

Trending_Videos_On_Youtube

Bill_Wu

03/10/2022

```
library(tidyverse)
library(ggplot2)
library(utils)

youtube_data <- read_csv("data/Trending_videos_on_youtube_dataset.csv")

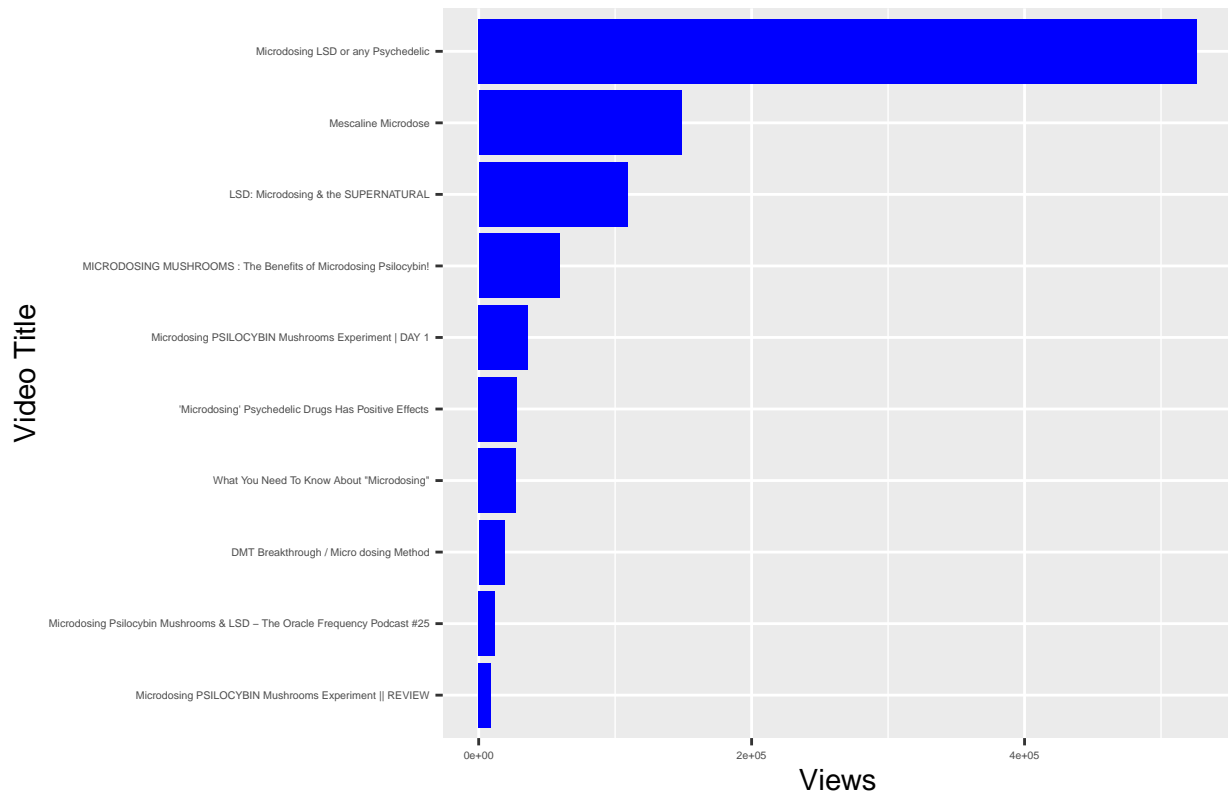
top10video <- youtube_data %>%
  arrange(desc(viewCount)) %>%
  head(10)

ggplot(data = top10video, aes(x = reorder(videoTitle, viewCount), y = viewCount)) +
  geom_bar(stat = "identity", fill = "blue") +
  coord_flip() +
  ggtitle("Top 10 Most Viewed Videos") +
  ylab("Views") +
  xlab("Video Title") +

  # changing all labels font
  #theme(text = element_text(size = 8)) +

  theme(axis.text = element_text(size = 4))
```

