

STELLA OLIVEIRA

WEB SCRAPING PROJECT

MOTIVATION

- ▶ TED Conferences LLC posts talks online for free distribution
- ▶ Every week 5-7 new talks are published
- ▶ It is watched worldwide
- ▶ Wide range of topics

QUESTIONS

- ▶ How topics changed over time
- ▶ Which topics are most watched
- ▶ Which topics have more videos
- ▶ Is there a relation between the number of videos and views for each topic?


SCRAPING METHOD

- ▶ Main page: <https://www.ted.com/talks>
- ▶ Pattern for the following pages: <https://www.ted.com/talks?page=2>
- ▶ Let's use Scrapy!

SCRAPING METHOD

- ▶ Video page: https://www.ted.com/talks/cecile_richards_the_political_progress_women_have_made_and_what_s_next

← → ↺ 🔒 https://www.ted.com/talks/cecile_richards_the_political_progress_women_have_made_and_what_s_next



The political progress women have made — and what's next

▶ 17:05 🔊 🗨 ⚙ ↗

Details **Transcript** **Reading List** **Footnotes** **Comments (8)**

About the talk 1 language Further learning Notes + references Join the conversation

Women have made enormous progress over the last century -- challenging the status quo, busting old taboos and changing business from the inside out. But when it comes to political representation, there's still a long way to go, says activist Cecile Richards. In this visionary talk, Richards calls for a global political revolution for women's equality and offers her ideas for how we can build it.

This talk was presented at an official TED conference, and was featured by our editors on the home page.

ABOUT THE SPEAKER



Cecile Richards · Activist

Cecile Richards is a national leader for women's rights and social and economic justice, and the author of the bestselling book "Make Trouble."

896,761 views

TEDWomen 2018 |
November 2018

Related tags

[Politics](#)
[Government](#)
[Feminism](#)
[Women](#)
[History](#)
[Gender](#)
[Gender Equality](#)
[Activism](#)
[Social Change](#)
[Inequality](#)
[Leadership](#)

DATA – SCRAPED AT 01/25/2019

▶ 3471 rows

▶ Columns:

▶ Title

▶ Description

▶ Keywords (Related tags)

