

Project Proposal

Project Title: Ted Talks Timeline

Names: Gauri Joshi
Neha Bandal
Emails: grjoshi@dons.usfca.edu
npbandal@dons.usfca.edu
UID: Gauri: 20451007/grjoshi
Neha: 20404212/npbandal

Link to The Project Repository: <https://github.com/nehabandal/DataViz-Final-Project>

Background and Motivation:

TED is a nonprofit devoted to spreading ideas, usually in the form of short, powerful talks (18 minutes or less). TED began in 1984 as a conference where Technology, Entertainment and Design converged, and today covers almost all topics — from science to business to global issues — in more than 100 languages. Meanwhile, independently run TEDx events help share ideas in communities around the world.

We have always been fascinated by TED Talks and the immense diversity of content that it provides for free. We were also thoroughly inspired by a TED Talk that visually explored TED Talks stats and were motivated to do the same thing, on a much less grander scale.

Project Objectives:

The primary objective of this project is to show the timeline of different Ted Talks based on different categories. We are also planning to provide filters for Speakers and Talks. Along with these primary objectives some of the questions that can be answered from this project are:

1. How is each TED Talk related to every other TED Talk?
2. Which are the most viewed and most favorited Talks of all time? What does this tell us?
3. What kind of topics attract the maximum discussion and debate (in the form of comments)?
4. Which months are most popular among TED and TEDx chapters?
5. Diversity of speakers(Different category of individuals sharing their related or unrelated experience)?

Data:

The Ted Talks Timeline is based on talks subject and speaker statistics aggregated from <https://www.kaggle.com/rounakbanik/ted-talks>.

These datasets contain information about all audio-video recordings of TED Talks uploaded to the official TED.com website until September 21st, 2017. It contains information about all talks including number of views, number of comments, descriptions, speakers and titles. If require we will find more relevant data.

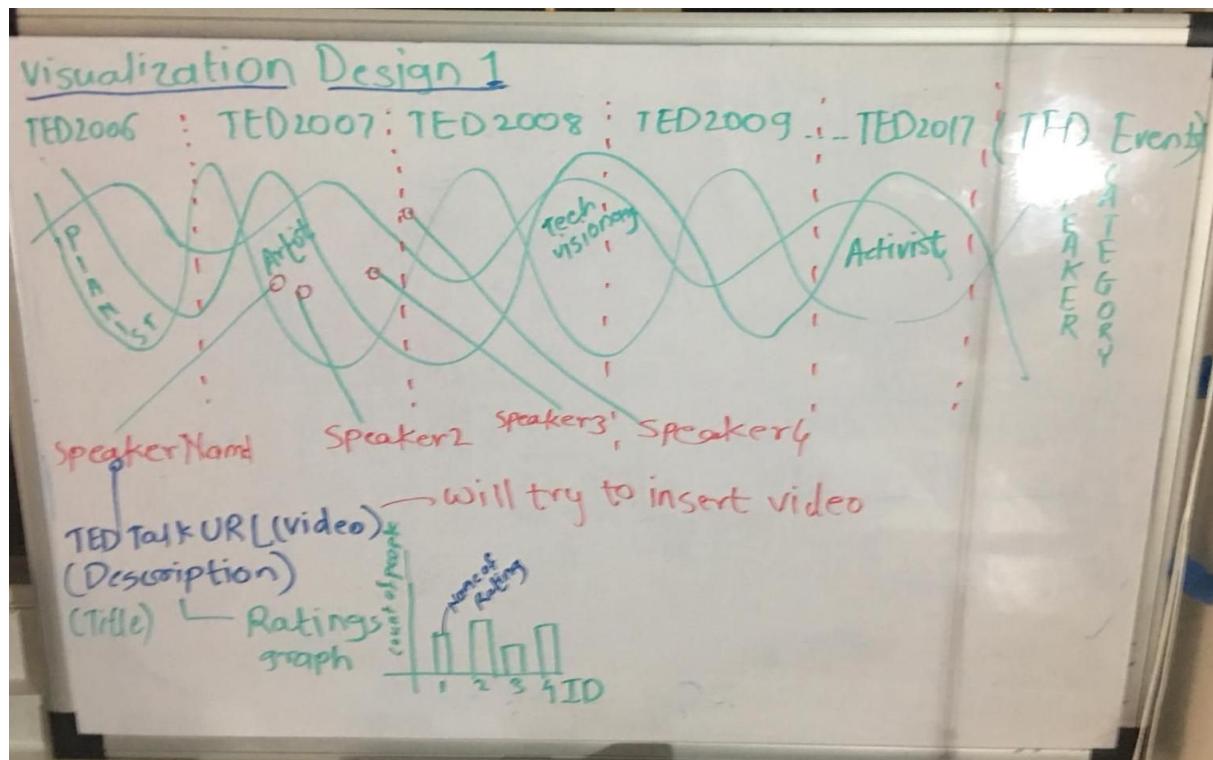
Data Processing:

