Trend Analysis in TED Talks

By

- *Nikhil Gudugula*
- *Siddarth Maredu*
- *Sathish Komire*

What is a TED Talk and Why Did we choose it?

TED (Technology, Entertainment, Design) is a global platform that hosts a wide range of influential speakers who deliver engaging and thought-provoking talks on diverse topics.

Text Analysis

TED Partnerships

- Opportunity to be part of global culture
- Not just buying ad space
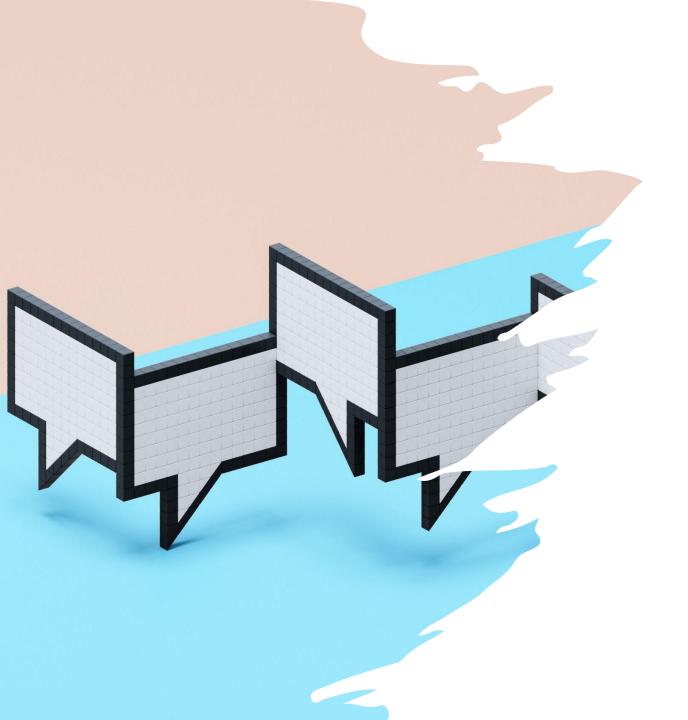- Lasting relationship

Tools Used

Data Cleaning and filtering

R – Data Wrangling and Analysis

Tableau - Visualization

Data Understanding

#Views

#Comments

Descriptions

Speakers

Titles

Transcripts

Data Cleaning

row_id	ratings
1	[{'id': 7, 'name': 'Funny', 'count': 19645}, {'id': 1, 'name': 'Beautiful', 'count': 4573}, {'id': 9, 'name': 'Ingenious', 'count': 6073
2	[{'id': 7, 'name': 'Funny', 'count': 544}, {'id': 3, 'name': 'Courageous', 'count': 139}, {'id': 2, 'name': 'Confusing', 'count': 62}, {
3	[{'id': 7, 'name': 'Funny', 'count': 964}, {'id': 3, 'name': 'Courageous', 'count': 45}, {'id': 9, 'name': 'Ingenious', 'count': 183}, {
4	[{'id': 3, 'name': 'Courageous', 'count': 760}, {'id': 1, 'name': 'Beautiful', 'count': 291}, {'id': 2, 'name': 'Confusing', 'count': 32

row_id	id	name	count
1	7	Funny	19645
1	1	Beautiful	4573
1	9	Ingenious	6073
1	3	Courageous	3253
1	11	Longwinded	387
1	2	Confusing	242
1	8	Informative	7346
1	22	Fascinating	10581
1	21	Unconvincing	300
1	24	Persuasive	10704