▶ Date

▶ Duration of the video

▶ Number of views

▶ Number of languages

▶ Speaker name

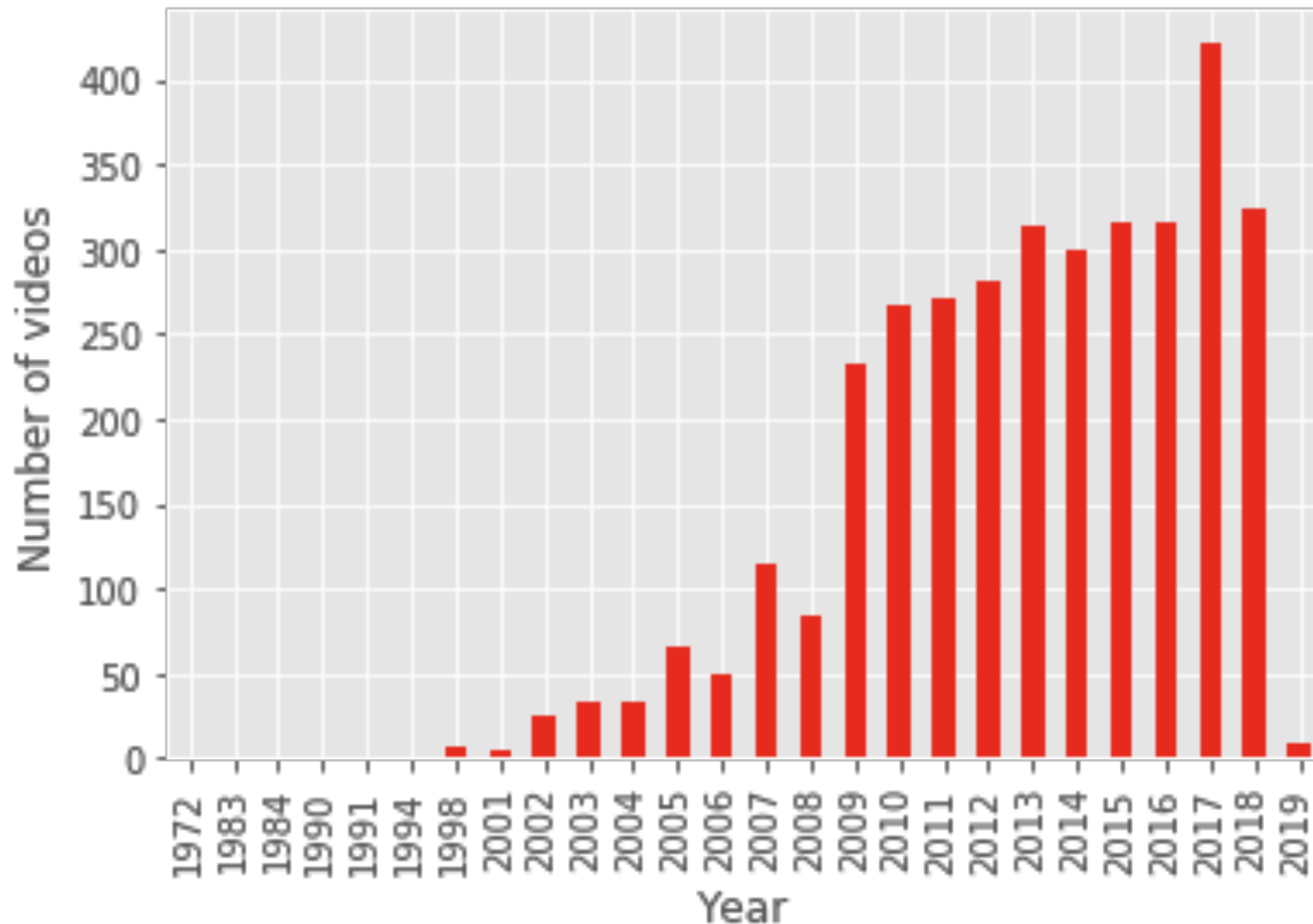
▶ Speaker job

▶ About the speaker

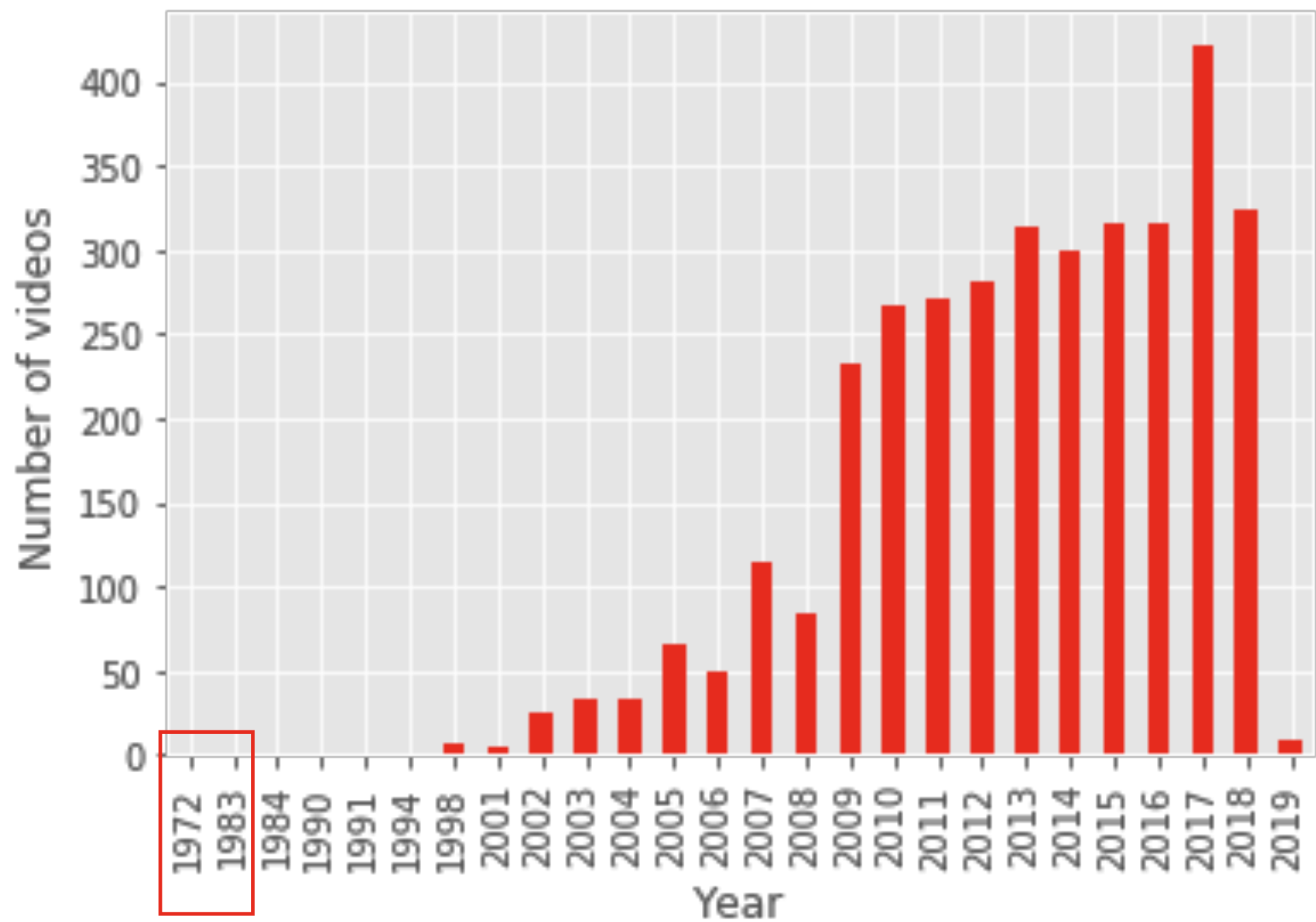
▶ Number of comments

	duration	views	n_lang	n_comments
count	3471.000000	3.471000e+03	3471.000000	3005.000000
mean	737.832901	1.723646e+06	23.646788	171.512146
std	352.669288	2.901919e+06	11.807034	272.120724
min	0.000000	0.000000e+00	0.000000	2.000000
25%	471.500000	6.869205e+05	18.000000	46.000000
50%	749.000000	1.141044e+06	25.000000	100.000000
75%	994.000000	1.747263e+06	31.000000	201.000000
max	3608.000000	5.573491e+07	72.000000	6418.000000

NUMBER OF VIDEOS PER YEAR



NUMBER OF VIDEOS PER YEAR



TALK FROM 1972

Viktor Frankl | Toronto Youth Corps

Why believe in others

- Details

About the talk
- Comments (377)

Join the conversation

In this rare clip from 1972, legendary psychiatrist and Holocaust-survivor Viktor Frankl delivers a powerful message about the human search for meaning -- and the most important gift we can give others.

1,263,574 views

Toronto Youth Corps | May 1972

ABOUT THE SPEAKER



Viktor Frankl · Psychiatrist, neurologist, author
Neurologist and psychiatrist Viktor Frankl pioneered an approach to psychotherapy that focuses on the human search for meaning.

Related tags

- Humanity
- Mind
- Peace
- Psychology
- War

ABOUT BEST OF THE WEB

Best of the Web, just like the rest of TED.com, involves riveting talks by remarkable people. But unlike TEDTalks, these talks don't come from TED or any of our partner conferences. These talks come from all over the Web. We'll draw from any source — from lectures at little-known forums to famous speeches that made history — so long as the video is available for free, and so long as the talk meets our most important benchmark: that it's an Idea Worth Spreading.

TALK FROM 1983



Details

About the talk

Comments (148)

Join the conversation

In this archival footage from BBC TV, celebrated physicist Richard Feynman explains what fire, magnets, rubber bands (and more) are like at the scale of the jiggling atoms they're made of. This accessible, enchanting conversation in physics reveals a teeming nano-world that's just plain fun to imagine.

621,904 views

BBC TV | July 1983

Related tags

- [Astronomy](#)
- [Physics](#)
- [Science](#)

ABOUT THE SPEAKER



Richard Feynman · Physicist

One of the best known and most renowned scientists in history, Richard Feynman pioneered quantum mechanics. His knack for accessible explanations made him a popularizer of physics of equal distinction to laypeople.

