#### **INTRODUCTION:**

This report intends to find the Best videos based on various criterias, posted in the video sharing website giant – YouTube and create visualizations based on the trending videos list maintained by them.

## **DESCRIPTION OF DATASETS:**

The dataset was sourced from the website Kaggle - <a href="https://www.kaggle.com/datasnaek/youtube-new">https://www.kaggle.com/datasnaek/youtube-new</a>

Multiple youtube video statistics datasets of many countries are available. This dataset includes several months (and counting) of data on daily trending YouTube videos from the years 2005 to 2012.

For the purpose of this dataset, I have chosen three different datasets of the following three countries and combined them into a single data frame:

- India
- Canada
- USA

The data also includes a category\_id field, which varies between different countries. To retrieve the categories for a specific video, there are different JSON files for each country. So, I have downloaded and combined the JSON file as well, for the three countries of interest.

## **Pre-processing**

Extensive data cleaning, wrangling, filtering, grouping were performed on the dataset to get the data in the desired format. The details of the data cleaning part are included in the Appendix section along with the code and appropriate comments explaining each step that were done.

#### User

I am a Data Scientist working in YouTube – the world's largest video sharing website. I have been asked to find out statistics and insights based on the list of trending videos maintained by the organization. wants to launch its first ever awards ceremony

#### **Problem**

The dataset maintained by YouTube has the attributes such as the date of the video when it was trending, title of the video, total number of views garnered, likes on the video, counts of dislikes, total number of comments on the video, category, channel that posted this video. This information is maintained separately for each country in the world.

A good visualization on the dataset will help to find out the videos that were trending based on the number of views, likes, comments, the channel that was viewed the most, etc.

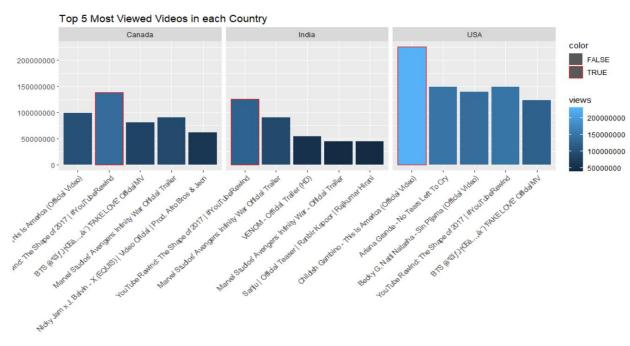
#### Scenario

The organization wants to launch its first ever awards ceremony and wants to find out the best videos till date in the countries Canada, India and USA. Awards will be given out based on many categories such as: The most viewed video till date, most liked video till date, most commented video till date, best channel that garnered the most views, the same videos that were in trending for many days.

For the purpose of this report, I will be showing three visualizations and tables that were created by me, which were also showcased in the YouTube awards ceremony.

### **Visualisation:**

## **First Visualisation:**



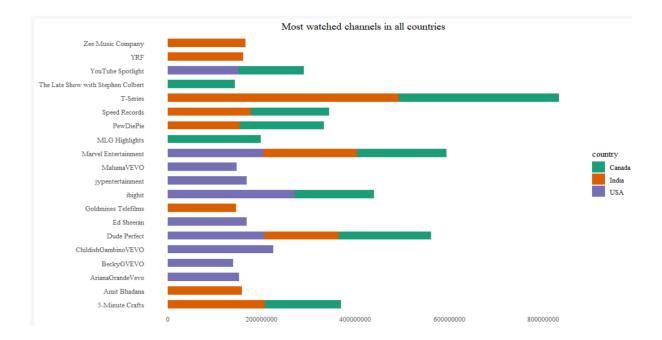
This visualisation intends to find the top 5 most viewed videos from every country and to pick the number one most viewed video among them. The total number of views of every video were taken into consideration and the top 5 most viewed videos were plotted.

As seen in the visualisation, the tallest bar graph and the red outline indicates the most viewed video of all times, from each country.

The most viewed video in USA is "**This is America**" and the award was given away to this video.

The most viewed video from both Canada and India are "YouTube rewind, #Shape of 2017" video. Thus, two awards were given away to the same video.

### Second visualisation:



This visualisation was done by grouping channels in each country and by calculating the total sum of views, each channel received. The top 10 most viewed channels from every country were taken and formed into a same data frame.

