YouTube Statistics

Presented by: Adrian Sealy

Video Statistics

- The dataset contains YouTube video statistics from videos published on YouTube's website.

- There are 1882 records ranging from 2007 – 2022, with four columns being used for this analysis, as follows:

- Keyword: The keyword associated with the video

- Likes

- Comments

- Views

Research question

- For this analysis I examined two questions for this dataset

1. Is there a significant difference in a video's views with videos that receive high likes and comments versus videos that receive high views?

2. Is there a correlation between views, likes, and comments a video receives?

Hypotheses

- H1: The mean of keyword Google views will be <> then the mean of keyword Mrbeast views
- Ha: The mean of keyword Google views will be = to the mean of keyword Mrbeast views

- H2: There is a strong correlation with views between likes and comments. This means that the more likes and comments a video generates the more views it will obtain.

How the data tests the hypotheses

- Visualizations and a table will show correlations between views, comments and likes
- T-test will show the significance of keyword Google’s mean views versus keyword Mrbeast’s mean views

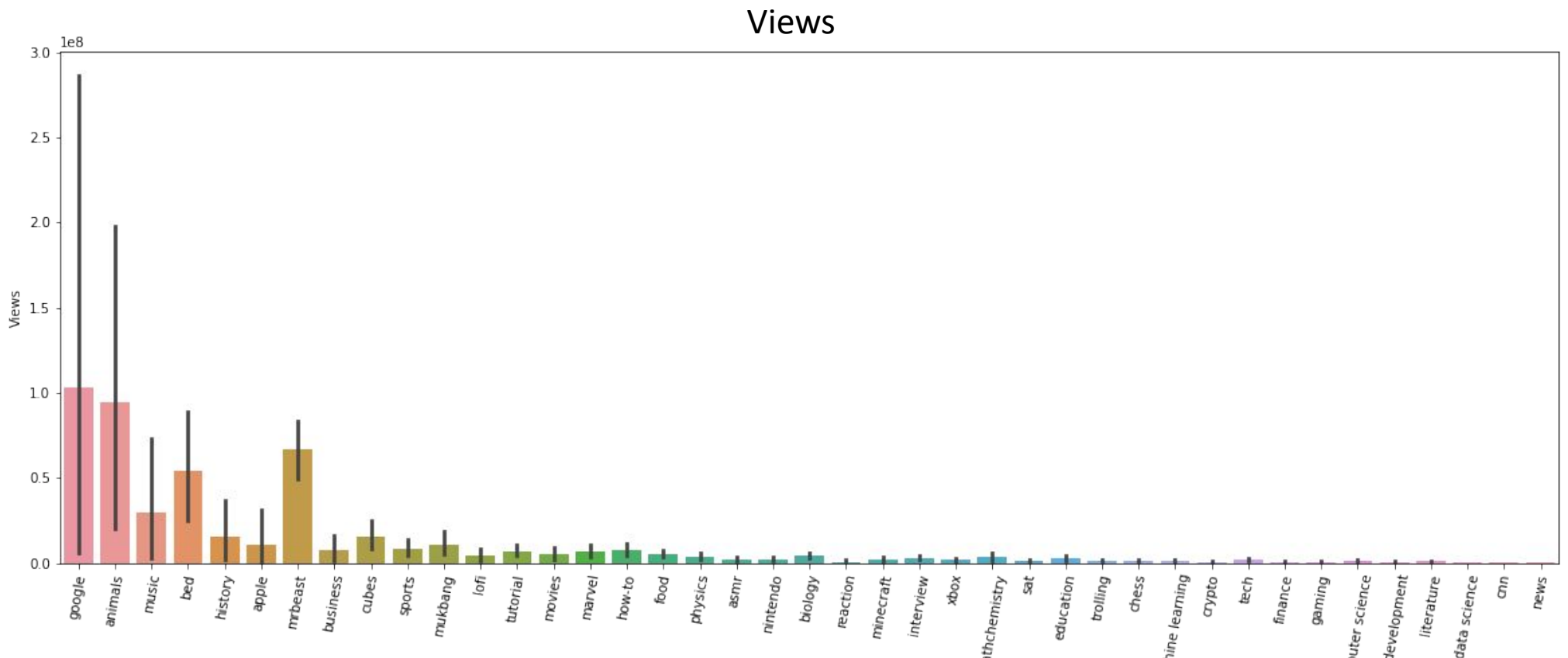
Uses for the findings

- Businesses will be able to use the findings to build criteria to help select content creators to promote their products