Top 10 Most Viewed Videos

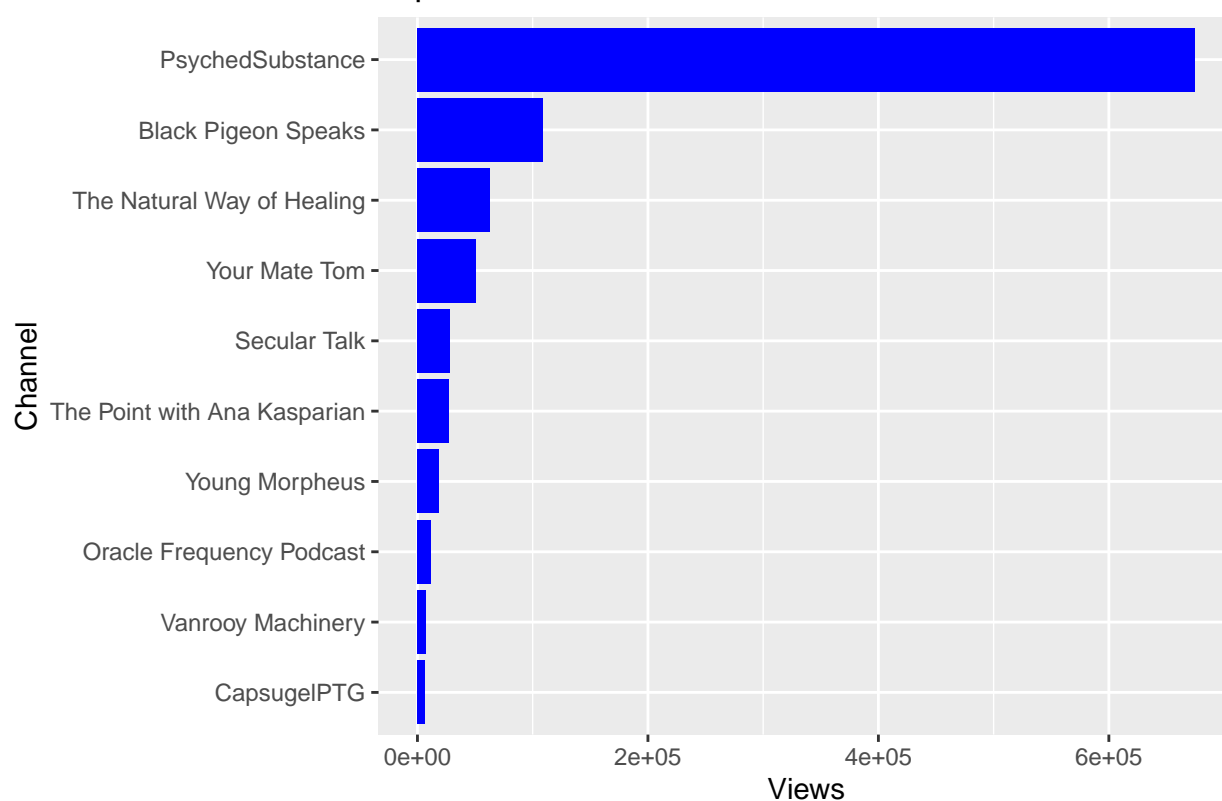


```
youtube_data1 <- group_by(youtube_data, channelTitle)

top10channel <- summarise(youtube_data1, TotalViews = sum(viewCount)) %>%
  arrange(desc(TotalViews)) %>%
  head(10)

ggplot(data = top10channel, aes(x = reorder(channelTitle, TotalViews), y = TotalViews)) +
  geom_bar(stat = "identity", fill = "blue") +
  coord_flip() +
  ggtitle("Top 10 Most Viewed Channel") +
  ylab("Views") +
  xlab("Channel")
```