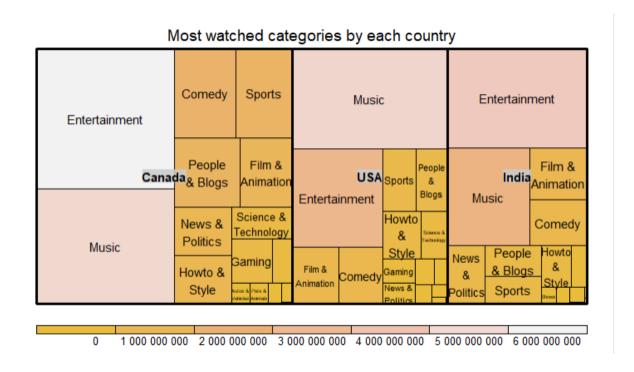
This is an useful graph, which shows the difference in countries based on the colors and by stacking the countries one upon the other, if the same channels were viewed in the other countries as well. In some cases, a channel was mostly viewed by a single country alone. That information is also included in this country.

From the plot, it is extremely clear that the channel "T-Series" garnered the most number of views, from the countries India and Canada combined, and thus bagging the title "Best Channel Award – T-series".

Some interesting observations can be found here, that the channels "Marvel Entertainment" and "Dude Perfect" were viewed in all three countries.

## **Third Visualisation:**

This visualisation was created as more of an effort to get an insight, to see which category of video was most viewed in each country. This was created by grouping the dataset based on the category title watched in every country and getting the total sum of the views in that category title from that country.



This is a treemap plot, which shows the most viewed category of video watched in every country. The area size and color of every category is to denote the number of views they receive.

Based on the color and size, we can clearly deduct that "Entertainment" in Canada, is the most viewed category of all. It is the only category with white color, indicating its significance very easily.

In USA, the category "Music" is most watched.
In India too, the category "Entertainment" is most watched.

# **General Insights:**

# Table with most likes and comments in Canada:

| TITLE  | CHANNEL TITLE          | LIKES                 | COMMENTS             |
|--|------------------------|-----------------------|----------------------|
| BTS (ë°©if"소ë"단) 'FAKE LOVE' Official MV                                       | ibighit                | 5053338               | 1114800              |
| Childish Gambino - This Is America (Official Video)                            | ChildishGambinoVEVO    | 3037318               | 31 <mark>9502</mark> |
| YouTube Rewind: The Shape of 2017   #YouTubeRewind                             | YouTube Spotlight      | 3014479               | 817582               |
| Marvel Studios' Avengers: Infinity War Official Trailer                        | Marvel Entertainment   | 2606665               | 347982               |
| BTS (ë°@ $f_n$ i†Œë,ë $\varsigma$ ") 'MIC Drop (Steve Aoki Remix)' Official MV | ibighit                | 25 <mark>42863</mark> | 519092               |
| Drake - God's Plan   | DrakeVEVO              | 24 <mark>69057</mark> | 1582 <mark>62</mark> |
| BTS (ë°©íf"소ë"ëç¨) LOVE YOURSELF 轉 Tear 'Singularity' Comeback Trailer         | ibighit                | 24 <mark>07419</mark> | 3 <mark>40125</mark> |
| j-hope 'Daydream (ë°±ì¼ë³½)' MV  | ibighit                | 23 <mark>92594</mark> | 437036               |
| Ariana Grande - No Tears Left To Cry   | ArianaGrandeVevo       | 2195120               | 192 <mark>685</mark> |
| Maroon 5 - Girls Like You ft. Cardi B  | Maroon5VEVO            | 2178332               | 1288 <mark>10</mark> |
| BTS (ë°©íf"소ë"ë<¨) 'FAKE LOVE' Official MV                                     | ibighit                | 5053338               | 1114800              |
| YouTube Rewind: The Shape of 2017   #YouTubeRewind                             | YouTube Spotlight      | 3014479               | 817582               |
| BTS (ë°@ $f_n$ i†Œë,ë $\varsigma$ °) 'MIC Drop (Steve Aoki Remix)' Official MV | ibighit                | 25 <mark>42863</mark> | 519092               |
| CHá° Y NGAY ÄI   RUN NOW   SÆ N TÙNG M-TP   Official Music Video               | Sơn Tùng M-TP Official | 827026                | 445251               |
| j-hope 'Daydream (ë°±ì¼ë³½)' MV  | ibighit                | 23 <mark>92594</mark> | 437036               |
| Melting Every Lipstick From Sephora Together                                   | Safiya Nygaard         | 335949                | 3 <mark>49112</mark> |
| Marvel Studios' Avengers: Infinity War Official Trailer                        | Marvel Entertainment   | 26,06665              | 347982               |
| BTS (방탄소년ëç¨) LOVE YOURSELF 轉 Tear 'Singularity' Comeback Trailer              | ibighit                | 24 <mark>07419</mark> | 3 <mark>40125</mark> |
| Childish Gambino - This Is America (Official Video)                            | ChildishGambinoVEVO    | 3037318               | 31 <mark>9502</mark> |
| we broke up  | David Dobrik           | 196 <mark>7910</mark> | 3 <mark>11525</mark> |

