

Script Analysis

1. Introduction

When it comes to posting video content, YouTube is the ultimate destination. In fact, at this point, YouTube is almost synonymous with internet video. YouTube has over 1.9 billion monthly active users and they watch a billion hours of video every day. To put that in perspective, that's over 114,000 years of video consumed EVERY DAY. Therefore, the impact of youtube should not be overlooked.

Here is the structure of data set:

```
Data columns (total 5 columns):
title          700 non-null object
url            700 non-null object
content        700 non-null object
views          700 non-null int64
posting_date   700 non-null object
```

We scraped 700 scripts of 700 videos from 16 different YouTube channels, the variables are titles, url, content, views, and postdate.

2. Statistical Data

For the 700 scripts(after removed stopwords), we got:

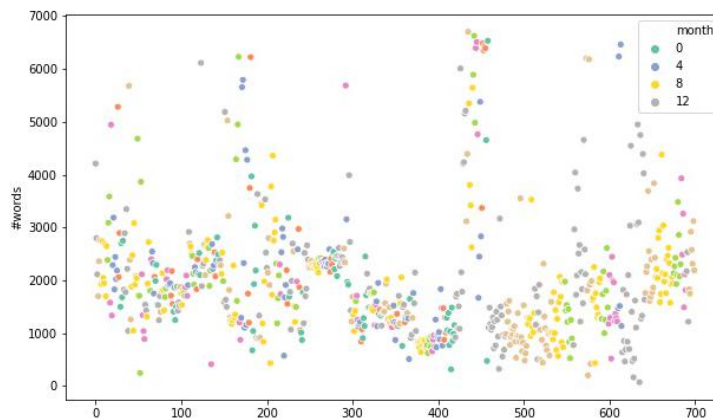
Total words	1409,748 words
Average words per video	2,013 words

After removed all the stopwords:

Total verbs	63,065 words	Average verbs	90 words
Total adjectives	114,657 words	Average adjectives	163 words
Total adverbs	96,778 words	Average adverbs	138 words
Total nouns	293,880 words	Average nouns	419 words

For our videos, the average words should be around 10,760 words. And normally, youtubers don't use too many verbs, but more nouns, which means the videos are more delivery information instead of interacting with audiences.

1. Total words distribution:



From the word distribution, we can see that most of the videos have 1000-3000 words that are meaningful(after remove stopwords).

2. Common verbs word cloud:



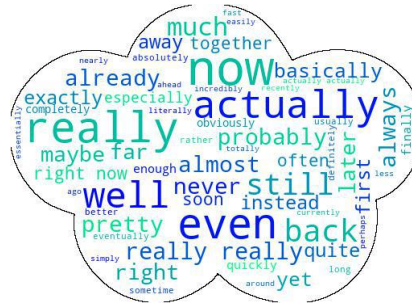
Verbs are most focused on positive verbs which can delivery information efficiently.

3. Common adjectives word cloud:



Adjectives for a great video to choose is more like much, many, great, good and also are more positive and active.

4. Common adverbs word cloud:

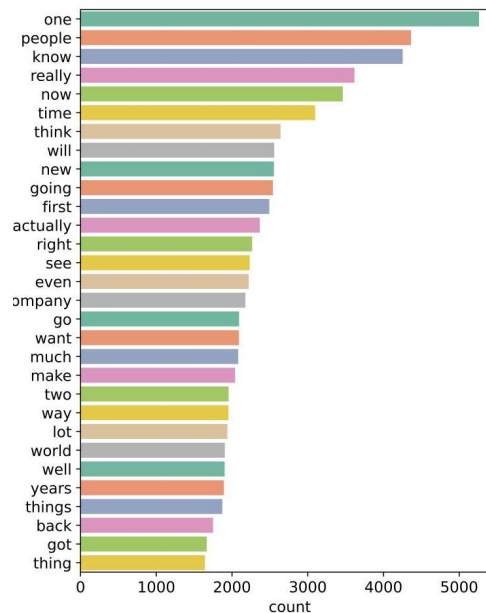


5. Common nouns word cloud:



For the nouns, we should choose those more related to people and thing. Those nouns are more relevant to people's interest which can catch audiences.