Data Transformation

```
Rows: 2,550
Columns: 17
$ comments         <int> 4553, 265, 124, 200, 593, 672, 919, 46, 852, 900, 7…
$ description      <chr> "Sir Ken Robinson makes an entertaining and profoun…
$ duration         <int> 1164, 977, 1286, 1116, 1190, 1305, 992, 1198, 1485,…
$ event            <chr> "TED2006", "TED2006", "TED2006", "TED2006", "TED200…
$ film_date        <int> 1140825600, 1140825600, 1140739200, 1140912000, 114…
$ languages        <int> 60, 43, 26, 35, 48, 36, 31, 19, 32, 31, 27, 20, 24,…
$ main_speaker     <chr> "Ken Robinson", "Al Gore", "David Pogue", "Majora C…
$ name             <chr> "Ken Robinson: Do schools kill creativity?", "Al Go…
$ num_speaker      <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,…
$ published_date   <int> 115136706O, 1151367060, 1151367060, 1151367060, 115…
$ ratings          <chr> "[{'id': 7, 'name': 'Funny', 'count': 19645}, {'id'…
$ related_talks    <chr> "[{'id': 865, 'hero': 'https://pe.tedcdn.com/images…
$ speaker_occupation <chr> "Author/educator", "Climate advocate", "Technology …
$ tags             <chr> "['children', 'creativity', 'culture', 'dance', 'ed…
$ title            <chr> "Do schools kill creativity?", "Averting the climat…
$ url              <chr> "https://www.ted.com/talks/ken_robinson_says_school…
$ views            <int> 47227110, 3200520, 1636292, 1697550, 12005869, 2068…
```
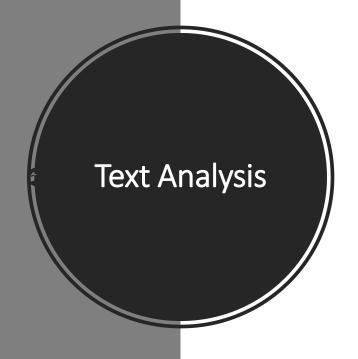
→

```
Rows: 2,550
Columns: 18
$ comments         <int> 4553, 265, 124, 200, 593, 672, 919, 46, 852, 900, 7…
$ description      <chr> "Sir Ken Robinson makes an entertaining and profoun…
$ duration         <int> 1164, 977, 1286, 1116, 1190, 1305, 992, 1198, 1485,…
$ event            <chr> "TED2006", "TED2006", "TED2006", "TED2006", "TED200…
$ film_date        <dttm> 2006-02-25, 2006-02-25, 2006-02-24, 2006-02-26, 20…
$ languages        <int> 60, 43, 26, 35, 48, 36, 31, 19, 32, 31, 27, 20, 24,…
$ main_speaker     <chr> "Ken Robinson", "Al Gore", "David Pogue", "Majora C…
$ name             <chr> "Ken Robinson: Do schools kill creativity?", "Al Go…
$ num_speaker      <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,…
$ published_date   <dttm> 2006-06-27 00:11:00, 2006-06-27 00:11:00, 2006-06-…
$ ratings          <chr> "[{'id': 7, 'name': 'Funny', 'count': 19645}, {'id'…
$ related_talks    <chr> "[{'id': 865, 'hero': 'https://pe.tedcdn.com/images…
$ speaker_occupation <chr> "Author/educator", "Climate advocate", "Technology …
$ tags             <chr> "['children', 'creativity', 'culture', 'dance', 'ed…
$ title            <chr> "Do schools kill creativity?", "Averting the climat…
$ url              <chr> "https://www.ted.com/talks/ken_robinson_says_school…
$ views            <int> 47227110, 3200520, 1636292, 1697550, 12005869, 2068…
$ transcript       <chr> "Good morning. How are you?(Laughter)It's been grea…
```
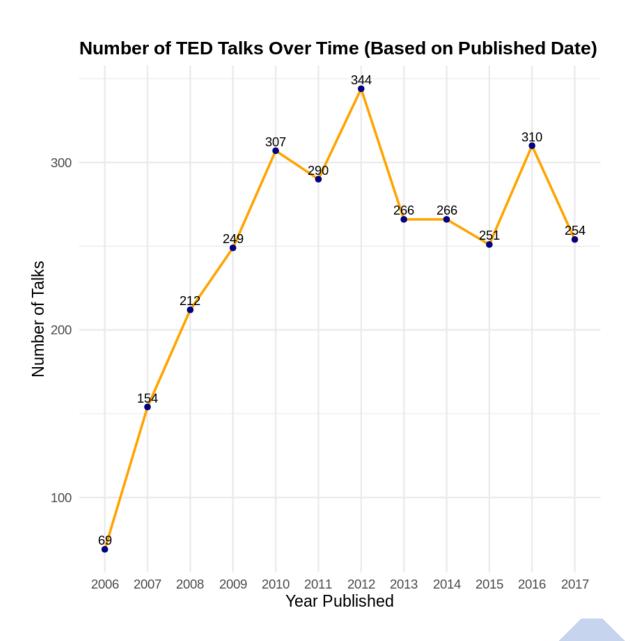
Statistical Analysis


```
       TED_ID          event           num_speaker    main_speaker
 Min.    :   1.0   Length:2972       Min.   :1.00    Length:2972
 1st Qu.: 632.8   Class :character   1st Qu.:1.00   Class :character
 Median :1280.5   Mode  :character   Median :1.00   Mode  :character
 Mean   :1284.4                      Mean   :1.03
 3rd Qu.:1952.2                       3rd Qu.:1.00
 Max.   :2550.0                       Max.   :5.00

 speaker_occupation description         duration
 Length:2972        Length:2972      Min.    : 135.0
 Class :character   Class :character   1st Qu.: 582.0
 Mode  :character   Mode  :character   Median : 848.5
                                       Mean    : 830.1
                                       3rd Qu.:1047.5
                                       Max.    :5256.0

   film_date                    published_date
 Min.    :1972-05-14 00:00:00.00   Min.    :2006-06-27 00:00:00.00
 1st Qu.:2009-11-06 00:00:00.00   1st Qu.:2010-03-05 00:00:00.00
 Median :2012-04-14 12:00:00.00   Median :2012-07-05 12:00:00.00
 Mean    :2011-12-05 13:51:28.83   Mean    :2012-08-09 09:49:10.74
 3rd Qu.:2014-11-19 18:00:00.00   3rd Qu.:2015-04-06 06:00:00.00
 Max.    :2017-08-27 00:00:00.00   Max.    :2017-09-22 00:00:00.00
 NA's    :14
   languages       comments          tags           title
 Min.    : 0.00   Min.    :   2.0   Length:2972      Length:2972
 1st Qu.:23.00   1st Qu.:  63.0   Class :character   Class :character
 Median :28.00   Median : 117.0   Mode  :character   Mode  :character
 Mean    :27.16   Mean    : 193.4
 3rd Qu.:33.00   3rd Qu.: 224.0
 Max.    :72.00   Max.    :6404.0

   url             views          transcript       ratings_id
 Length:2972      Min.    :   50443   Length:2972      Length:2972
 Class :character   1st Qu.:  761876   Class :character   Class :character
 Mode  :character   Median : 1132964   Mode  :character   Mode  :character
                    Mean    : 1705681
                    3rd Qu.: 1734962
                    Max.    :47227110
```
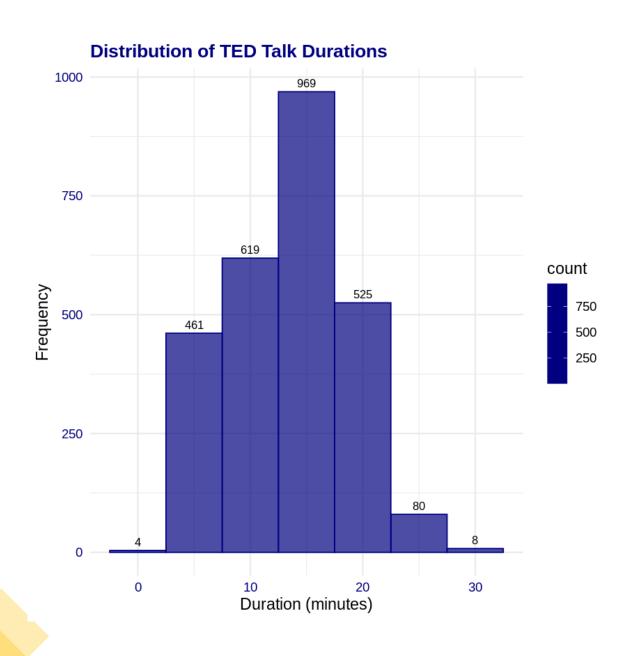