# Table with most likes and comments in India:

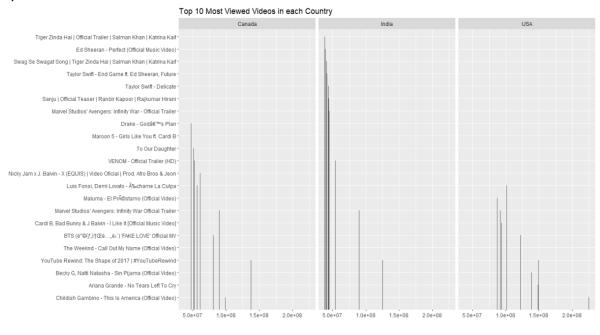
| ПП.Е   | CHANNEL TITLE               | LIKES                 | COMMENTS             |
|--|-----------------------------|-----------------------|----------------------|
| YouTube Rewind: The Shape of 2017   #YouTubeRewind   | YouTube Spotlight           | 2912710               | 807558               |
| Marvel Studios' Avengers: Infinity War Official Trailer  | Marvel Entertainment        | 2606663               | 347982               |
| Taylor Swift - End Game ft. Ed Sheeran, Future   | TaylorSwiftVEVO             | 1804377               | 1469 <mark>17</mark> |
| Ed Sheeran - Perfect (Official Music Video)  | Ed Sheeran                  | 1721384               | 903 <mark>5</mark> 2 |
| Taylor Swift - Delicate  | TaylorSwiftVEVO             | 1658756               | 1485 <mark>48</mark> |
| Marvel Studios' Avengers: Infinity War - Official Trailer  | Marvel Entertainment        | 14 <mark>94997</mark> | 186 <mark>005</mark> |
| VENOM - Official Trailer (HD)  | Sony Pictures Entertainment | 1250535               | 1354 <mark>05</mark> |
| Selena Gomez - Back To You   | SelenaGomezVEVO             | 1101522               | 6415 <mark>6</mark>  |
| Ed Sheeran - Happier (Official Video)  | Ed Sheeran                  | 1095526               | 5894 <mark>8</mark>  |
| BB Ki Vines-   The Sacrifice   | BB Ki Vines                 | 1084173               | 1257 <mark>40</mark> |
| YouTube Rewind: The Shape of 2017   #YouTubeRewind   | YouTube Spotlight           | 2912710               | 807558               |
| OnePlus 6 Top Features and GIVEAWAY $\delta \ddot{Y}"\mbox{$\Psi$}-$ OnePlus 6 Avengers Edition Giveaway!! $\delta \ddot{Y}"\mbox{$\Psi$}$ | Technical Guruji            | 37053                 | 382685               |
| Samsung Galaxy J6 Unboxing and Giveaway ōŸ"¥ōŸ"¥ōŸ"¥   | Technical Guruji            | 40277                 | 377912               |
| Marvel Studios' Avengers: Infinity War Official Trailer  | Marvel Entertainment        | 2606663               | 347982               |
| Samsung Galaxy S9 Top Features and Tips Tricks - Galaxy S9 Mega Giveaway ðŸ"¥ðŸ"¥ðŸ"¥  | Technical Guruji            | 252602                | 2 <mark>95139</mark> |
| Xiaomi Redmi Note 5 Pro Unboxing and Giveaway ōŸ"¥ōŸ"¥ōŸ"¥   | Technical Guruji            | 237090                | 192 <mark>229</mark> |
| Marvel Studios' Avengers: Infinity War - Official Trailer  | Marvel Entertainment        | 14 <mark>94997</mark> | 186 <mark>005</mark> |
| OnePlus 5T Star Wars Limited Edition Unboxing and First Look *GIVEAWAY*  | Technical Guruji            | 173149                | 153049               |
| Logan Paul - SANTA DISS TRACK (Official Music Video)   | Logan Paul Vlogs            | 882821                | 1508 <mark>56</mark> |
| Taylor Swift - Delicate  | TaylorSwiftVEVO             | 1658756               | 1485 <mark>48</mark> |