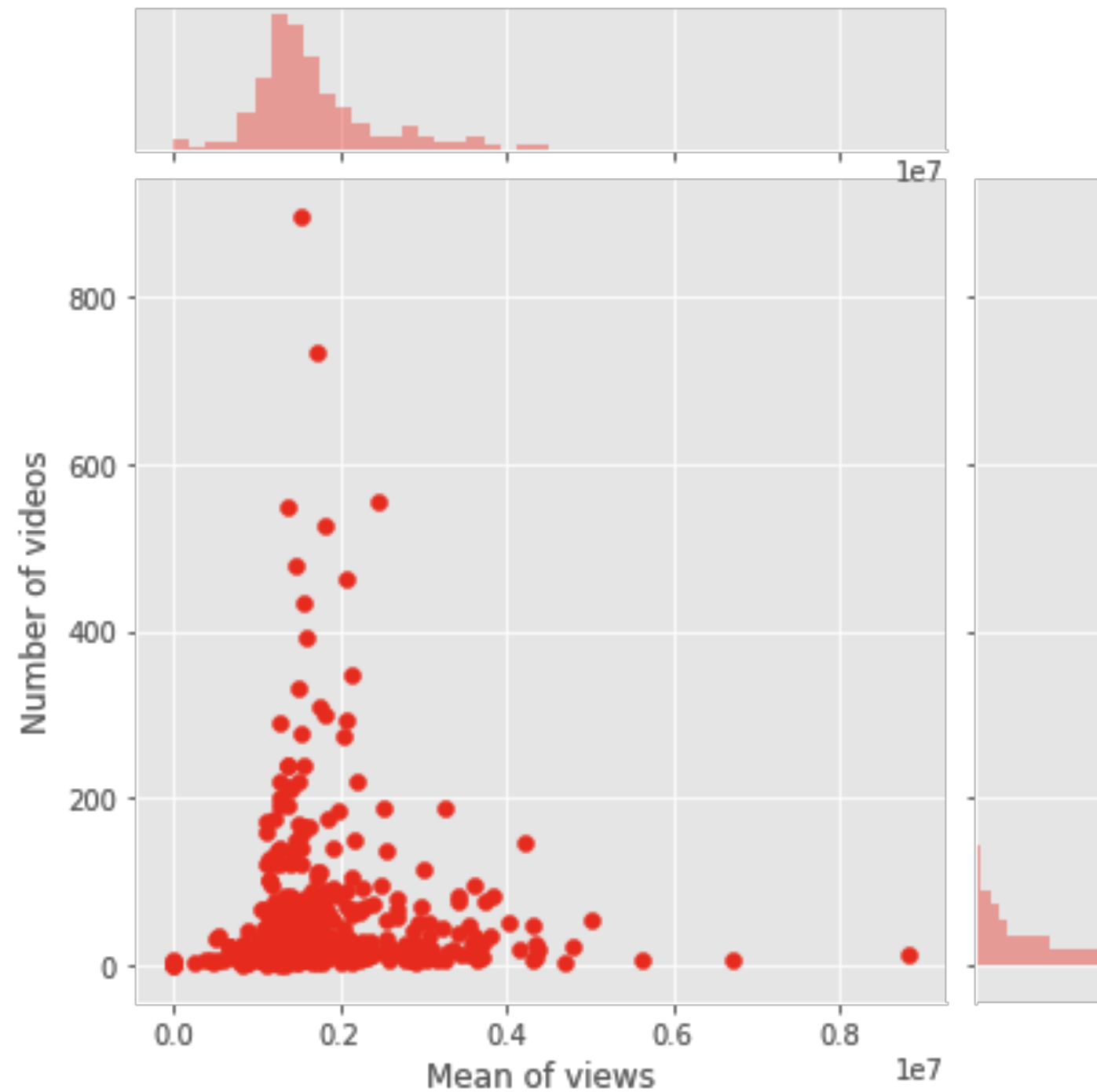
ABOUT BEST OF THE WEB

Best of the Web, just like the rest of TED.com, involves riveting talks by remarkable people. But unlike TEDTalks, these talks don't come from TED or any of our partner conferences. These talks come from all over the Web. We'll draw from any source — from lectures at little-known forums to famous speeches that made history — so long as the video is available for free, and so long as the talk meets our most important benchmark: that it's an Idea Worth Spreading.