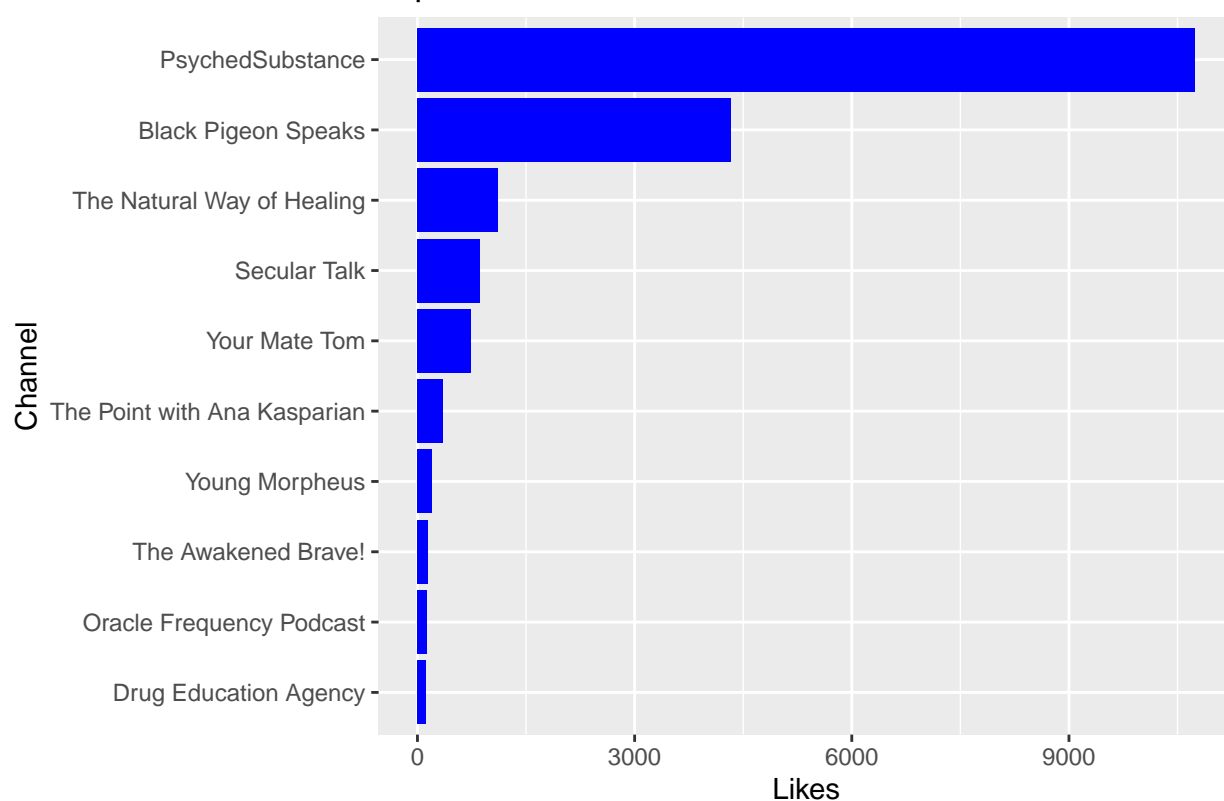
Top 10 Most Viewed Channel



```
top10likechannel <- summarise(youtube_data1, TotalLike = sum(likeCount)) %>%
  arrange(desc(TotalLike)) %>%
  head(10)

ggplot(data = top10likechannel, aes(x = reorder(channelTitle, TotalLike), y = TotalLike)) +
  geom_bar(stat = "identity", fill = "blue") +
  coord_flip() +
  ggtitle("Top 10 Most Liked Channel") +
  ylab("Likes") +
  xlab("Channel")
```