# Table with most likes and comments in USA:

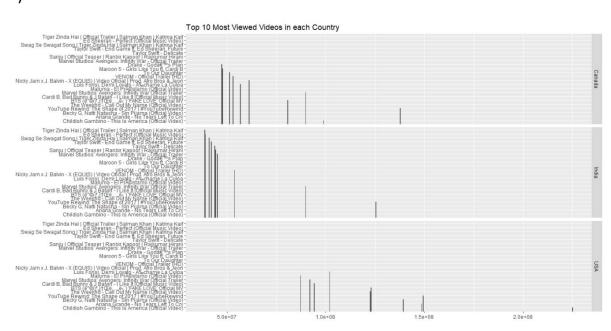
| TITLE   | CHANNEL TITLE        | LIKES                 | COMMENTS             |
|---|----------------------|-----------------------|----------------------|
| BTS (방탄소년ë∢") 'FAKE LOVE' Official MV   | ibighit              | 5613827               | 1228655              |
| Childish Gambino - This Is America (Official Video)                                     | ChildishGambinoVEVO  | 5023450               | 517232               |
| Ariana Grande - No Tears Left To Cry  | ArianaGrandeVevo     | 3094021               | 242                  |
| YouTube Rewind: The Shape of 2017   #YouTubeRewind                                      | YouTube Spotlight    | 3093544               | 810698               |
| BTS (ë°©í $f$ "소ë"ë<br>ć") 'MIC Drop (Steve Aoki Remix)' Official MV                    | ibighit              | 2729292               | 546100               |
| BTS (ë°©í $f_n$ 소ë,ë $\epsilon$ '') LOVE YOURSELF 轉 Tear 'Singularity' Comeback Trailer | ibighit              | 2700800               | 37 <mark>1864</mark> |
| j-hope 'Daydream (ë°±i1⁄4ë <sup>s</sup> ½)' MV  | ibighit              | 2672431               | 477233               |
| Marvel Studios' Avengers: Infinity War Official Trailer                                 | Marvel Entertainment | 26 <mark>25661</mark> | 35 <mark>0458</mark> |
| Maroon 5 - Girls Like You ft. Cardi B   | Maroon5VEVO          | 24 <mark>88565</mark> | 142410               |
| Luis Fonsi, Demi Lovato - Échame La Culpa   | LuisFonsiVEVO        | 2376636               | 13422                |
| So Sorry.   | Logan Paul Vlogs     | 1402578               | 1361580              |
| BTS (ë°©í $f_*$ 소ë,ë $\epsilon$ °) 'FAKE LOVE' Official MV                              | ibighit              | 5613827               | 1228655              |
| YouTube Rewind: The Shape of 2017   #YouTubeRewind                                      | YouTube Spotlight    | 3 <mark>093544</mark> | 810698               |
| Suicide: Be Here Tomorrow.  | Logan Paul Vlogs     | 198 <mark>8746</mark> | 658130               |
| BTS (ë°©í $f_n$ i†Œë,ë<br>ë') 'MIC Drop (Steve Aoki Remix)' Official MV                 | ibighit              | 2729292               | 546100               |
| Childish Gambino - This Is America (Official Video)                                     | ChildishGambinoVEVO  | 5023450               | 517232               |
| j-hope 'Daydream (ë°±i¼ë <sup>8</sup> ½)' MV  | ibighit              | 2672431               | 477233               |
| LOGAN PAUL IS BACK!   | Logan Paul Vlogs     | 1216340               | 4 <mark>36698</mark> |
| Melting Every Lipstick From Sephora Together  | Safiya Nygaard       | 402569                | 38 <mark>3762</mark> |
| BTS (ë°©í $f_n$ 소ë,eć°) LOVE YOURSELF 轉 Tear 'Singularity' Comeback Trailer             | ibighit              | 2700800               | 37 <mark>1864</mark> |

