**TED Talk Analysis**

**TABLEAU LINKS:**

Initial: <https://public.tableau.com/profile/sarah3638#!/vizhome/TEDTalksAnalysis/Story1?publish=yes>  
Final: <https://public.tableau.com/profile/sarah3638#!/vizhome/TEDTalksStory/Story1?publish=yes>

**SUMMARY:**The graphs and illustrations in the Tableau show popularity of TED Talks. Initially, I tried to get a sense of what factors might affect the popularity of a talk. I chose to focus on number of speakers, views, comments, languages, duration of the talk and date of the talk to investigate which features play the most significant role in the popularity of a talk. Through the various visualizations, it can be seen that most of TED talks are short and effective. Most popular talks have a high number of comments. Moreover, popular talks are translated to too many languages, so it can reach more audience, the opposite is true as well. secondly, I found that most of the talks were held in February, and least were held in January and August.

**DESIGN:**

1. First of all, I wanted to give a clear graph that shows number of speakers for each talk. A Bar chart was my best approach here.
2. Next, to give the user a better understanding to what makes a talk popular, I decided to use a bar chart to represent duration of each talk. However, duration was in seconds in my original data, so to make it readable, I a new calculations duration in minutes which is duration in seconds divided by 60.
3. To further strengthen my analysis and findings, I decided to see when most TED talks were held. To facilitate this, I used a bar chart with month of film date.
4. Finally, I created a line chart to illustrates the correlation between comments, views and languages over time. However, the range of each dimension is different; therefore, I divided the graph to three different line charts to make it clear for the reader.

**FEEDBACK:**

After my initial thoughts, I showed my illustrations to a data engineer and his feedbacks were really helpful!  
Here were his pointers:

1. Suggestion: Y-axis can’t tell so much about the data since total number of TED talks in the data is unknown.

Implementation: I changed y-axis to represent the percentage instead of count.

1. Suggestion: A line chart is not really helpful when showing duration of each talk. Implementation: I changed it to a bar chart.
2. Suggestion: To show the relationship between comments, views and available languages, it would be better if you added them in the same graph.   
   Implementation: A created a new graph with three line charts to represent average comments, views and languages.

**RESOURCES:**

None