**Python Problem Statement**

**Youtube Video Statistics**

**Abstract:**

YouTube (the world-famous video sharing website) maintains a list of the top trending videos on the platform. According to Variety magazine, to determine the year's top- trending videos, YouTube uses a combination of factors including measuring users interactions (number of views, shares, comments and likes). Note: that they’re not the most-viewed videos overall for the calendar year.

**Problem Statement:** Read the youtube data and perform exploratory data analysis.

**Dataset Information:**

This dataset is the daily record from the top trending YouTube videos. Top 200 trending videos of a given day. Original Data was collected during 14th November 2017 & 5th March 2018(though, data for January 10th & 11th of 2018 is missing). Original dataset was collected by Youtube API.

**Variable Description:**

|  |  |
| --- | --- |
| Column | Description |
| Video\_id | Unique Identity which tells the video\_id of each subscribed video |
| Category\_id | Unique number assigned to each category of the video |
| Channel\_title | Unique number assigned to each category of the video |
| Subscriber | Count of all subscribers for the respective video |
| Title | Title of the video |
| Tags | Tags are descriptive keywords you can add to your video to help viewers find your content |
| Description | Description of the respective video |

|  |  |
| --- | --- |
| Trend\_day\_count | Trending videos count for respective video |
| Tag\_count | Tag count for the respective video |
| Trend\_tag\_count | Tag count for respective trending video |
| Comment\_count | Count of comments for particular video |
| Comment\_disabled | It represents the boolean value. True represents comments are enabled and False represents comments are disabled |
| Like dislike disabled | It represents the boolean value. True represents comments are enabled and False represents comments are disabled |
| Likes | Number of likes for particular video |
| Dislikes | Number of dislikes for particular video |
| Tag appeared in title | It represents the boolean value. True represents that the respective tag has appeared in a particular video. False represents that the respective tag has not appeared in particular video. |
| Views | Target variable. Number of views for particular video |

**Scope:**

* Sentiment analysis in a variety of forms
* Categorising YouTube videos based on their comments and statistics.
* Training ML algorithms like RNNs to generate their own YouTube comments.
* Analysing what factors affect how popular a YouTube video will be.
* Statistical analysis over time