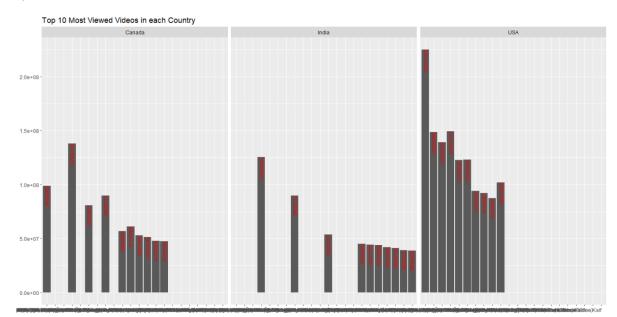
## **Previous Iterations of first visualisation:**

1)

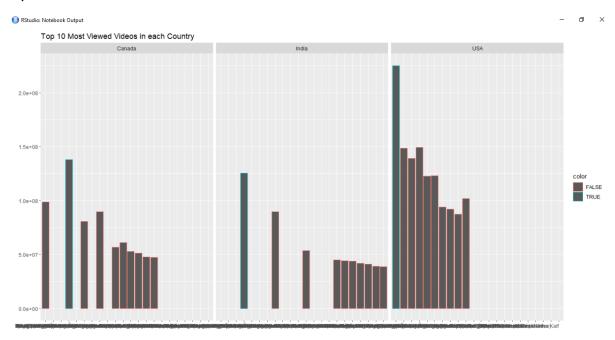


2)



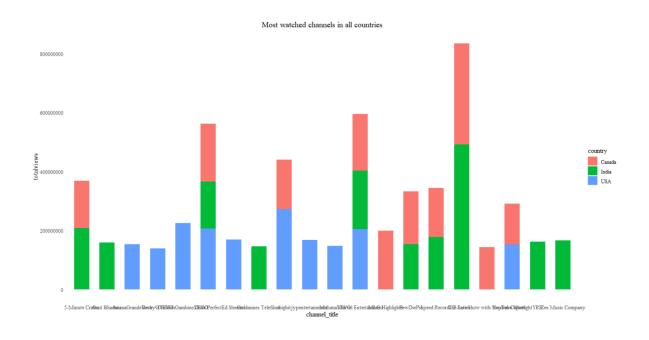






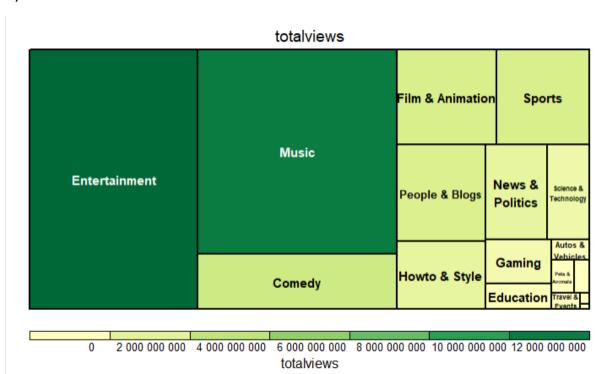
## **Previous Iterations of second visualisation:**

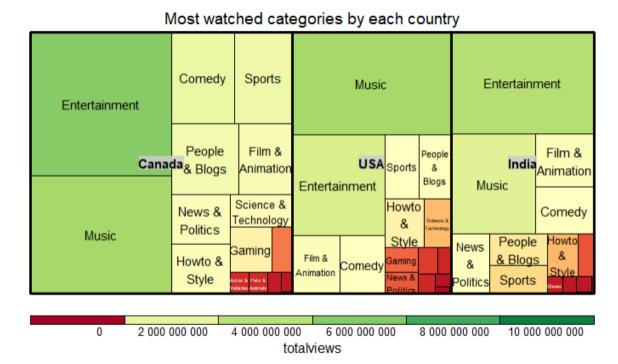
1)



## **Previous Iterations of third visualisation:**

1)





## **Appendices with Code, Comments:**

## **Libraries used:**

library(ggthemes)

library(treemap)

library(rjson)

library(jsonlite)

library(formattable)

library(sqldf)

library(data.table)

library(pastecs)

library(dplyr)

library(DT)

library(ggplot2)

library(plotrix)

library(tidytext)

library(stringr)

library(tm)