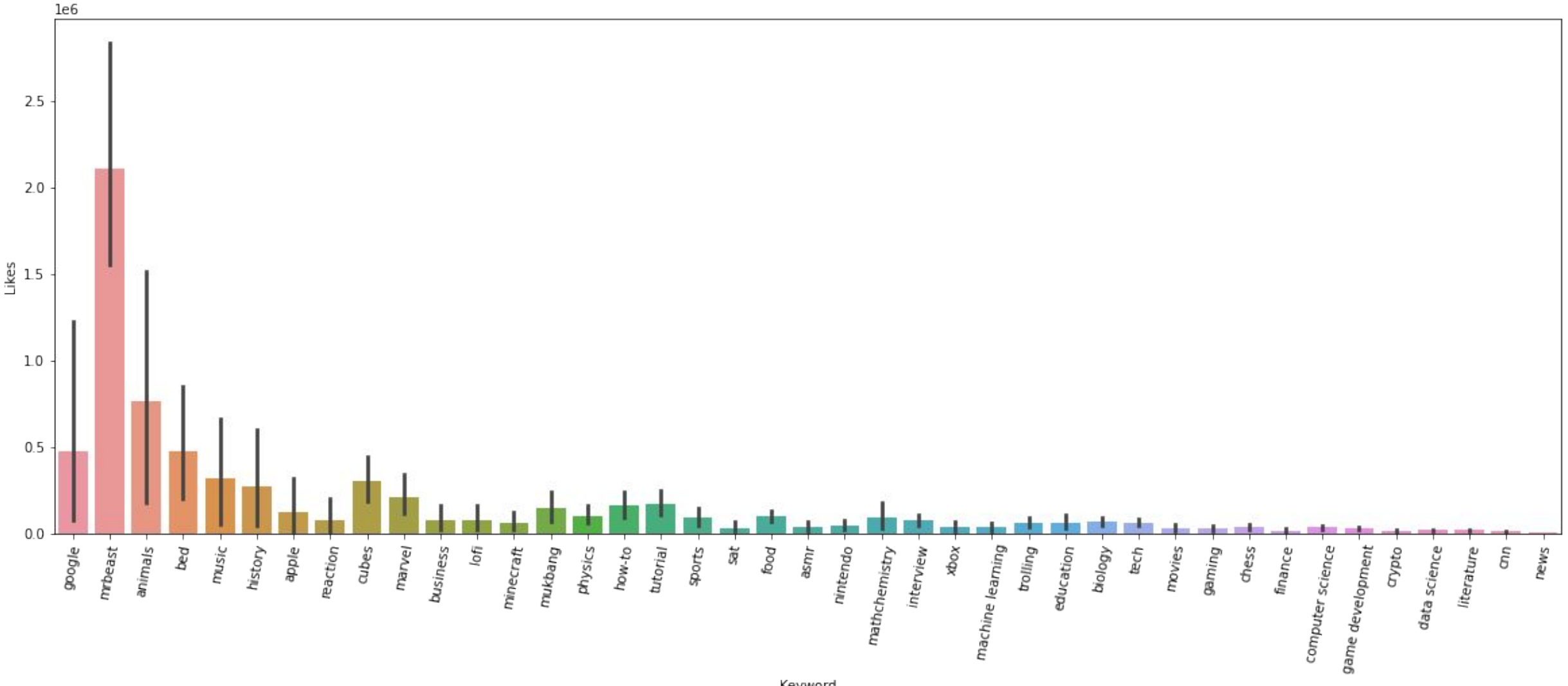
- It will show the importance of keywords used when looking for a content creator with their target audience

