.** (2017). Implications of social network sites for teaching and learning. Where we are and where we want to go. *Education and Information Technologies,* 22, 2, 605–622. https://doi.org/10.1007/s10639-015-9429-x

24. **Mandal, N., Das, A., Monda, D., & Das, S.** (2021). Analysis of the literature growth and usability of YouTube videos related to Moodle. *Library Philosophy and Practice,* 2021(October).

25. **Neumann, M. M. and Herodotou, C.** (2020). Young Children and YouTube: A global phenomenon. *Childhood Education,* 96, 4, 72–77. https://doi.org/10.1080/00094056. 2020. 1796459

26. **Pallavi et al.** (2019). Case study on Vidya-Mitra (e-Learning MHRD Initiative): A quantitative analysis on multimedia content (Video). CALIBER Conference Procedings, 2019, 114–123. https://ir.inflibnet.ac.in/handle/1944/2342

27. **Parabhoi, L and Chand, P.** (2018). Content analysis of Youtube videos related Drupal, Joomla And Wordpress 11Th Coventional PLANNER-2018 Proceding, 10–16.

28. **Serdaroglu, E.** (2020). Exploring the Use of Youtube by Symphonic Orchestras as An Educational Platform During the Pandemic of Covid-19. *European Journal of Social Science Education and Research,* 7, 3, 59–66.

29. **Siersdorfer, S. et al.** (2010). How useful are your comments? Analyzing and predicting YouTube comments and comment ratings. Proceedings of the 19th International Conference on World Wide Web, WWW '2010, 891–900. https://doi.org/10.1145/ 1772690. 1772781

30. **Singh, A. K. and Mahawar, K. L.** (2020). Online studies with YouTube. *Library Progress (International),* 40, 2, 231–235. https://doi.org/10.5958/2320-317x.2020. 00026.4

31. **Temban, M. et al.** (2021). Exploring informal learning opportunities via youtube kids among children during COVID-19. *Academic Journal of Interdisciplinary Studies,* 10, 3, 272–287. https://doi.org/ 10.36941/AJIS-2021-0083

32. **Wikipedia** (2022). https://en.wikipedia.org/ wiki/YouTube.

33. **Yaacob, Z. et al.** (2020). Acceptance of YouTube as a Learning Platform during the Covid-19 Pandemic: The Moderating Effect of Subscription Status. *TEM Journal,* 9, 4, 1732–1739. https://doi.org/10.18421/ TEM94-54

34. **Zhou, X. and Ordonez, C.** (2021). Programming Languages in Data Science: A aomparison from a database angle. In 2021 IEEE International Conference on Big Data (Big Data),15-18 December 2021, (pp. 3147-3154). IEEE. https://ieeexplore.ieee.org/ abstract/document/9672007 (accessed on 25 May 2022)