## Pre-processing data and cleaning, wrangling:

```
#Loading the dataset
canadaData <- read.csv("CAvideos.csv",header = TRUE)</pre>
usaData <- read.csv("USvideos.csv",header = TRUE)
indiaData <- read.csv("INvideos.csv",header = TRUE)</pre>
#Checking for NA values
sum(is.na(canadaData))
sum(is.na(usaData))
sum(is.na(indiaData))
#No NA values
#Getting the Category ID and Category title from a separate json file
CanadaCategory <- fromJSON("CA category id.json", flatten = TRUE)
CanadaCategorydf <- as.data.frame(CanadaCategory)</pre>
CanadaCategorydf <- CanadaCategorydf[,c("items.id","items.snippet.title")]
USCategory <- from JSON ("US category id. json", flatten = TRUE)
USCategorydf <- as.data.frame(USCategory)
USCategorydf <- USCategorydf[,c("items.id","items.snippet.title")]
IndiaCategory <- fromJSON("India cat id.json", flatten = TRUE)</pre>
IndiaCategorydf <- as.data.frame(IndiaCategory)</pre>
IndiaCategorydf <- IndiaCategorydf[,c("items.id","items.snippet.title")]</pre>
#Renaming category id and category title
colnames(CanadaCategorydf)[1] <- "category id"
colnames(USCategorydf)[1] <- "category id"
colnames(IndiaCategorydf)[1] <- "category_id"</pre>
colnames(CanadaCategorydf)[2] <- "category title"
colnames(USCategorydf)[2] <- "category title"
colnames(IndiaCategorydf)[2] <- "category title"
```

```
#Converting category id from the json file to integer format CanadaCategorydf$category_id <- as.integer(CanadaCategorydf$category_id) IndiaCategorydf$category_id <- as.integer(IndiaCategorydf$category_id) USCategorydf$category_id <- as.integer(USCategorydf$category_id)
```

#Joining category title from the json file to the main data set based on category id

```
Canada <- left_join(canadaData, CanadaCategorydf, by="category_id")
USA <- left_join(usaData, USCategorydf, by="category_id")
India <- left_join(indiaData, IndiaCategorydf, by="category_id")
```

#Checking for NA values after combining the two data frames. There are NA values in two data frames.

```
sum(is.na(Canada)) #69 NA values
sum(is.na(USA)) #0
sum(is.na(India)) #41 NA values
```

#Removing rows that have NA values, because we don't have the information of category title for that category id and we cannot replace it with any other values, because these are absolute values.

```
Canada <- na.omit(Canada, Canada $ category_title)
India <- na.omit(India, India $ category_title)
```

#Removing rows that have comments disabled, ratings disabled and videos that were removed

Canada <-

```
subset(Canada,(toupper(Canada$video_error_or_removed)=="FALSE" &
toupper(Canada$ratings_disabled) == "FALSE" &
toupper(Canada$comments disabled) == "FALSE"))
```

#ADDING A NEW COLUMN COUNTRY WITH VALUE CANADA Canada\$country<-"Canada"

```
India <- subset(India,(toupper(India$video_error_or_removed)=="FALSE" &
toupper(India$ratings_disabled) == "FALSE" &
toupper(India$comments_disabled) == "FALSE"))</pre>
```

#ADDING A NEW COLUMN COUNTRY WITH VALUE INDIA India\$country <-"India"

USA <- subset(USA,(toupper(USA\$video\_error\_or\_removed)=="FALSE" & toupper(USA\$ratings\_disabled) == "FALSE" & toupper(USA\$comments\_disabled) == "FALSE"))

#ADDING A NEW COLUMN COUNTRY WITH VALUE USA USA\$country <- "USA"

#Getting the unique video sequence id, title based on the latest trending date/maximum number of views. The same video's id sometimes appears more than once, because it was trending on more than one day.

CanadaViews <- arrange(Canada, desc(Canada\$views))

USAViews <- arrange(USA, desc(USA\$views))

IndiaViews <- arrange(India, desc(India\$views))

UniqueCanadaViews <- CanadaViews[match(unique(CanadaViews\$video\_id), CanadaViews\$video\_id),]

UniqueUSAViews <- USAViews[match(unique(USAViews\$video\_id), USAViews\$video\_id),]

UniqueIndiaViews <- IndiaViews[match(unique(IndiaViews\$video\_id), IndiaViews\$video\_id),]

