



## **Business Intelligence and Visualisation (B9BA104\_2223\_TMD3)**

### **Aim:**

Design and develop an effective data visualisation for a dataset of your choosing, and compile a short report (1200 words) on the process you used to create this visualisation.

**Submitted to - Rudi O'Reilly Meehan**

**Submitted By - Mamta Devi  
Student ID - 10616210**



### Source:

<https://www.kaggle.com/datasets/thegupta/ted-talk>

### Abstract:

The dataset is about 'TED Talks' are short, impactful lectures that present a variety of viewpoints and themes. It was established in 1984, "TED" stands for Technology, Entertainment, and Design. However, TED Talks have now expanded their topic from physics and psychology to art, education, and more. TED Talks have gained a lot of popularity over the years, largely as a result of their inspiring and thought-provoking nature. Since they were widely shared on a number of sites, including as the official TED website, YouTube, and other social media platforms, they are now accessible to a global audience. There are TEDx events, which are independently organised conferences, in addition to the standard TED Talks. TEDx events are separate from TED but are TED-approved, locally organised, and feature speakers from the host region. TED Talks have grown into a powerful platform for sharing cutting-edge ideas, igniting social change, and raising people's awareness of important topics. The vast collection of inspirational and educational resources available to the general public has increased significantly as a result of the many well-known professionals, thought leaders, and significant figures from a variety of professions who have delivered TED Talk.

### About Dataset:

Each column includes Metadata (general/speaker/talk related information) as well as the Transcript, and each row corresponds to a Talk on TED.com

### About Dashboard:

The dashboard shows total number of talks, views, comments, tags, view count by speaker name, comment count by speaker name and count of video type by speaker description.



Card Chart has been used to visualise the total talks, views, comments and tags because a card chart is a type of visual display that only shows one value, making it perfect for displaying a certain number.

**Talks-** The term "talk" when referring to TED Talks denotes a speaker's presentation or speech at a TED event. The term "TED Talks" stands for "Technology, Entertainment, Design," and they are short, impactful talks that cover a variety of subjects, such as science, technology, the arts, culture, global issues, and personal growth. TED Talks are recognised for their structure, which enables presenters to quickly and compellingly communicate their ideas, viewpoints, and life lessons. The purpose of TED Talks is to inspire audiences, promote ideas and knowledge, and start thoughtful conversations. Speakers from a wide range of fields, including scientists, executives, artists, activists, educators, and many more, give talks on the TED platform. The TED Talks brand has gained enormous global recognition. TED conferences and events throughout the world regularly record the talks, and the recordings are then freely available online on the official TED website and YouTube channel.

However, the total number of talks as per sorted dataset on dashboard on TED talks are 3585.

**Comments-** A comment is a written comment or piece of feedback that viewers or other users who have seen the video have left underneath it. Viewers can express their thoughts, queries, and reactions to the video's subject matter through comments. Online video platforms must have video comments because comments can open up a line of communication between producers and their viewers, and they can have a significant impact on how people view and react to videos.

However, the total number of comments on the ted talks are 546K which is a significant and implying an engaging presentation with a large audience.

**Views-** The number of "Views" a video has indicates how frequently users have seen or played it. Every time a viewer clicks on the video and watches it, the overall number of views increases. Views are a critical metric for determining the value and appeal of online content. The presentation or video has undoubtedly resonated with a larger audience if it has had more views, which typically denotes higher attention and interaction. Views are essential to the success and impact of online content.

However, according to visuals there are 8bn views on TED talks videos and this remarkable accomplishment highlights the TED talks platform's enormous appeal, wide reach, and substantial influence as a platform for the spread of insightful ideas and inspirational presentations. The 8 billion times that TED Talks have been watched show how popular they are on a worldwide scale.

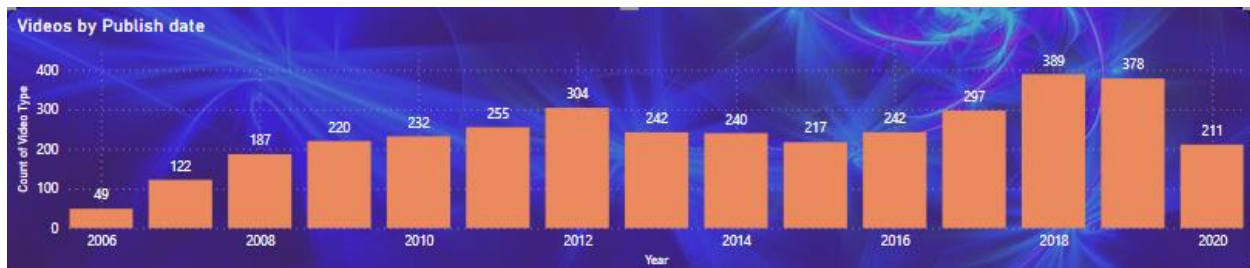
**Tags-** The terms or labels used to categorise and clarify each talk's topic are known as "tags" on TED Talks. By using tags, users and viewers can quickly find speeches that address certain topics they are looking for or that align with their interests. This makes it simpler for viewers to find and interact with relevant and engaging information.

However, the total number of tags as per visuals are 28K. The 28,000 tags on TED Talks show the depth and breadth of the platform's content. It exemplifies TED's commitment to presenting a wide

range of speakers and opinions, making TED Talks an indispensable tool for education, inspiration, and investigating a variety of topics.