Hypothesis 1

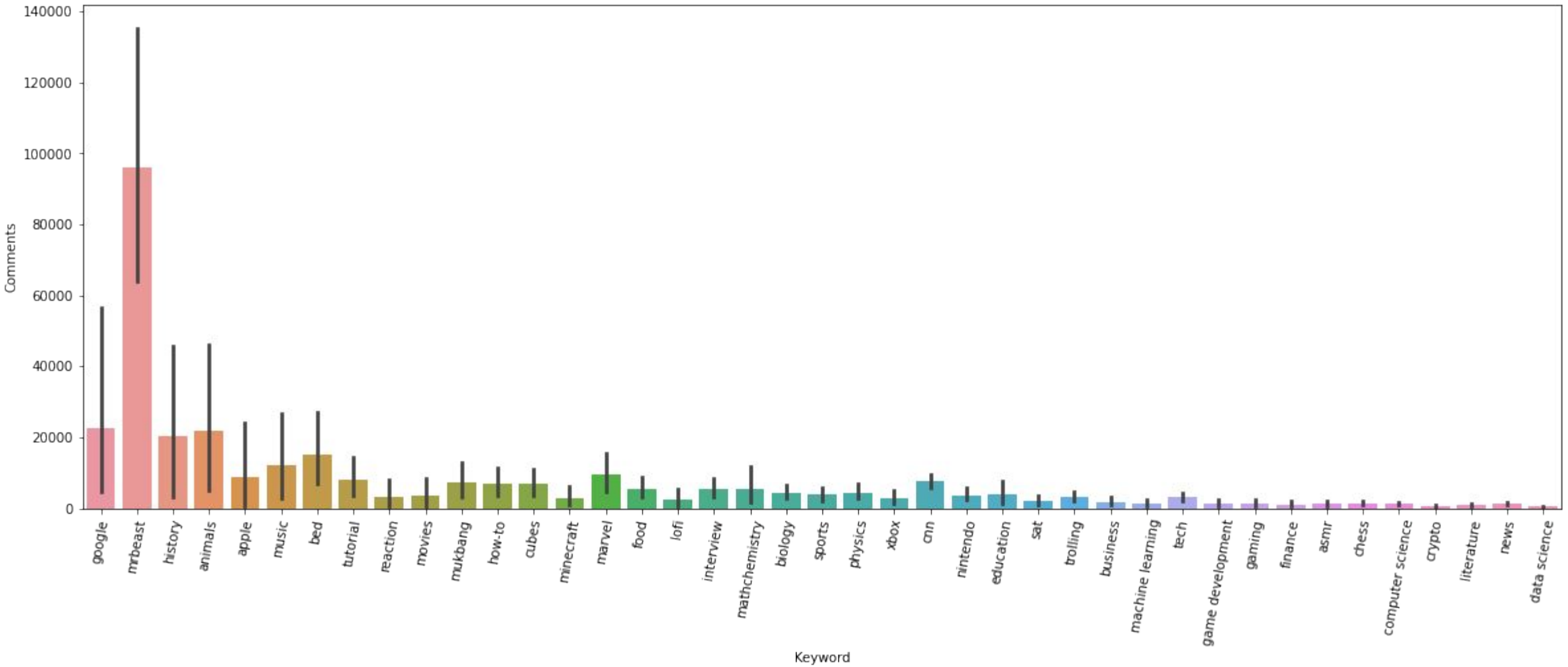
- Keywords help YouTube understand the type of content you produce and who your target audience is
- H1: The mean of Google's views will be <> Then the mean of mrbeast views
- Ha: The mean of Google's views will be = to the mean of mrbeast views



Likes


1e6
2.5
2.0
1.5
1.0
0.5
0.0
Likes
google
mrbeast
animals
bed
music
history
apple
reaction
cubes
marvel
business
lofi
minecraft
mukbang
physics
how-to
tutorial
sports
sat
food
asmr
nintendo
mathchemistry
interview
xbox
machine learning
trolling
education
biology
tech
movies
gaming
chess
finance
computer science
game development
crypto
data science
literature
cnn
news
Keyword

Comment


Comments
140000
120000
100000
80000
60000
40000
20000
0
google
mrbeast
history
animals
apple
music
bed
tutorial
reaction
movies
mukbang
how-to
cubes
minecraft
marvel
food
lofi
interview
mathchemistry
biology
sports
physics
xbox
cnn
nintendo
education
sat
trolling
business
machine learning
tech
game development
gaming
finance
asmr
chess
computer science
crypto
literature
news
data science
Keyword

Hypothesis 1 results

The p-value of 0.67 is greater than the alpha of .05, therefore we can not reject the null hypothesis.

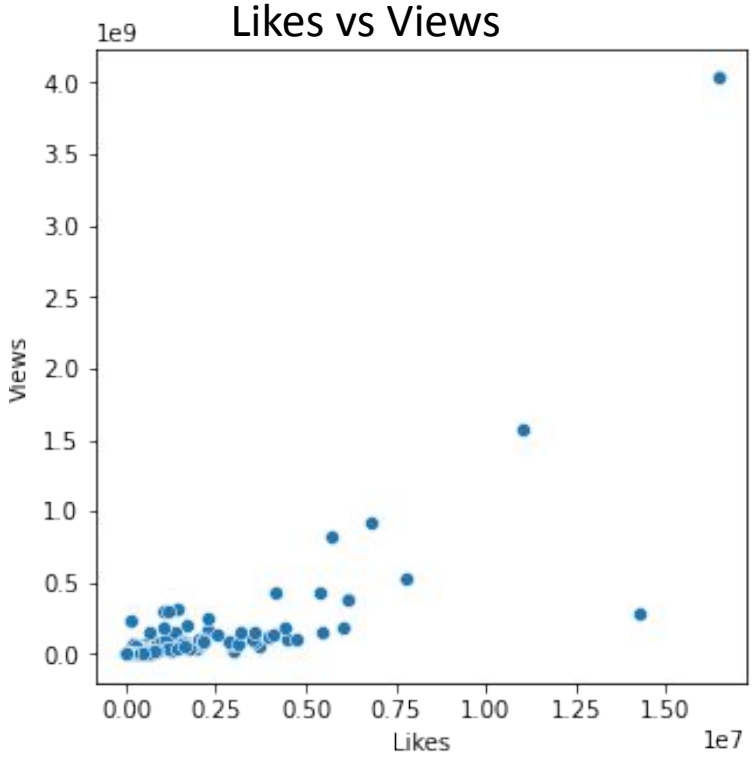
There is a 95% confidence that the difference in means is between -200M and 139M

