

# TED Talk Project Report

Summer He , Salar Saljoughi

- **Description**

On this project we decided to analyse TED Talk dataset.

The Questions we answered in this process are :

1. what's the 20 most viewed Ted Talks of all time? (cause the nbr views is a good indicator of the popularity of the talk)
2. Comments [mean, max, min,]
3. Does comments have relationship with views?
4. which talks tend to attract more comments?
5. Month (which month has more talks) and (which day has most talks as well) & Year analysis
6. TED Talk vs. TEDx Talk
7. Ted Speakers -- Most appearances and occupation
8. Ted Language -- is a good indicator to say the popularity in the world
9. durations
10. tags -- choose several and know the top 5 talks

- **(Extract)**

For this project we used a CSV file to import data we need to work.

This dataset is about all TED talks from 1972 to 2017 which provides information about talk title , main speaker , talk date , publish date , duration and related links and tags.

- **Transform**

We import this CSV file to python jupyter notebook and used pandas to import information to DataFrame so we can easily do our filtering, aggregation and to complete process and find proper answers to questions that we defined.

We also used matplotlib and seaborn to visually present out data.

- **Load**

At last , all our result are return as pandas dataframe and visually presented by different plots