A tibble: 1 × 4

negative	positive	sentiment_score	neutral
<dbl>	<dbl>	<dbl>	<dbl>
4425	6531	2106	0

A matrix: 4 × 4 of type dbl

	duration	views	comments	languages
duration	1.00000000	0.04171085	0.1369176	-0.3005083
views	0.04171085	1.00000000	0.5130712	0.3799774
comments	0.13691758	0.51307120	1.0000000	0.3106607
languages	-0.30050830	0.37997737	0.3106607	1.0000000
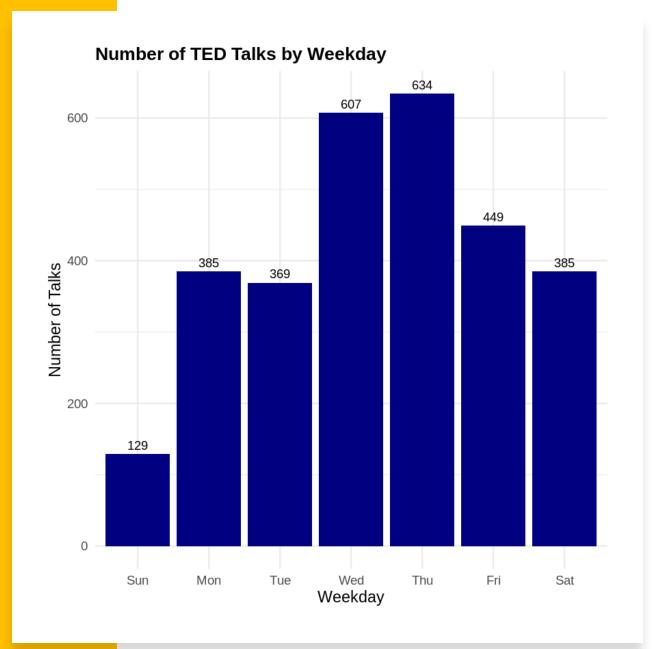
Text Analysis