(-212,911,769.9 and 139,709,236.5)

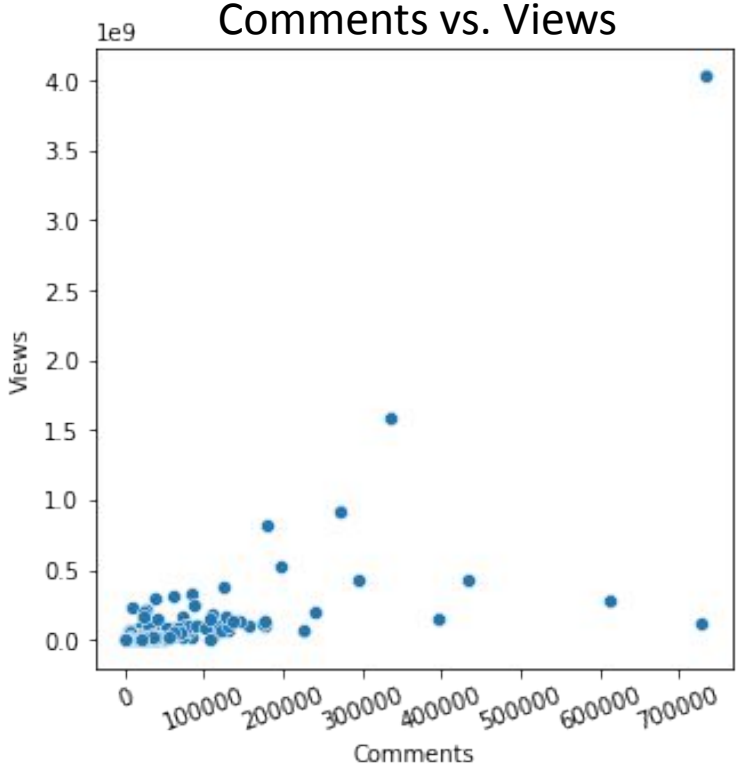
Hypothesis 2

- Is there a strong correlation with views between likes and comments

Likes vs Views



Comments vs. Views



Like vs. Comments

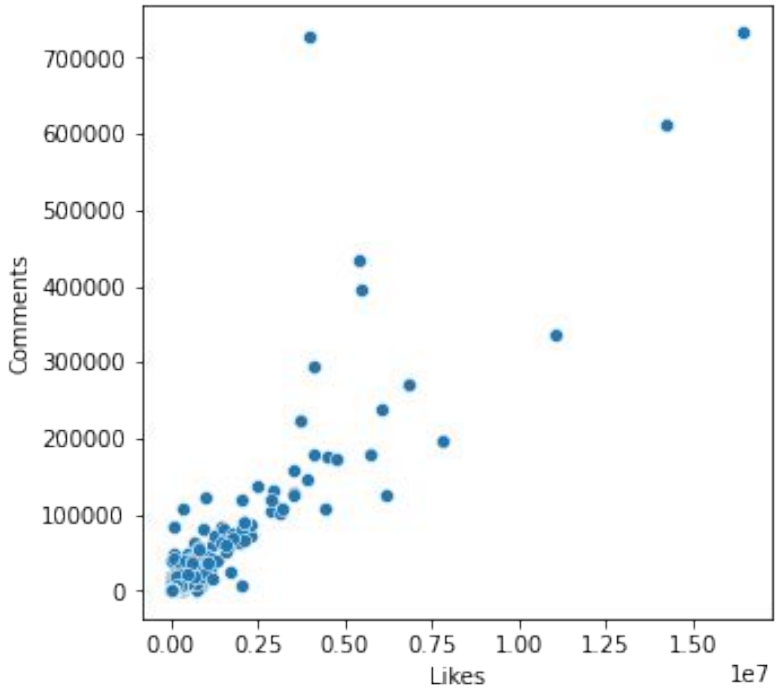


Table of Correlations

	Likes	Comments	Views
Likes	1.000000	0.891903	0.756817
Comments	0.891903	1.000000	0.660877
Views	0.756817	0.660877	1.000000

Recommendations

1. Keywords play a big role in the views a video will get. Therefore, a business must prioritize videos with keywords that match a viewer they want to advertise towards and that will get the most views.

2. When selecting a content creator, I would suggest looking at the views but also the likes a creator gets.

Additional Information

- Link to Colab