BA 64060 assignment1

2023-09-08

Download Reference

The data set have been downloaded from kaggle
https://www.kaggle.com/datasets/nelgiriyewithana/global-youtube-statistics-2023

#load the dataset In order to import the dataset in RStudio, I used the read.csv function.

```
data =
read.csv("C:/Users/gdurg/Downloads/Global_YouTube_Statistics.csv",header =
TRUE, sep = ",")

data = na.omit(data)

head(data, 10)
```

```
##    rank                   Youtuber subscribers  video.views       category
## 1     1                   T-Series   245000000 228000000000          Music
## 3     3                    MrBeast   166000000  28368841870  Entertainment
## 4     4 Cocomelon - Nursery Rhymes   162000000 164000000000      Education
## 5     5                  SET India   159000000 148000000000          Shows
## 9     9                Like Nastya   106000000  90479060027 People & Blogs
## 10   10              Vlad and Niki    98900000  77180169894  Entertainment
## 11   11          Zee Music Company    96700000  57856289381          Music
## 12   12                        WWE    96000000  77428473662         Sports
## 14   14                  BLACKPINK    89800000  32144597566 People & Blogs
## 16   16                   Sony SAB    83000000 101000000000          Shows
##                         Title uploads       Country Abbreviation
channel_type
## 1                    T-Series   20082         India           IN
Music
## 3                     MrBeast     741 United States           US
Entertainment
## 4  Cocomelon - Nursery Rhymes     966 United States           US
Education
## 5                   SET India  116536         India           IN
Entertainment
## 9            Like Nastya Vlog     493        Russia           RU
People
## 10              Vlad and Niki     574 United States           US
Entertainment
## 11          Zee Music Company    8548         India           IN
Music
## 12                        WWE   70127 United States           US
```

```
Sports
## 14                  BLACKPINK     543   South Korea           KR
Music
## 16                   Sony SAB   71270         India           IN
Entertainment
##    video_views_rank country_rank channel_type_rank
## 1                 1            1                1
## 3                48            1                1
## 4                 2            2                1
## 5                 3            2                2
## 9               630            5               25
## 10                8            5                6
## 11               12            3                2
## 12                7            6                1
## 14               32            1                3
## 16                4            5                7
##    video_views_for_the_last_30_days lowest_monthly_earnings
## 1                        2258000000                 564600
## 3                        1348000000                 337000
## 4                        1975000000                 493800
## 5                        1824000000                 455900
## 9                          48947000                  12200
## 10                        580574000                 145100
## 11                        803613000                 200900
## 12                        714614000                 178700
## 14                        498930000                 124700
## 16                       1657000000                 414300
##    highest_monthly_earnings lowest_yearly_earnings highest_yearly_earnings
## 1                   9000000                6800000                108400000
## 3                   5400000                4000000                 64700000
## 4                   7900000                5900000                 94800000
## 5                   7300000                5500000                 87500000
## 9                    195800                 146800                  2300000
## 10                  2300000                1700000                 27900000
## 11                  3200000                2400000                 38600000
## 12                  2900000                2100000                 34300000
## 14                  2000000                1500000                 23900000
## 16                  6600000                5000000                 79600000
##    subscribers_for_last_30_days created_year created_month created_date
## 1                       2000000         2006           Mar           13
## 3                       8000000         2012           Feb           20
## 4                       1000000         2006           Sep            1
## 5                       1000000         2006           Sep           20
## 9                        100000         2016           Jan           14
## 10                       600000         2018           Apr           23
## 11                      1100000         2014           Mar           12
## 12                       600000         2007           May           11
## 14                       700000         2016           Jun           29
## 16                      1100000         2007           Aug            4
##    Gross.tertiary.education.enrollment.... Population Unemployment.rate
```

```
## 1                                    28.1 1366417754              5.36
## 3                                    88.2  328239523             14.70
## 4                                    88.2  328239523             14.70
## 5                                    28.1 1366417754              5.36
## 9                                    81.9  144373535              4.59
## 10                                   88.2  328239523             14.70
## 11                                   28.1 1366417754              5.36
## 12                                   88.2  328239523             14.70
## 14                                   94.3   51709098              4.15
## 16                                   28.1 1366417754              5.36
##    Urban_population Latitude Longitude
## 1         471031528 20.59368  78.96288
## 3         270663028 37.09024 -95.71289
## 4         270663028 37.09024 -95.71289
## 5         471031528 20.59368  78.96288
## 9         107683889 61.52401 105.31876
## 10        270663028 37.09024 -95.71289
## 11        471031528 20.59368  78.96288
## 12        270663028 37.09024 -95.71289
## 14         42106719 35.90776 127.76692
## 16        471031528 20.59368  78.96288
```

