EDA & Data Cleaning of Youtube Trending Videos

# Description

YouTube maintains a list of the top trending videos on the platform. This data set includes several months of data on daily trending YouTube videos.This list is determined by using user interactions such as views, comments and likes to identify which videos are user preferred and displays them on the trending page. Ranking of these videos is done based on a ratio of views, likes, comments and shares, in order to display the best videos at the top of the page

# Dataset

<https://www.kaggle.com/datasets/datasnaek/youtube-new>

# Attribute Info:

**No of Data Points :37352**

**No of Attributes: 16**

**Attribute Name**  **Data Type**

0 video\_id object

1 trending\_date object

2 title object

3 channel\_title object

4 category\_id int64

5 publish\_time object

6 tags object

7 views int64

8 likes int64

9 dislikes int64

10 comment\_count int64

11 thumbnail\_link object

12 comments\_disabled bool

13 ratings\_disabled bool

14 video\_error\_or\_removed bool

15 description object

.

# Packages

* Pandas
* Matplotlib
* Seaborn
* NLTK

## Observation of Analysis:

* Histogram
* Wordcloud
* Barplot
* Heatmap