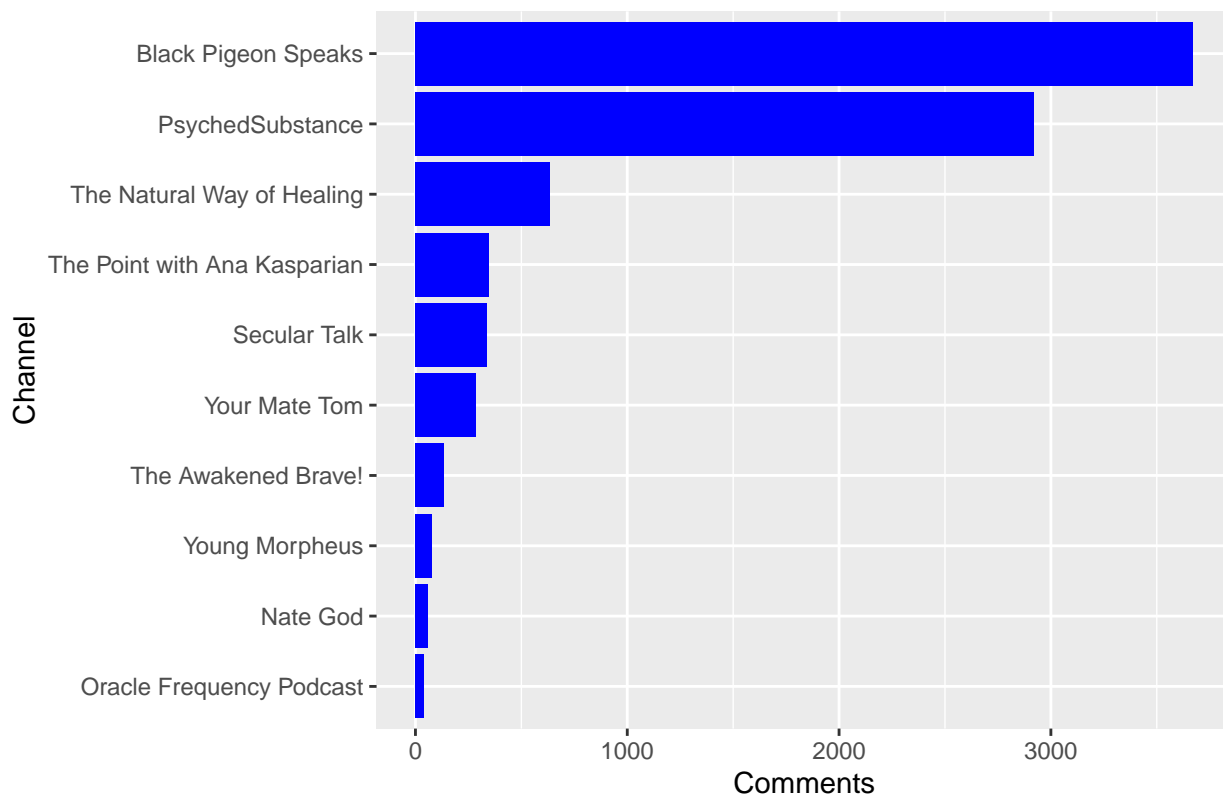
Top 10 Most Liked Channel



```
top10commentchannel <- summarise(youtube_data1, TotalComment = sum(commentCount)) %>%
  arrange(desc(TotalComment)) %>%
  head(10)

ggplot(data = top10commentchannel, aes(x = reorder(channelTitle, TotalComment), y = TotalComment)) +
  geom_bar(stat = "identity", fill = "blue") +
  coord_flip() +
  ggtitle("Top 10 Most Comment Channel") +
  ylab("Comments") +
  xlab("Channel")
```

