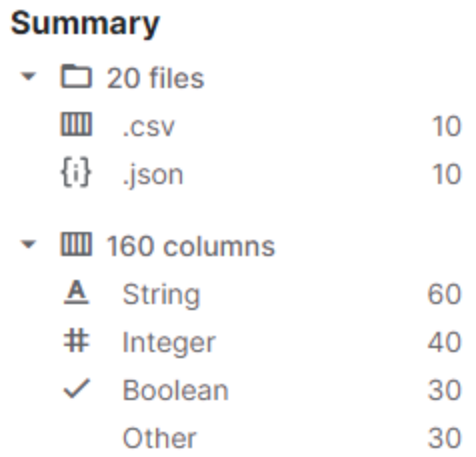
Project Proposal - Analysing YouTube Trending Videos

For this project we attempt to analyze what YouTube videos are trending. Find common characteristics of trending videos and attempt to predict and reproduce results. Our target clients are companies and/or individuals who are interested in using YouTube as an avenue for content creation.

This information could be useful for these content creators who are interested in increasing their view rates.

The data we will start with is from a kaggle dataset: [Trending YouTube Video Statistics]

We have a set of 10 CSV files and 10 JSON files



Initial exploration of the data columns give us a few ideas:

```
ctry = united_states
ctry.columns
```

```
Index(['video_id', 'trending_date', 'title', 'channel_title', 'category_id',
       'publish_time', 'tags', 'views', 'likes', 'dislikes', 'comment_count',
       'thumbnail_link', 'comments_disabled', 'ratings_disabled',
       'video_error_or_removed', 'description'],
      dtype='object')
```

We can find out which videos have the most views by using sort_values() to show which videos in descending order. We could also use value_counts() to show which category type has the most views. This will be useful for our client/s to guide them in creating content that will get more views from the data. We could see if there are correlations between views and numbers of likes and dislikes, comment count, etc … We could create some predictive models to test whether a particular video and/or video category will get many views. We could attempt to predict exactly how many average daily views a view will get.