Descriptive Statistics

```
summary(data)
```

```
##       rank         Youtuber          subscribers         video.views
##  Min.   :  1.0   Length:588         Min.   : 12300000   Min.   :2.634e+03
##  1st Qu.:212.8   Class :character   1st Qu.: 14700000   1st Qu.:4.920e+09
##  Median :453.5   Mode  :character   Median : 18600000   Median :8.770e+09
##  Mean   :470.3                      Mean   : 24284524   Mean   :1.282e+10
##  3rd Qu.:722.2                      3rd Qu.: 26400000   3rd Qu.:1.512e+10
##  Max.   :995.0                      Max.   :245000000   Max.   :2.280e+11
##    category            Title               uploads           Country
##  Length:588         Length:588         Min.   :     1.0   Length:588
##  Class :character   Class :character   1st Qu.:   434.5   Class :character
##  Mode  :character   Mode  :character   Median :  1194.5   Mode  :character
##                                        Mean   : 13968.7
##                                        3rd Qu.:  3886.5
##                                        Max.   :301308.0
##  Abbreviation       channel_type       video_views_rank   country_rank
##  Length:588         Length:588         Min.   :      1   Min.   :   1.0
##  Class :character   Class :character   1st Qu.:    224   1st Qu.:  10.0
##  Mode  :character   Mode  :character   Median :    596   Median :  44.0
##                                        Mean   : 131375   Mean   : 185.1
##                                        3rd Qu.:   1622   3rd Qu.: 107.0
##                                        Max.   :4054962   Max.   :7683.0
##  channel_type_rank video_views_for_the_last_30_days
lowest_monthly_earnings
##  Min.   :   1.0    Min.   :3.000e+00               Min.   :     0
##  1st Qu.:  19.0    1st Qu.:4.867e+07               1st Qu.: 11500
```

```
##  Median :  49.0    Median :1.145e+08                Median : 28200
##  Mean   : 248.2    Mean   :2.450e+08                Mean   : 54352
##  3rd Qu.: 115.0    3rd Qu.:2.443e+08                3rd Qu.: 59025
##  Max.   :7670.0    Max.   :6.589e+09                Max.   :850900
##  highest_monthly_earnings lowest_yearly_earnings highest_yearly_earnings
##  Min.   :       0         Min.   :       0       Min.   :        0
##  1st Qu.:  183200         1st Qu.:  137400       1st Qu.:  2200000
##  Median :  451500         Median :  338650       Median :  5400000
##  Mean   :  868974         Mean   :  651725       Mean   : 10435433
##  3rd Qu.:  944300         3rd Qu.:  708225       3rd Qu.: 11300000
##  Max.   :13600000         Max.   :10200000       Max.   :163400000
##  subscribers_for_last_30_days  created_year  created_month
created_date
##  Min.   :      1              Min.   :1970   Length:588         Min.   :
1.00
##  1st Qu.: 100000              1st Qu.:2010   Class :character   1st Qu.:
8.00
##  Median : 200000              Median :2013   Mode  :character   Median
:15.00
##  Mean   : 357302              Mean   :2013                      Mean
:15.21
##  3rd Qu.: 400000              3rd Qu.:2016                      3rd
Qu.:22.25
##  Max.   :8000000              Max.   :2022                      Max.
:31.00
##  Gross.tertiary.education.enrollment....   Population
Unemployment.rate
##  Min.   :  7.60                           Min.   :2.025e+05   Min.   :
0.750
##  1st Qu.: 28.10                           1st Qu.:1.081e+08   1st Qu.:
5.360
##  Median : 60.00                           Median :3.282e+08   Median :
5.930
##  Mean   : 60.86                           Mean   :4.911e+08   Mean   :
8.934
##  3rd Qu.: 88.20                           3rd Qu.:3.282e+08   3rd
Qu.:14.700
##  Max.   :113.10                           Max.   :1.398e+09   Max.
:14.720
##  Urban_population       Latitude        Longitude
##  Min.   :    35588   Min.   :-38.42   Min.   :-172.105
##  1st Qu.: 61300442   1st Qu.: 20.59   1st Qu.: -95.713
##  Median :270663028   Median : 30.59   Median :  -3.593
##  Mean   :239499682   Mean   : 26.20   Mean   :  -6.382
##  3rd Qu.:270663028   3rd Qu.: 37.09   3rd Qu.:  78.963
##  Max.   :842933962   Max.   : 61.52   Max.   : 138.253
```