#TOP 10 VIEWS, LIKES, COMMENTS OF CANADA
CanadaLikes <- arrange(UniqueCanadaViews, desc(UniqueCanadaViews\$likes))

```
CanadaComments <- arrange(UniqueCanadaViews,
desc(UniqueCanadaViews$comment count))
UniqueCanadaViews1 <- head(UniqueCanadaViews,10)</pre>
TopCanadaLikes<- head(CanadaLikes,10)
TopCanadaComments<-head(CanadaComments,10)
#Creating a dataset by combining top likes and comments
TopVideosCanada <-
as.data.frame(rbind(TopCanadaLikes,TopCanadaComments))
#TOP 10 OF USA
USALikes <- arrange(UniqueUSAViews, desc(UniqueUSAViews$likes))
USAComments <- arrange(UniqueUSAViews,
desc(UniqueUSAViews$comment_count))
UniqueUSAViews1 <- head(UniqueUSAViews,10)</pre>
UniqueUSALikes1<- head(USALikes,10)</pre>
UniqueUSAComments1<-head(USAComments,10)
TopVideosUSA <- rbind(UniqueUSALikes1,UniqueUSAComments1)</pre>
TopVideosUSATable = TopVideosUSA[, c(3, 4, 9, 11)]
colnames(TopVideosUSATable) = c("TITLE", "CHANNEL
TITLE","LIKES","COMMENTS")
#Creating a table with all the views, comments and likes and varying the
colours based on their counts
widgetUSA = formattable(TopVideosUSATable, list(
      LIKES = color_bar('red'),
      COMMENTS = color_bar('darkblue')
))
```

```
#TOP 10 OF INDIA
IndiaLikes <- arrange(UniqueIndiaViews, desc(UniqueIndiaViews$likes))
IndiaComments <- arrange(UniqueIndiaViews,
desc(UniqueIndiaViews$comment count))
UniqueIndiaViews1 <- head(UniqueIndiaViews,10)
UniqueIndiaLikes1<- head(IndiaLikes,10)
UniqueIndiaComments1<-head(IndiaComments,10)
TopVideosIndia <- rbind(UniqueIndiaLikes1,UniqueIndiaComments1)
TopVideosIndiaTable = TopVideosIndia[, c(3, 4, 9, 11)]
colnames(TopVideosIndiaTable) = c("TITLE", "CHANNEL
TITLE","LIKES","COMMENTS")
#Creating a table with all the views, comments and likes and varying the
colours based on their counts
widgetIndia = formattable(TopVideosIndiaTable, list(
      LIKES = color bar('green'),
      COMMENTS = color_bar('orange')
))
```

#This is done so that scientific notation representation will not appear options(scipen = 999)

## Code for visualisation 1 with iterations:

#COMBINING ALL THE RECORDS OF THE MOST VIEWED VIDEOS FROM EVERY COUNTRY

TopViewedVideos<-

rbind(head(UniqueUSAViews1,5),head(UniqueCanadaViews1,5),head(UniqueIn diaViews1,5))

```
#This is done so that, the most viewed video and its title will be highlighted
with red colour separately
topViewed <- TopViewedVideos %>%
 group by(country) %>%
 mutate(color = max(views) == views)
#First visualization
ggplot(topViewed, aes(reorder(title,-views),views,fill=views))+
facet_grid(~country,scales = "free")+ geom_bar(stat="identity",aes(color =
color))+theme(axis.text.x = element text(angle = 45,hjust = 1))
+scale color manual(values = c(NA, "red"))+ labs(title="Top 5 Most Viewed
Videos in each Country")+xlab(NULL)+ylab(NULL)
#previous iterations
#1
ggplot(topViewed, aes(y=reorder(title,-views),x=views))+
facet_grid(country~.)+ geom_bar(stat="identity") + labs(title="Top 10 Most
Viewed Videos in each Country")
#2
ggplot(topViewed, aes(y=reorder(title,-views),x=views))+ facet_grid(~country)+
geom bar(stat="identity") + labs(title="Top 10 Most Viewed Videos in each
Country")+xlab(NULL)+ylab(NULL)
#3
ggplot(topViewed, aes(reorder(title,-views),views))+ facet_grid(~country)+
geom bar(stat="identity")+theme(axis.text.x = element text(hjust = 1)) +
labs(title="Top 10 Most Viewed Videos in each
Country")+xlab(NULL)+ylab(NULL)+ geom text(aes(label = views,
```