Based on our initial analysis there are different columns which requires cleanup and processing. For e.g. Film Date and Published Date are UNIX timestamps we will need to convert it into Javascript/D3 Date format. Also we will also require to normalize data to some extent as few columns consist of JSON objects. For e.g. ratings, related talks etc.

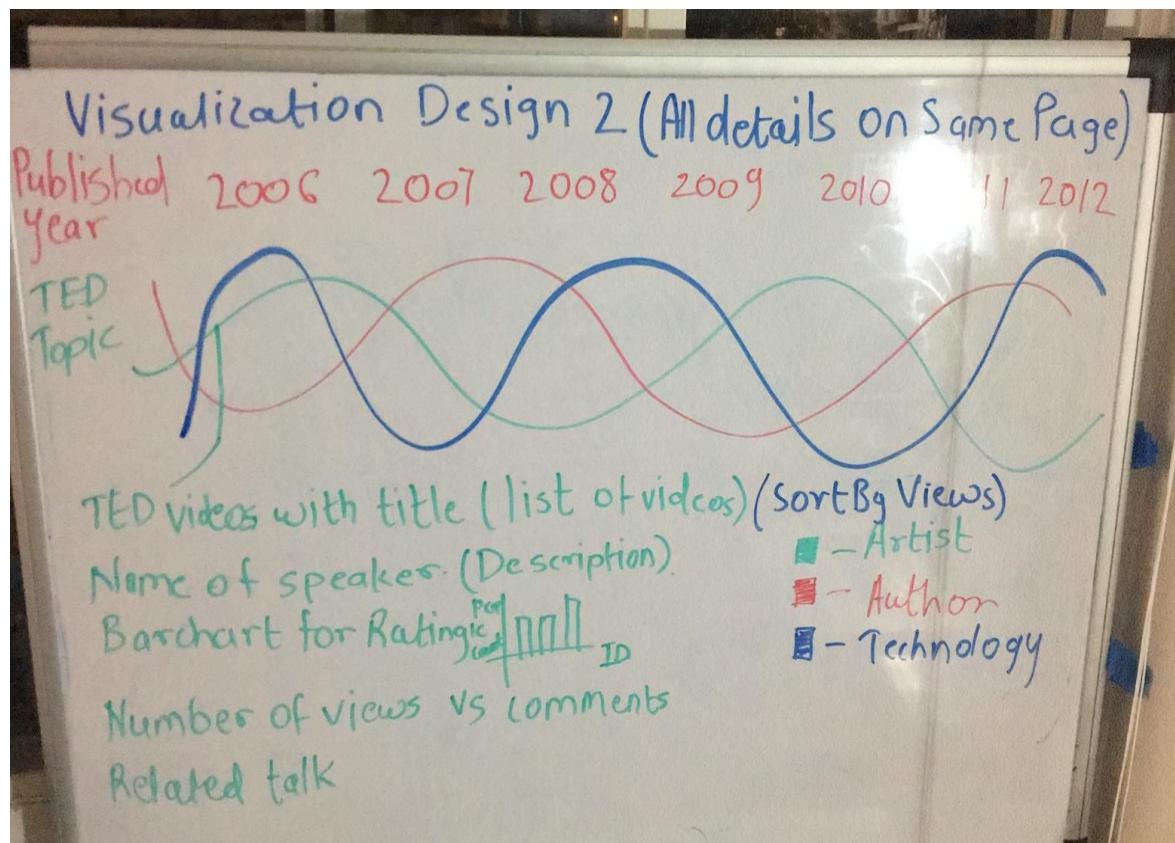
Visualization Design:

Please find below the screenshots of visualization designs for our project.