6. Word frequency:



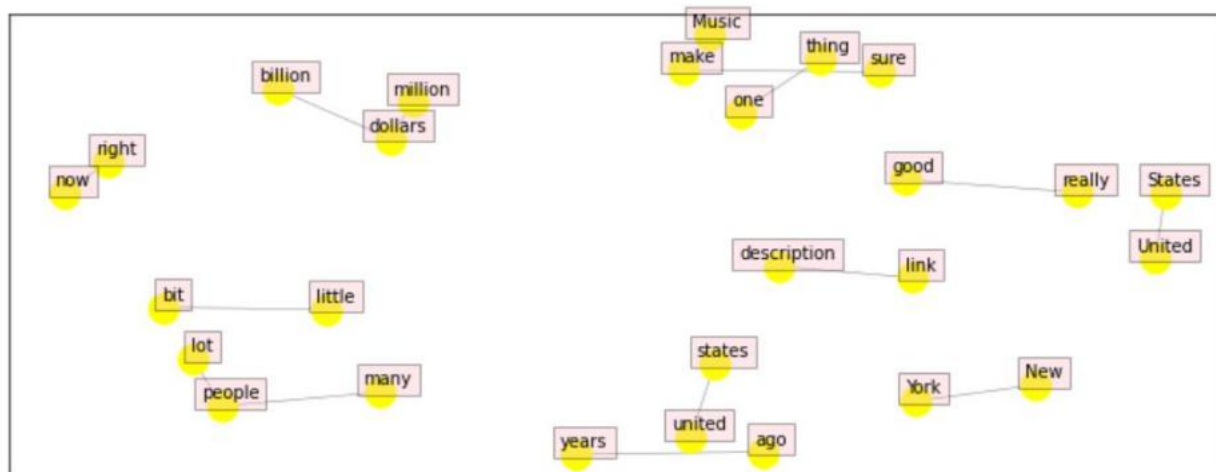
After remove all the stopwords, we got 679,749 words left. Among those words, we have common words that we may consider to use.

7. What, where, why, when, how:



For what, where, when, why, how, those words have a relative big proportion of a video, we should consider to use those words more to lead audiences.

8. Bigram Correlation:



3. Text Summary:

Because all the scripts that we have were not punctuated, so we need to add punctuation before process text summary using punctuator (Python API).

An example of the first video of coldfusion:

Original text:

"this video is brought to you by morning brew hi welcome to today's episode of cold fusion before i start this episode isn't a recommendation to buy any of apple's products in fact i'm a windows and android user myself the following is simply my documentation of what i perceive to be the biggest moment..."

After punctuation:

"this video is brought to you by morning, brew. hi. Welcome to today's episode of cold fusion. Before i start, this episode isn't a recommendation to buy any of apple's products. In fact, i'm a windows and android user myself....."

After text summary, we got:

"You again soon for the next episode cheers guys have a good one cold fusion, it's new thinking. The first thing you notice is just the scrubbing in the playback right, so this is a fanless teeny, tiny, iphone, 12 and just butter. Some of the complex instructions can take a dozen clock cycles or more to complete, though, contrary to popular belief, both kinds of processes were flourishing during the 1990s. That's intel the computer inside [, Music] in 1991 apple, ibm and motorola combined forces to create the power pc chip. The iphone launched in 2007 using a simple instruction arm cpu if intel's ceo, paul ottolini, had made the deal happen. Amd has shown lots of progress lately, however, it'll be incredibly hard to catch up with apple's power efficiency. So what i'm going to do is take that export, make sure we're in 4k on 20 battery hdr, keep in mind this thing's still going."

