

TED Talk

Important columns

views
comments
Rating
Tag
Published date

Question

What makes a popular Ted talk?

total view = 4-16 B

total talk = 2406

Rating = 14 category

Tag = 423

Find

TOP 5 speaker and their occupation
when most Ted events happened (which month and week day)

Rating column

Rating column is a key value column (JSON data)
create python script to generate new records.

Example

Input

Index	comment	desc	rating
1	4553	abe	{id: 7, name: "Funny", count: 190}, {id: 1, name: "Beautiful", count: 4573}

Output

Index	comment	desc	Funny	Beautiful	12 Rating col
1	4553	abe	190	4573	...

Tag column

This is useful column we can understand what is the topic of that ted talk. But it is in List (array) format.

create a python script to generate new records with tags.

Example

Index	comm	tags
1	5	['children', 'teaching', 'Tedx', 'parenting']

O/P

Index	comm	theme
1	5	children
1	5	teaching
1	5	Tedx
1	5	parenting

Analysis

Most ted talks are rated as positive

positive: (inspiring, informative, fascinating, funny, jaw dropping, encourages, beautiful)

very less are rated as -ve

-ve: longwinded, obnoxious, confusing

Question

are most viewed also most commented as well?

create scatter plot

comment vs view

→ found no strong relationship

create additional feature

discussion quotient = comment/view

most commented ted talks are mostly controversial (Ex. faith and religion)

Why ^{some} ted talks are unpopular?

Rating on them are mostly positive.

so that is not a reason.

2006

Ted talk started

check the trend (line chart)

It shows Increase in count

most unpopular talk happened between (2008-2010)

check the speaker's occupation (Not ^{found} best occupations for ted talk)

topics (global issue, Africa, Ted X, compassion)
which are not relevant for that time
as 2008-2010 was a recession
season.