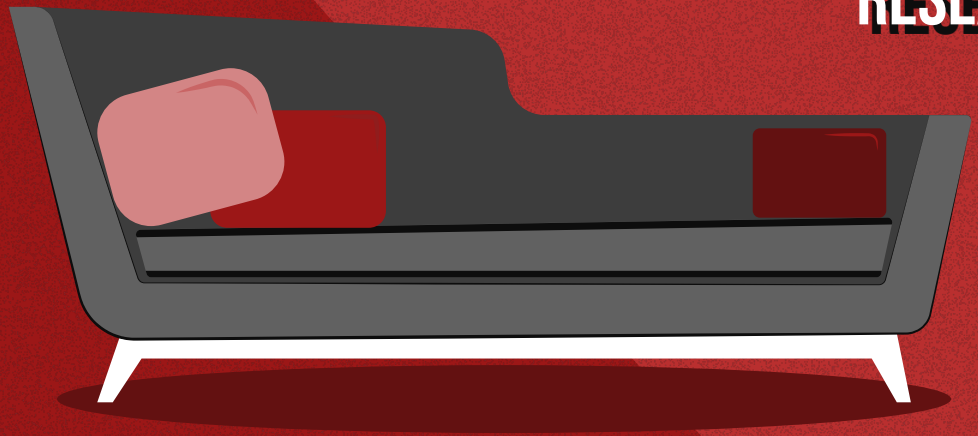
YOUTUBE GENRES

A STATISTICAL EXAMINATION

Brenna Somdahl-Sands, Ethen Kantu, Natalia Morales Flores and Soulai Vang

RESEARCH QUESTION

What is the relationship between a YouTuber's category and their popularity, as measured by the number of subscribers and video views?



CONTEXT & IMPACT

YouTube is the most influential media company of the past 20 years. But, why does some content receive more recognition than others? Understanding influential factors can contribute to the creation of strategies for:

- Audience engagement
- Enhancement of visibility and impact of the content
- Understanding of media cultures and digital trends' influence

WHY?

RELEVANT VARIABLES

X

PREDICTOR

Category:
The 10 largest categories were selected for this research paper

Y

OUTCOME

Popularity:
Measured by number of subscribers, views, subscribers in the last 30 days, and video views in the last 30 days

OUR LINEAR REGRESSION MODEL

$$E[\text{Subscribers}|\text{Category}] = \beta 0 + \beta 1\ \text{Comedy} + \beta 2\ \text{Education} + \beta 3\ \text{Entertainment} + \beta 4\ \text{FilmAndAnimation} + \beta 5\ \text{Gaming} + \beta 6\ \text{HowToAndStyle} + \beta 7\ \text{Music} + \beta 8\ \text{NewsAndPolitics} + \beta 9\ \text{Other} + \beta 10\ \text{PeopleAndBlogs}$$