Number of TED Talks Over Time (Based on Published Date)
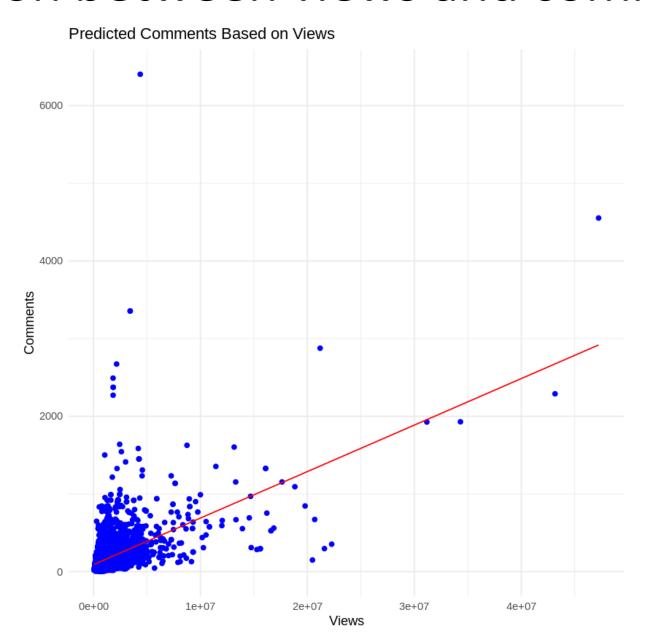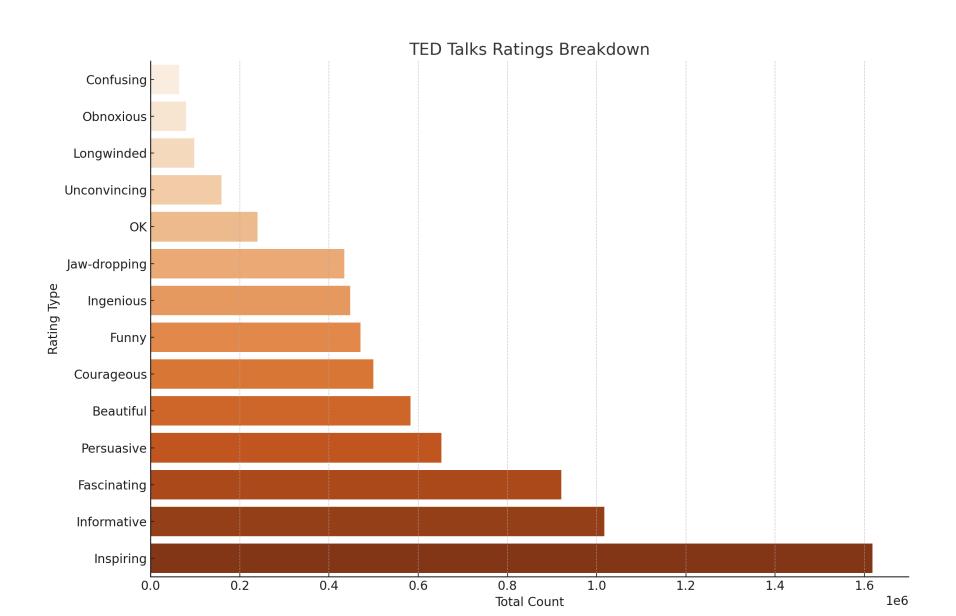
Top shares of talks by speaker occupation type

Number of talks published (WeekDay wise)

Relation between views and comments

Total count per Individual Rating Type

TED Talks Ratings Breakdown

Challenges

- Choosing the dataset Ted Talks
- Handling nested structures in the dataset
- Learning Tableau
- Collaboration challenges

Thank You ☺

Any Questions?