## Code for visualisation 2:

#Grouping by based on best channels - canada

color=I(red'), size = 3, hjust =1,angle=90)

```
cangroup <- group by(UniqueCanadaViews, channel title,country)</pre>
channelCanada <- summarise(cangroup, totalviews = sum(views)) %>%
arrange(desc(totalviews)) %>% head(10)
#Grouping by based on best channels - usa
usagroup <- group_by(UniqueUSAViews, channel_title,country)</pre>
channelUSA <- summarise(usagroup, totalviews = sum(views)) %>%
arrange(desc(totalviews)) %>% head(10)
#Grouping by based on best channels - india
Indiagroup <- group_by(UniqueIndiaViews, channel_title,country)</pre>
channelIndia <- summarise(Indiagroup, totalviews = sum(views)) %>%
arrange(desc(totalviews)) %>% head(10)
#Combining all the 10 best channels of each country to give us a spread-out
bar plot
TopChannels <- rbind(channelCanada,channelIndia,channelUSA)
ggplot(TopChannels, aes(x = channel_title, y = totalviews, fill = country)) +
geom_bar(stat = "identity", width = .6) + coord_flip() + labs(title="Most
watched channels in all countries") + theme tufte() + theme(plot.title =
```

element text(hjust = .5),axis.ticks = element blank()) +

scale\_fill\_brewer(palette = "Dark2") +xlab(NULL) +ylab(NULL)

## #Iteration

```
ggplot(TopChannels, aes(x = channel_title, y = totalviews, fill = country)) +
geom_bar(stat = "identity", width = .6) + labs(title="Most watched channels in
all countries") + theme(plot.title = element_text(hjust = .5),axis.ticks =
element_blank()) + scale_fill_brewer(palette = "Accent")
```

### Code for visualisation 3:

#Combining ALL the top viewed videos (with unique video-id) from every country

countries <- rbind(UniqueCanadaViews, UniqueUSAViews, UniqueIndiaViews)

#Grouping it based on the category title watched in every country and getting the total sum of the views in that category title from that country

BestCategory <- group\_by(countries, category\_title, country) %>%

summarise(totalviews = sum(as.double(views)))

```
treemap(BestCategory,
    index=c("country", "category_title"),
    vSize="totalviews",
    vColor="totalviews",
    type="value", title.legend = "",
    title="Most watched categories by each country",
    format.legend = list(scientific = FALSE, big.mark = " "),
    palette=terrain.colors(7))
```

### #Iterations

```
#1)
treemap(BestCategory,
    index=c("category_title"),
    vSize="totalviews",
    vColor="totalviews",
    type="value",
    format.legend = list(scientific = FALSE, big.mark = " "))
```

```
#2)
treemap(BestCategory,
   index=c("country", "category_title"),
   vSize="totalviews",
   vColor="totalviews",
   type="value",
   title="Most watched categories by each country",
   format.legend = list(scientific = FALSE, big.mark = " "),
    palette="RdYlGn",
    range=c(0,1100000000),
    mapping=c(0, 600000000, 11000000000))
Code for creating tables of most liked and commented videos:
Canada:
#Creating a dataset by combining everything
TopVideosCanada <-
as.data.frame(rbind(TopCanadaLikes,TopCanadaComments))
TopVideosCanadaTable = TopVideosCanada[, c(3, 4, 9, 11)]
colnames(TopVideosCanadaTable) = c("TITLE", "CHANNEL
TITLE","LIKES","COMMENTS")
#Creating a table with all the views, comments and likes and varying the
colours based on their counts
Canada = formattable(TopVideosCanadaTable, list(
      LIKES = color bar('red'),
      COMMENTS = color bar('orange')
```

))

```
India:
```

```
TopVideosIndia <- rbind(UniqueIndiaLikes1,UniqueIndiaComments1)</pre>
```

```
TopVideosIndiaTable = TopVideosIndia[, c(3, 4, 9, 11)] colnames(TopVideosIndiaTable) = c("TITLE", "CHANNEL TITLE", "LIKES", "COMMENTS")
```

#Creating a table with all the views, comments and likes and varying the colours based on their counts

## **USA:**

TopVideosUSA <- rbind(UniqueUSALikes1,UniqueUSAComments1)</pre>

```
TopVideosUSATable = TopVideosUSA[, c(3, 4, 9, 11)] colnames(TopVideosUSATable) = c("TITLE", "CHANNEL TITLE", "LIKES", "COMMENTS")
```

#Creating a table with all the views, comments and likes and varying the colours based on their counts