**Clustered Column Chart** – A clustered column chart is a type of chart that displays data in a column format with columns grouped together based on specific criteria. The clustered column chart is quite useful for comparing data across numerous categories or groups and examining how the data values change within each category. It provides a precise visual representation of how large or small each data series is in comparison to others in each category.

However, the below cluster visual is showing the number of videos that were made available annually between 2006 and 2020.



On X- axis it is showing the publish date by year and on Y-axis, the count of video published on that particular year.

Visual shows:

- In the year 2006- 49 videos were published
- In the year 2007-122 videos were published
- In the year 2008-187 videos were published and it shows data till 2020 as in the year 2020-211 videos were published. The sum of the total videos 3585 were published from 2006 to 2020.

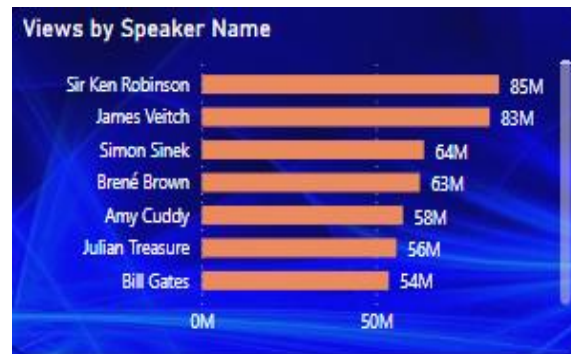
**Clustered Bar Chart**- A clustered bar chart is a form of bar chart that groups vertical bars next to one another to display data, with each bar representing a category or group. Similar to a standard bar chart, but with the bars grouped together rather than stacked on top of one another, it is easier to compare data from different categories.

The below visual shows the top seven popular speakers whose videos are watched frequently as the number of views on their talks or videos are tremendously high.

As per the visuals, Sir Ken Robinson is on the top on the list as his videos have total number of views across 85 Million which is a significant number. It shows the Popularity, Impact, Engaging Content and Influence on a large number of audiences.

James Veitch 83 M which are also a great number and shows a great impact on audiences and viewers.

We can see that all the popular speakers have amazing view counts on their videos or talks however, it shows audience have a greater level of interest and engagement with the speaker's content is typically indicated by a higher view count.



In the below visual on X-axis, it is showing the count of comments by Speaker Name on Y-axis. The visual is showing top 7 speakers who has the maximum numbers of the comments on their videos on TED talks. On top **Richard Dawkins** has 7K comments on his talks and videos on TED talks which is significant number and shows great engagement with the audience.



**Table Cart-** A table, often known as a table chart, is a methodical way to present data in rows and columns. The organisation of data in a tabular format is a fundamental type of data visualisation. The table's columns each list a different attribute or variable related to the distinct record or data point for each row. Table charts are widely used for information presentation that is planned and structured as well as for data entry and data comparison. They are often used in spreadsheets (like Excel and Google Sheets) and other data processing programmes. Additionally, tables can be used into books, articles, or websites to display data in a structured and clear manner.

| Speaker Description | Count of Video Type |
|---------------------|---------------------|
| Writer              | 54                  |
| Journalist          | 46                  |
| Entrepreneur        | 43                  |
| Artist              | 42                  |
| Architect           | 41                  |
| Designer            | 38                  |
| Neuroscientist      | 35                  |
| Educator            | 28                  |

However, Above table is showing how many video formats there are that are related to a given speaker's profile or occupation. The "Count of Video Type" column lists the number of videos that are accessible for each speaker description, and each row corresponds to a certain speaker description.

- Authors are speakers in 54 videos.
- There are 46 videos associated with Journalists.
- There are 43 films available that discuss business entrepreneurs.
- 42 videos are linked to the artist.
- There are 41 videos on the topic of architects.
- There are 38 videos associated with Designers.

### **Justifications of your choice of visualisation technique:**

The tool and software I have used for making this dashboard is power BI. I chose power BI over Tableau. Both Power BI and Tableau are well known for their user-friendly interfaces and drag-and-drop capabilities but I found Power BI more user friendly than Tableau as Power BI offers a big advantage in terms of data source connectivity because to its close integration with other Microsoft products including Excel, SQL Server and SharePoint.

Also, before data is visualised, it is easy to clean, modify, and shape it using the robust data preparation tool Power Query, which is a component of Power BI as in Tableau data preparation features are less robust.

### **Conclusion:**

This dataset is not only beneficial for the audience but it is also helpful for the ted talks business itself as this data set contains a lot of information which is beneficial for the organisation point of view as well as audience point of view.

- The data set is giving clear picture of business to TED talks, how popular are their talks by the total numbers of views and comments on their published videos from 2006 to 2020.
- They have popular speakers on their platforms which are making TED talks even more famous in the market.
- Also, it is giving opportunity to every domain or description of a speaker.
- The fact that TED Talks give speakers a platform to share innovative ideas, disseminate knowledge, and promote meaningful dialogues has had a big impact on many generations.
- TED talks inspiring and benefitting different generations by Inspiring and Motivating, Bridging Generational Gaps, Igniting Social Change, Personal development, Empowering Youth, Encouraging Public Speaking and Communication Skills.

While preparing the dataset for dashboard I had to clean, sort, filter and transform the data as there were a lot of semicolons and single quotes which were making it very complexed and unclear to understand. Also, Data set had mostly categorical columns and less integers so I had to sort them to make dashboard visuals clearly understandable.