##Further we need to check quantative data– mean,median,min,max

```
mean(data$uploads)
```

```
## [1] 13968.69
```

```
median(data$uploads)
```

```
## [1] 1194.5
```

```
min(data$uploads)
```

```
## [1] 1
```

```
max(data$uploads)
```

```
## [1] 301308
```

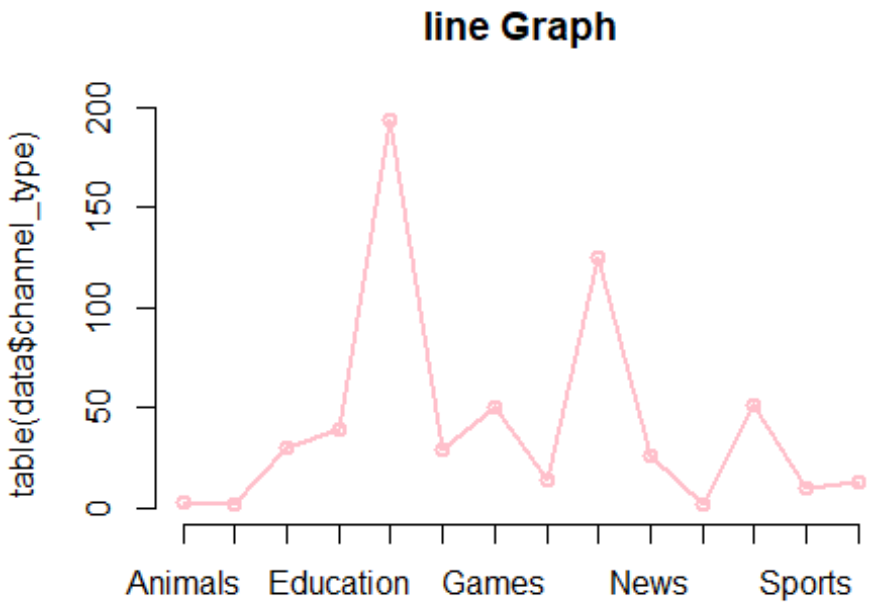
Transformation

```
log = log(data$video_views_rank)
head(log)
```

```
## [1] 0.0000000 3.8712010 0.6931472 1.0986123 6.4457198 2.0794415
```

Plot

```
plot(table(data$channel_type), type="o", col='pink',main="line Graph")
```




```
plot(x = data$uploads , y = data$video_views_rank,
     xlab = "UPLOADS",
     ylab = "VEDIO_VIEWS_RANK",
     xlim = c(0,1000),
     ylim = c(0,1000),
```

```
     main = "Youtube_stats"

)
```