4. Data Generator:

The original text:

"this video is brought to you by morning brew hi welcome to today s episode of cold fusion before i start this episode is not a recommendation to buy any of apple s products in fact I am a windows and android user myself the following is simply my documentation of"

Auto-generated text:

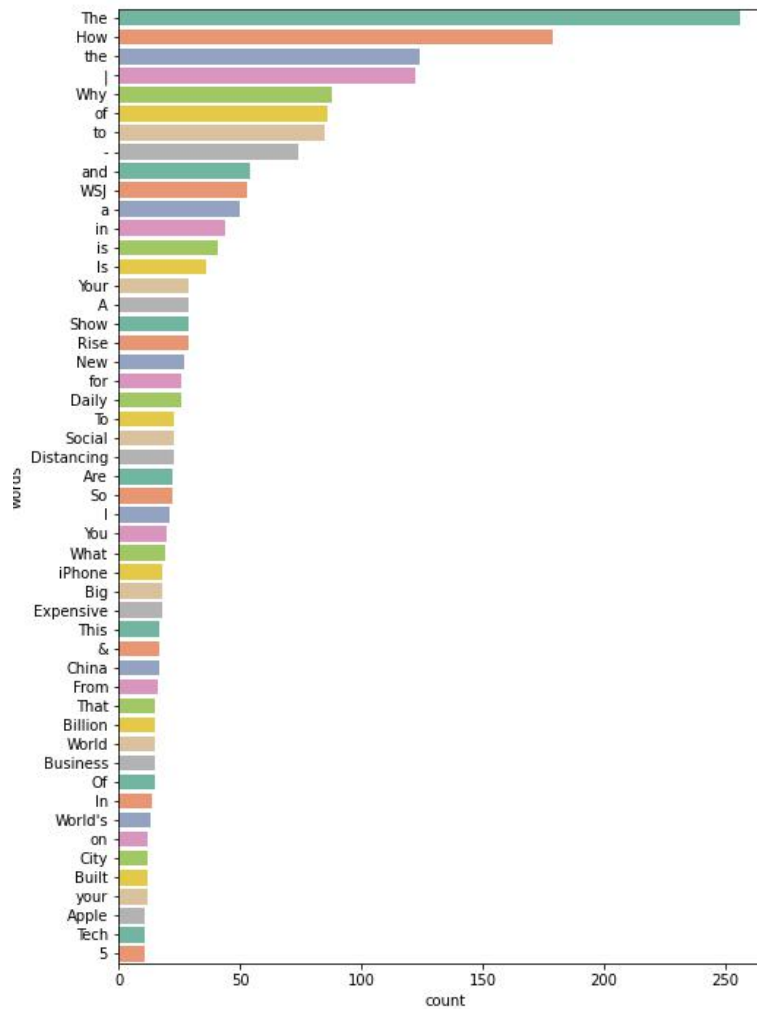
(what i do not think of themselves other technologies in the morning is accused ownership into lately then this episode we are going to do it is possible to think that we have a tremendous problem that you have been watching cold fusion tv and solar and meteoric biological grids we have a slower pace of cells and compete to stop ventilators directly into a yellow organised growth working doctor engineer sharing is a very neat cell phone x on a stanford classroom for a monopoly that a friend of the left and the bull was orion s first...)

According to the text generator, we can auto-generate our scripts based on what topics do we want and save a lot of time. The text generator that we have now is underperformed, because I did not spend a long time to train it yet.

5. Title analysis:

Total words	5650 words
Average words	8.07 words

Word frequency for titles:



6. Tips of how to make a successful video:

Here's everything, step-by-step, that you need to make a successful YouTube video:

Step 1. Create a YouTube video strategy

Step 2. Make sure your video is found on YouTube

Step 3. Find YouTube ideas and topics

Step 4. Understand YouTube equipment for beginners

Step 5. Learn the first YouTube video you should make

Step 6. Set up your video recording

Step 7. Record your desktop

Step 8. Edit your video

Step 9. Upload your video on YouTube

Step 10. Optimize your video for YouTube

Reference:

<http://bark.phon.ioc.ee/punctuator>

<https://backlinko.com/hub/youtube/planning>