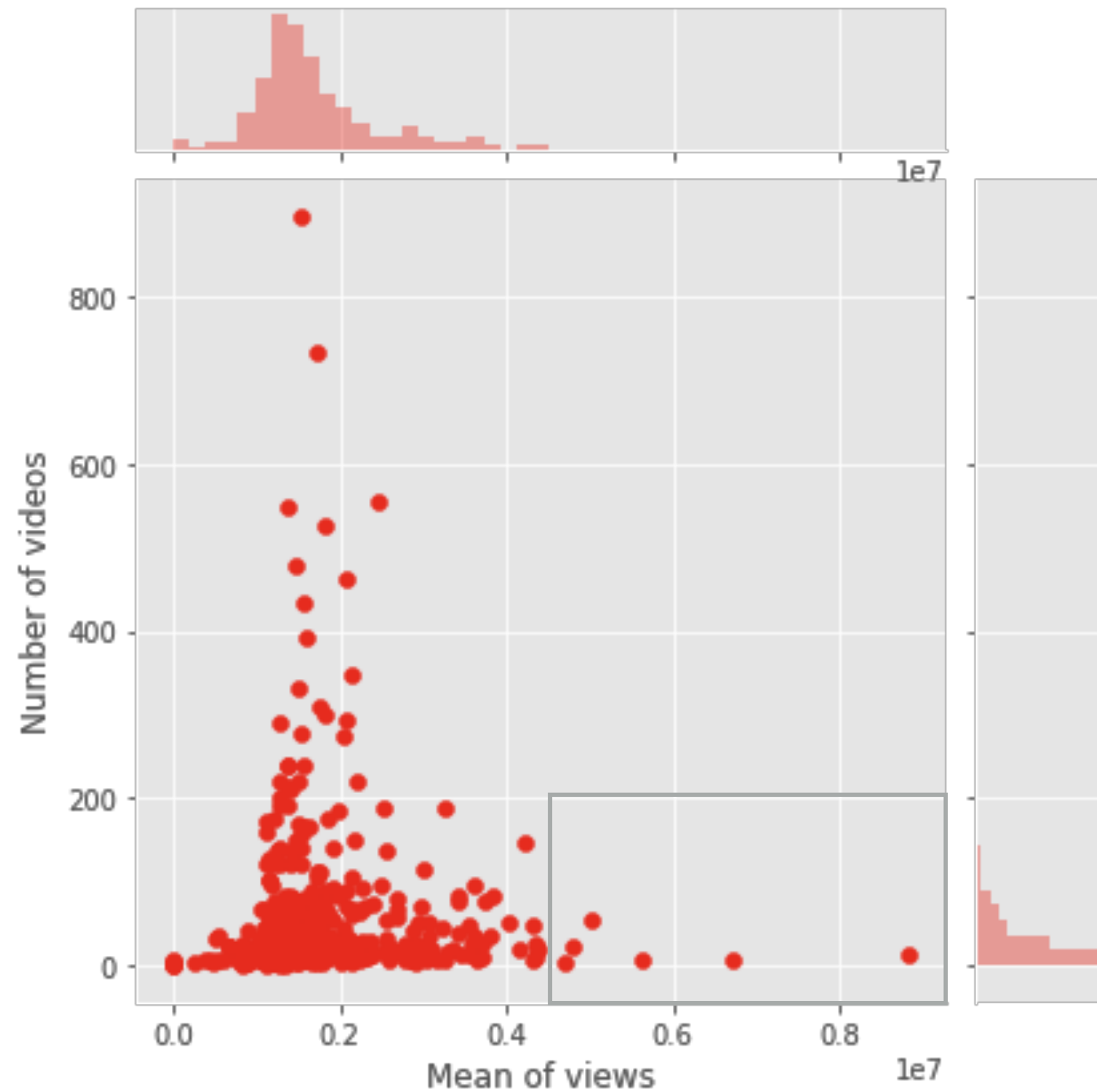
- 448 unique topics



TOPICS – VIEWS AND NUMBER OF VIDEOS

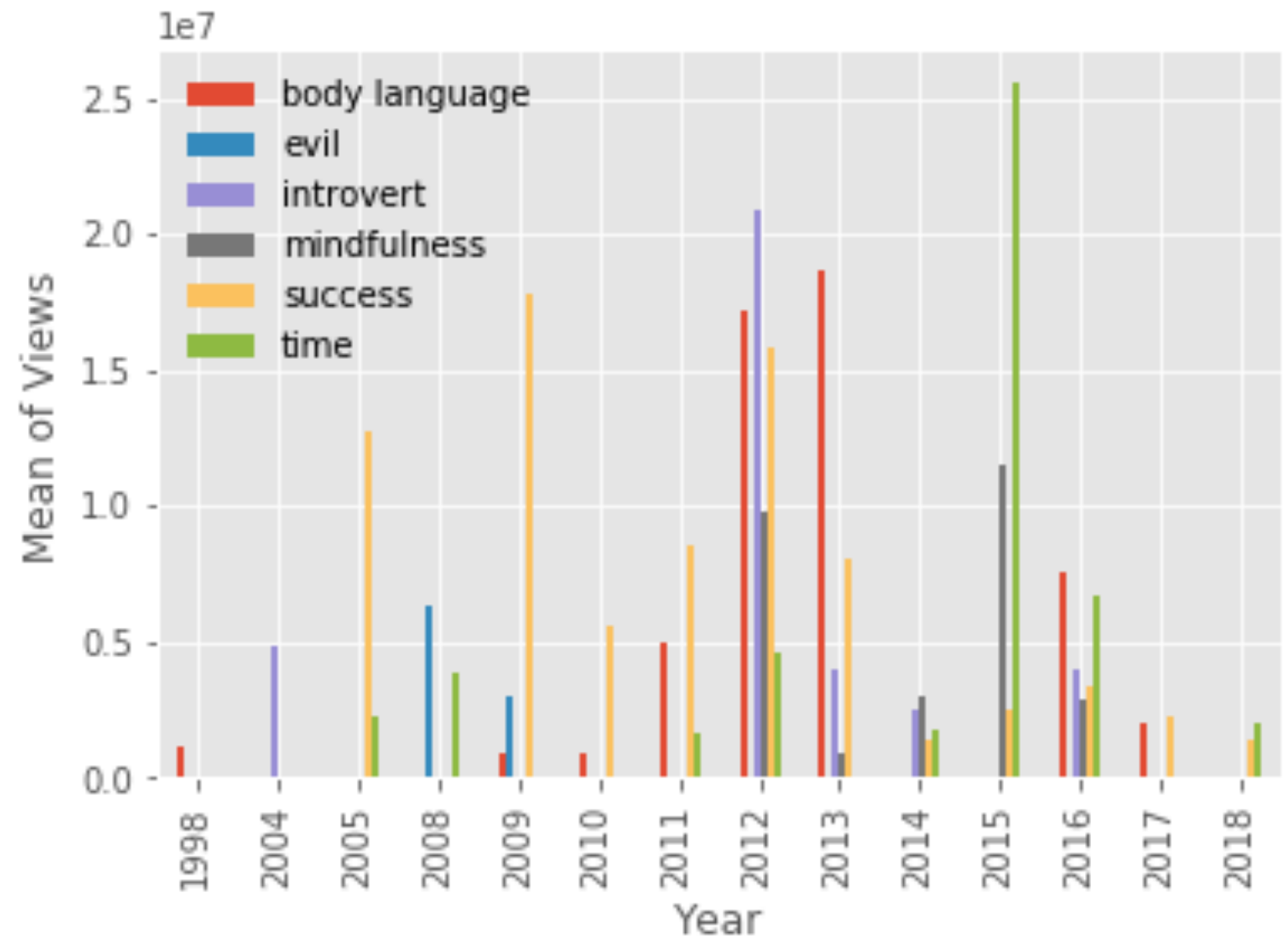


TOPICS – VIEWS AND NUMBER OF VIDEOS

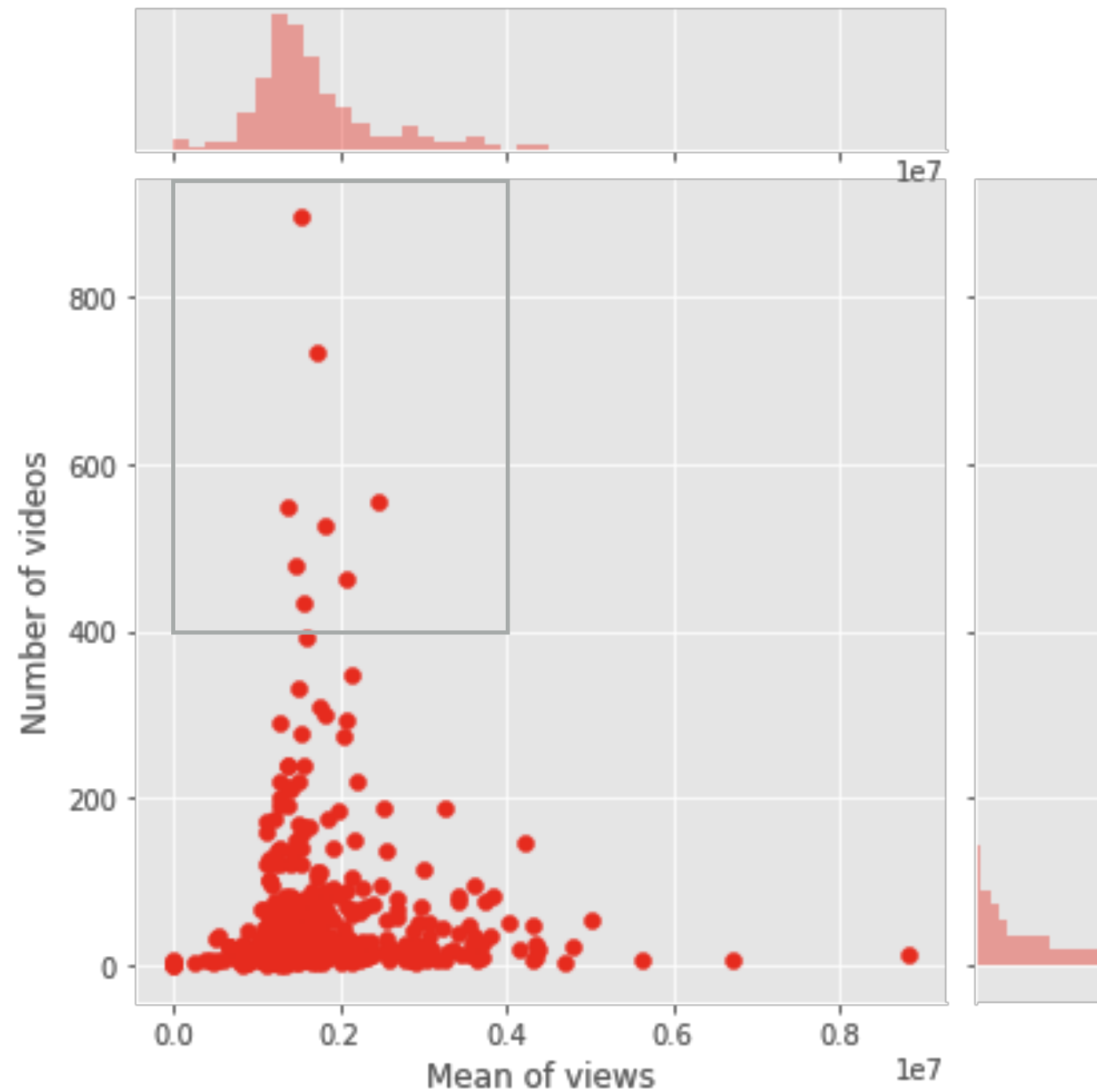


TOPICS - VIEWS > 4,000,000

	views	title
keywords		
body language	8.827407e+06	12
introvert	6.717924e+06	6
mindfulness	5.632883e+06	5
success	5.021557e+06	53
time	4.792291e+06	21
evil	4.705632e+06	2