Top 10 Most Comment Channel

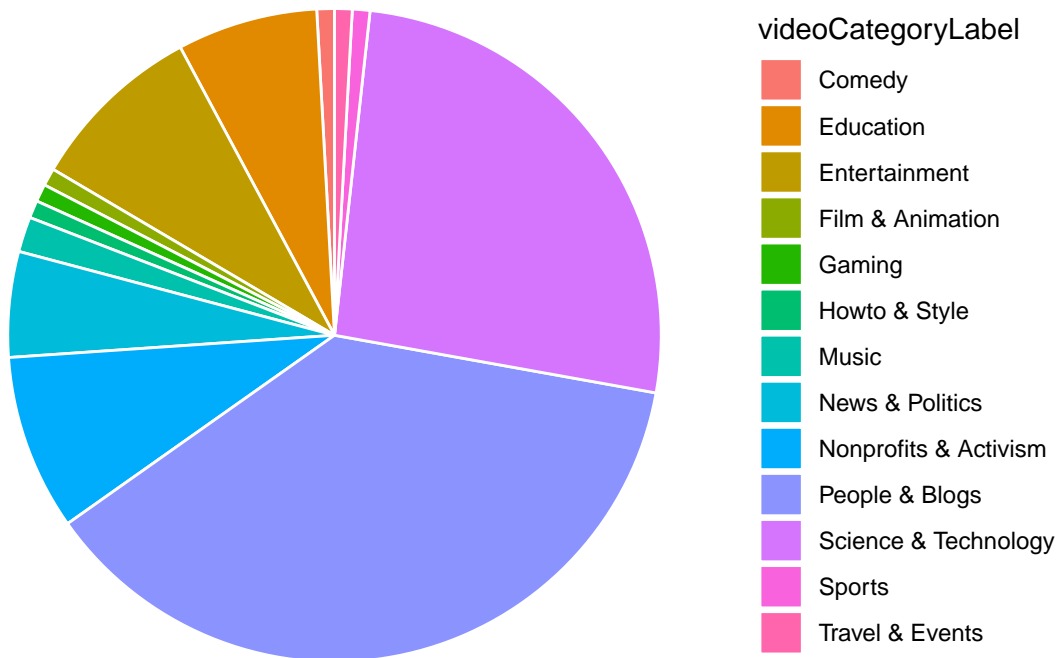


PsychedSubstance has the most views, likes and most viewed video in their channel. However, Black Pigeon Speaks seems to have a higher number in comments compared to the channel PsychedSubstance. We think the reason could be the difference in their channel category where Black Pigeon Speaks is a news&politics channel whereas PsychedSubstance is a Science&technology/entertainment channel.

```
youtube_Category <- youtube_data %>% count(videoCategoryLabel)

ggplot(youtube_Category, aes(x="", y=n, fill=videoCategoryLabel))+
  geom_bar(stat="identity", width = 1, color="white")+
  coord_polar("y", start = 0) +
  theme_void() +
  ggtitle("All Posted Videos in Different Categories")
```

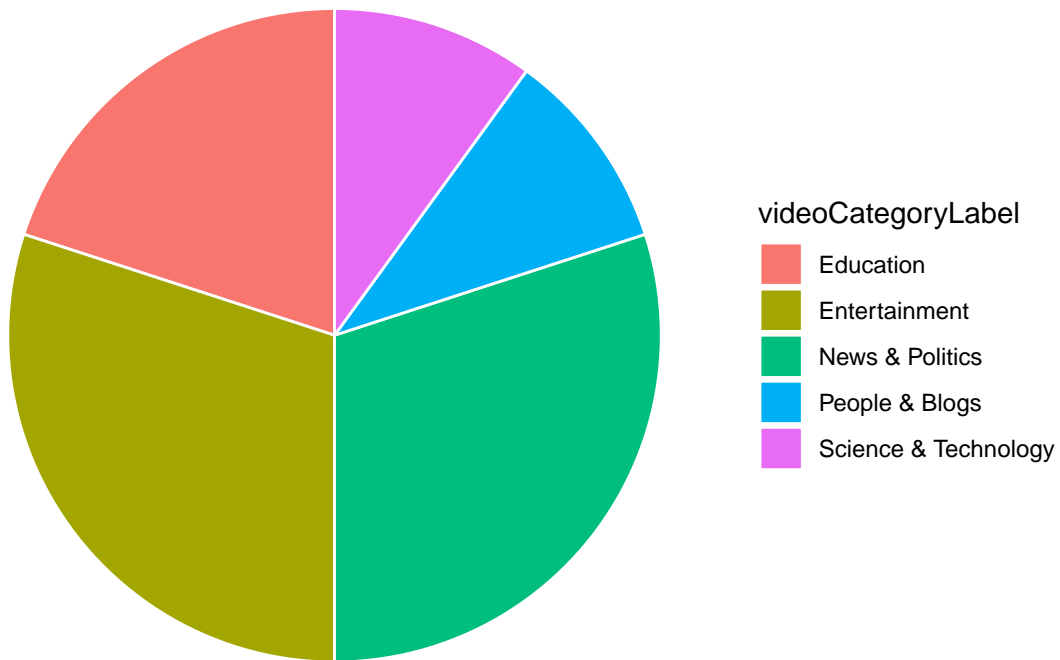
All Posted Videos in Different Categories



```
youtube_top10Category <- top10video %>% count(videoCategoryLabel)

ggplot(youtube_top10Category, aes(x="", y=n, fill=videoCategoryLabel))+
  geom_bar(stat="identity", width = 1, color="white")+
  coord_polar("y", start = 0) +
  theme_void() +
  ggtitle("Top 10 Most Viwed Videos in Different Categories")
```