MODEL STATEMENTS

01
VIDEO VIEWS

E[Views|Category] = β0 + β1Comedy+ ... + β10PeopleAndBlogs

02
SUBSCRIBERS IN 30 DAYS

E[Subscribers30| Category] = β0 + β1Comedy+ ... + β10PeopleAndBlogs

03
VIDEO VIEWS IN 30 DAYS

E[Views30|Category] = β0 + β1Comedy+ ... + β10PeopleAndBlogs

04
POPULATION

E[Subscribers|Population] = β0 + β1 Population

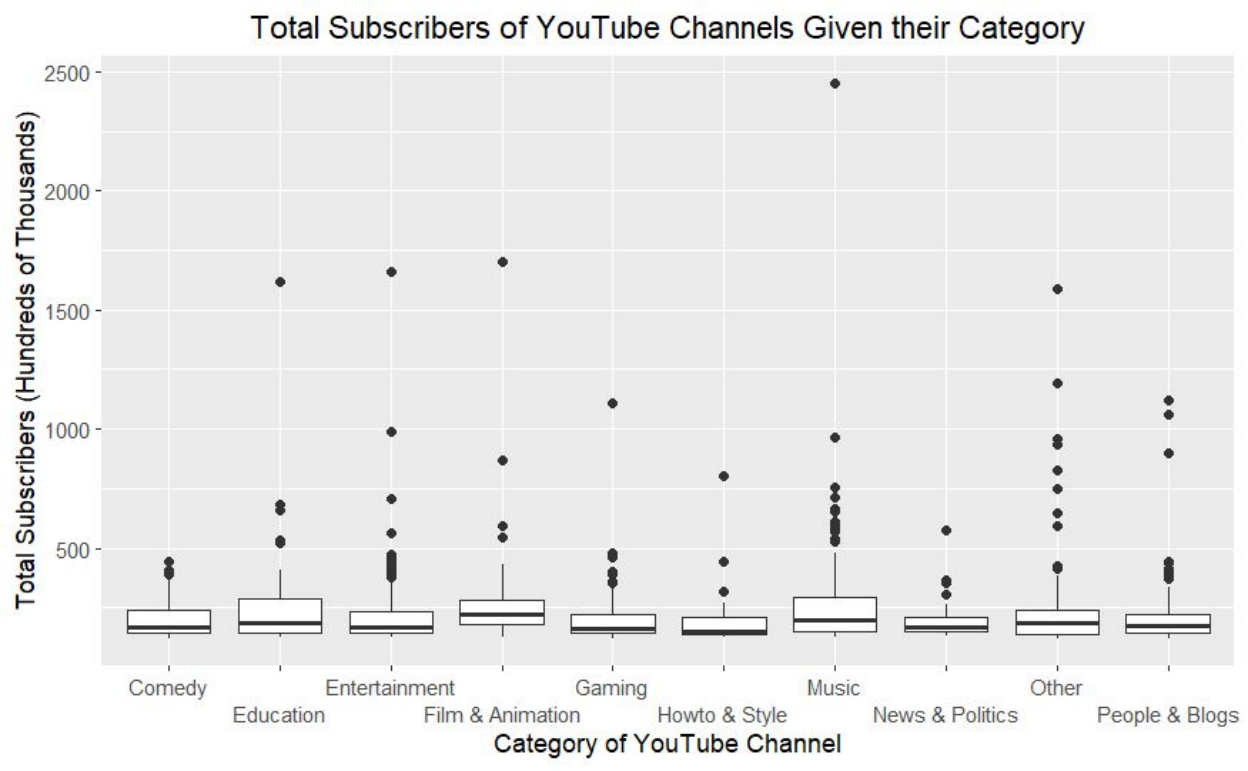
05
POPULATION & CATEGORY

E[Subscribers|Population, Category] = β0+ β1 Population + β2Comedy + ... + β11PeopleAndBlogs

06
INTERACTION TERM

E[Subscribers| Population, Category] = β0+β1 Population+ β2Comedy + ... +β11PeopleAndBlogs+ β12Population * Comedy+ ... + β21Population * PeopleAndBlogs

E[SUBSCRIBERS | CATEGORY]



EFFECT OF CATEGORY OF THE VIDEOS ON THE NUMBER OF SUBSCRIBERS

CATEGORY	ESTIMATE	STD ERROR	TEST STATISTIC	P-VALUE
Education	265	33.375	1.923	0.01169
Gaming	209	27.612	0.264	
Music	257	24.287	2.304	
People & Blogs	211	25.875	0.361	

*not all of the categories are included in this table

ALL THE P-VALUES ARE < 0.05

All our results reject the null hypothesis!

CONCLUSIONS

LIMITATIONS

- Focused on one main predictor (category)
- Possibility of people subscribing to a channel multiple times
- Bots
- Different generation have different habits
- Some categories the consumers are less likely to subscribe

ETHICAL CONSIDERATIONS

- Ambiguity of origin of data

The categories with the highest average number of subscribers and highest average number of video views across all categories were

MUSIC AND EDUCATION

although their growth has been slower than other categories.

The top category for growth was

PEOPLE AND BLOGS

measured by subscribers in the last 30 days

MULTIPLE REGRESSION MODELS

When adding an interaction term, the

POPULATION

of the

COUNTRY OF ORIGIN

affects the relationship between category and popularity.

QUESTIONS?