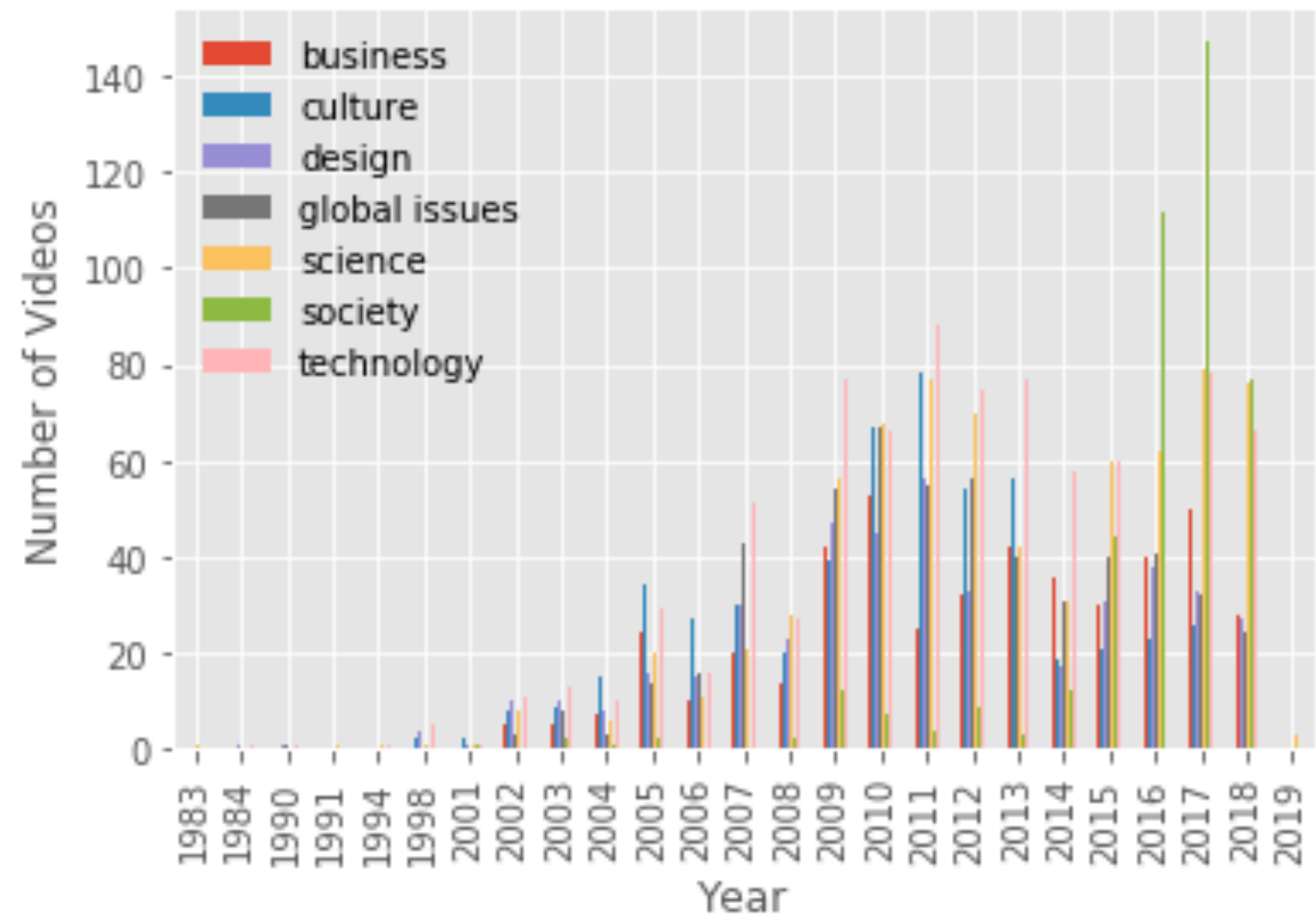


TOPICS – VIEWS AND NUMBER OF VIDEOS



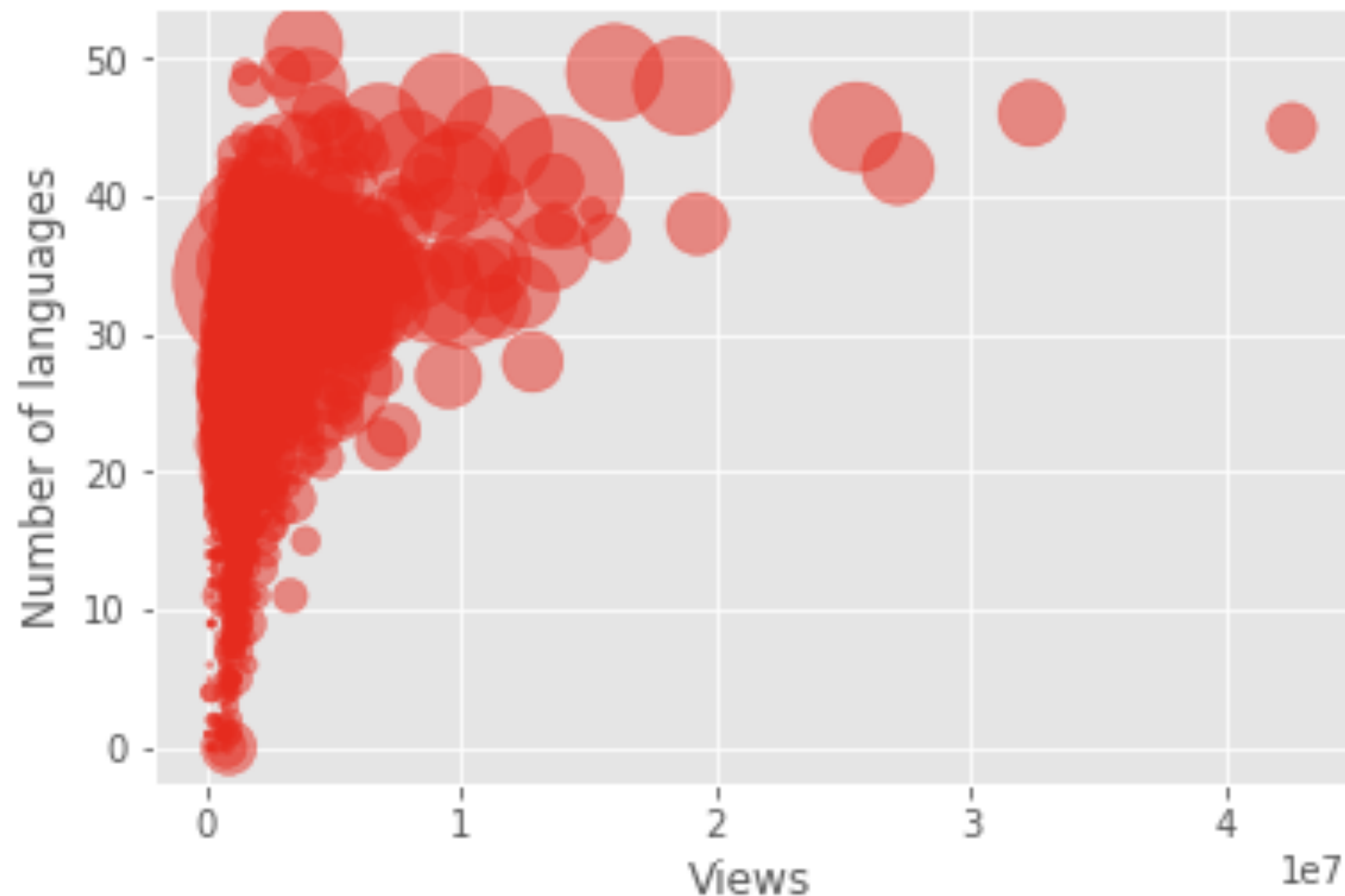
TOPICS – NUMBER OF VIDEOS > 400

	views	title
keywords		
technology	1.530273e+06	897
science	1.729308e+06	735
culture	2.474622e+06	556
global issues	1.383985e+06	548
design	1.456694e+06	480
business	2.064136e+06	463
society	1.580135e+06	435



VIEWS – NUMBER OF LANGUAGES – NUMBER OF COMMENTS

- ▶ Data from the last 5 years
- ▶ Size of circle represents the number of comments



FUTURE WORK

- ▶ Network among topics
- ▶ Scrape comments date
- ▶ Scrape books that are recommended by the speaker
- ▶ Relation between most view videos and trending topics in social media