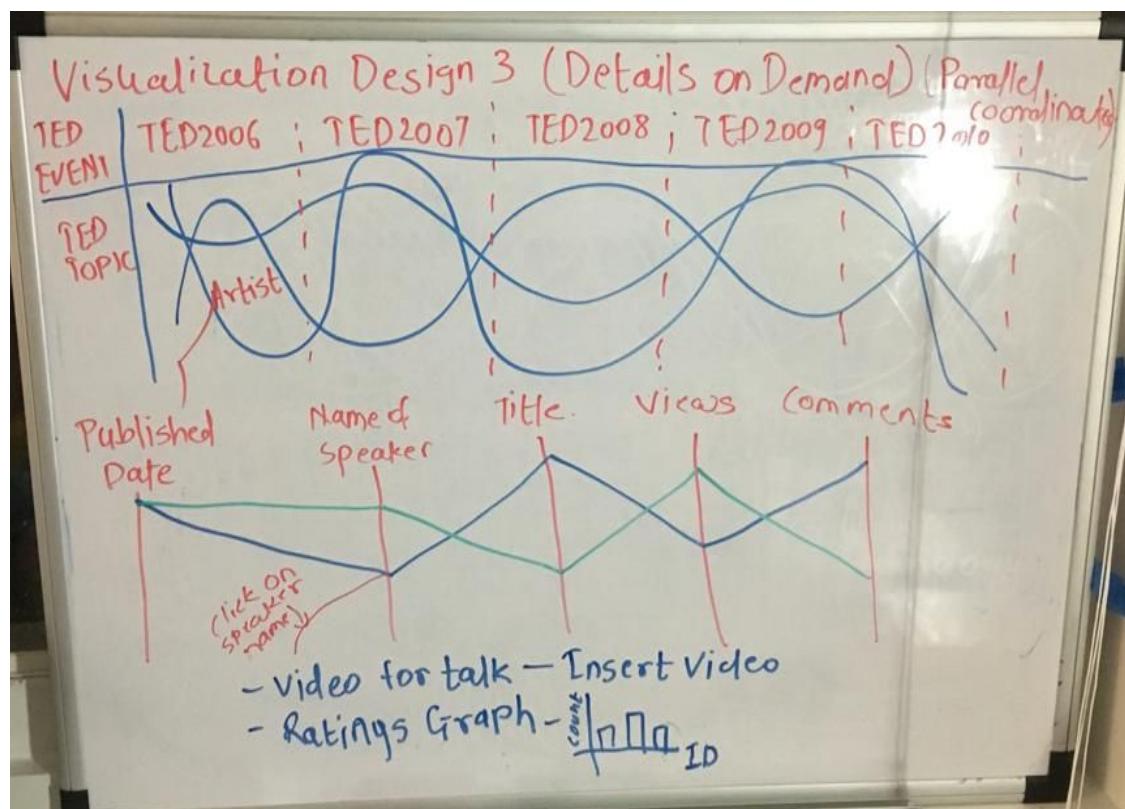
Design 1:



Design 2:



Design 3:



Must-Have Features:

1. Visualization with timeline to show different categories of Ted Talks or occupations of authors.
2. Details of each Speaker, Ted Talk will be shown in different visualizations.
3. When user will click/hover on any particular category the links will be drawn to speakers/related Ted Talks (Ted Talk Videos).

In this project we are focusing on Overview+Detail, Zooming, and Focus+Context Interfaces concepts. Along with these features we will have Interaction, Great storytelling(including effective use of colors, animation, annotation/ labels, depth in layers, simplicity and consistency) in our all visualizations.

Optional Features:

1. Filter for Ted Talk Title and/or Speaker.
2. Interactive visualization for Review analysis. The main purpose of this visualization is to see how audience feels about particular Ted Talk.
3. Visualization with relevant related Ted Talks. If user wants to get more details about any particular Ted Talk this visualization will help them to get more details about relevant Ted Talks.

As we progress in our project we may modify/add features a bit. We will keep everything updated in our Process Book.

Project Schedule:

We will divide the tasks once we finalize the modules. But tentative weekly deadlines for different modules are as below.

Week	Milestones / Weekly Deadlines
Week 1 (April 5-11)	 Brainstorming about Project Design.  Proposal, Website, Data Gathering Phase
Week 2 (April 12-16)	 Alpha Release Data Acquisition, Cleaning and Processing. Decide Data Structures
Week 3 (April 17-25)	 Beta Release Basic Visualizations Up and Running
Week 4 (April 26-30)	 Work on Interactions, Website
Week 5 (May 1-7)	 Project Presentation, Other Visualizations, Coding

Week 6 (May 8-14)	 Align Visualizations as per Design, Testing, Develop Visualizations if Any Remaining
Week 7 (May 15-16)	 Complete Process Book, Final Code Deployment, Testing