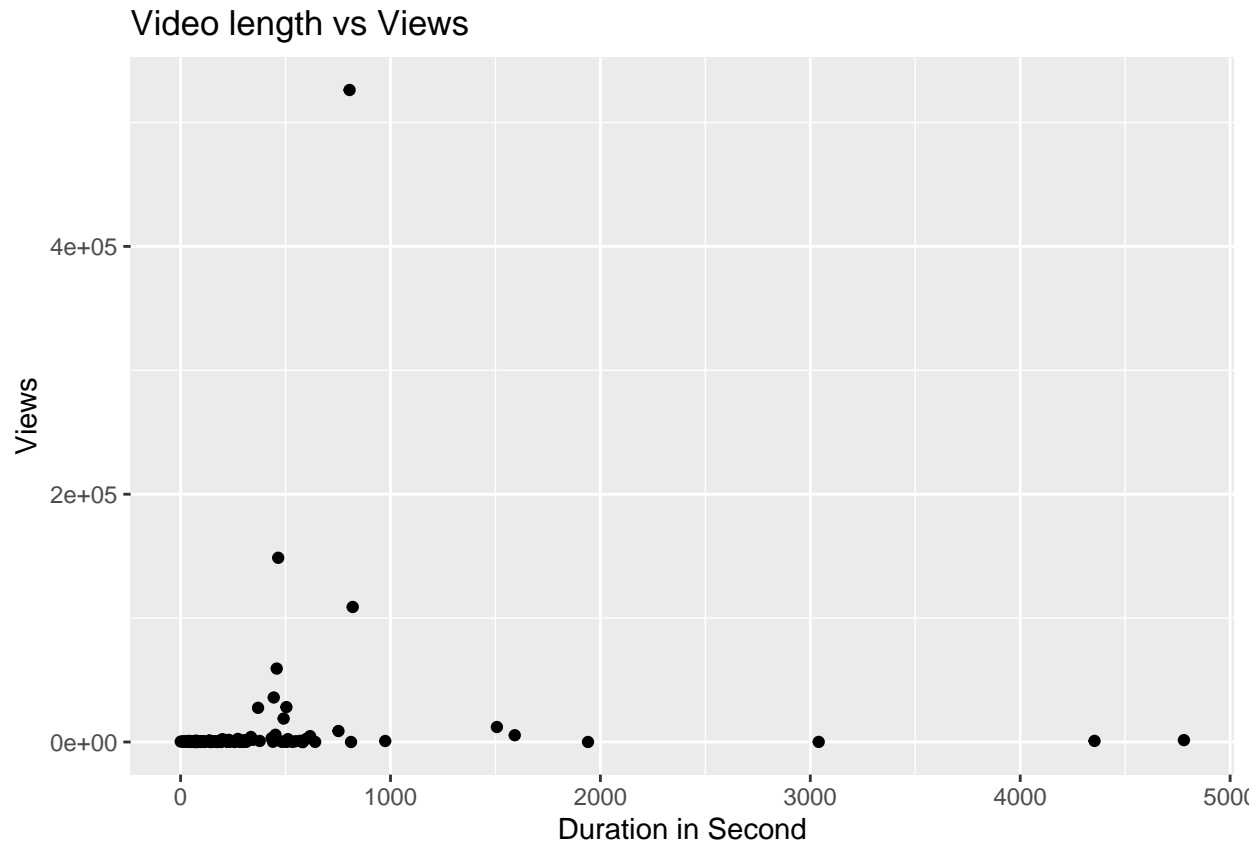
Top 10 Most Viwed Videos in Different Categories



Even though most videos are about People&Blogs or science&technology, only a small proportion got into the top 10 most viewed videos.

Q1: how does the duration of the video influences the views of the video?

```
graph1 <- ggplot(youtube_data, aes(x = durationSec, y = viewCount )) +  
  geom_point() +  
  ggtitle("Video length vs Views") +  
  xlab("Duration in Second") +  
  ylab("Views")  
graph1
```



As we can see from the chart, most viewed videos lie in between 400 seconds to 800 seconds. This is approximately 7 minutes to 13 minutes. According to narakeet, the average Youtube video length in 2022 is 11.7 minutes. (“What Is the Ideal Youtube Video Length in 2022?” 2022)

References

“What Is the Ideal Youtube Video Length in 2022?” 2022. *Narakeet*. <https://www.narakeet.com/blog/ideal-youtube-video-length.html#:~:text=Some%20creators%20have%20found%20success,%2D15%20minutes%20perform%20better.%E2%80%9D>.