Final Report of
Visualization Project

Causes of Death

Done by:
Temirlan Ulugbek uulu (t.ulugbekuulu@jacob-university.de)
Zhao Liu (zh.liu@jacobs-university.de)

Abstract

We have built a simple and easy to use visualization tool which allows to explore the causes of deaths in United States. The final product of our project is a web-based vis tool built by using HTML, CSS and Javascript/D3.

Motivation
(Background, tasks, users, and data)

The idea is to build a web-based tool that would allow to analyze different causes of death. Mostly people would like to avoid talking or thinking about death and it is not the nicest information one could get. Thus, no serious organization would like to invest into making a nice visualization tool, that would make exploration of death causes easier and more interesting. Those who are interested in data could only learn from articles or, even worse, raw data. We have decided to build a simple and easy to use visualization tool which allows to explore this dark area. Once everything was visualized nicely, one can easily be attracted to the exploration and get answers to the questions like, "What are the most popular causes of death?", "How did the rate of this cause change over the years?", "Which states have the most deaths due to cancer or murder?", "How did the rate change in LA comparing to the average rate across all states?". After exploring the data and getting more informed, one may even try to live better life and avoid the popular causes. For example after seeing that the Heart diseases are the most popular one across all the states, one may start doing sports and try to decrease the probability of getting heart problems.

User base of this tool is anyone who is interested in exploring the data and it doesn't need any expertise in the field. People can explore the causes of death in United states for the period of time starting from 1999 up to 2015. The dataset was taken from the source below:
https://catalog.data.gov/dataset/age-adjusted-death-rates-for-the-top-10-leading-causes-of-death-united-states-2013

Related work

As have been mentioned above, no visualization tools are publicly available to explore the causes of death (at least we weren't able to find one). The reason is that the majority of people would prefer not seeing this data. No serious organization would invest funds into creation of such a tool. All available sources are articles and raw datasets.

Another idea we had, was to incorporate another dataset with life expectancy at the birth in United States (https://data.cdc.gov/NCHS/NCHS-Death-rates-and-life-expectancy-at-birth/w9j2-

ggv5). The data had average expected life of a child according to the year he/she was born, race, and sex. We found this dataset very interesting and were trying to make the tool which would allow to explore both datasets. Later after milestone 1, we realized that it was bad idea, cause the core data in the second dataset was different and was out of the scope of this project. Adding that database would bring confusions and make tool more complicated. Moreover, we would deviate from our goal that we put for this project. Thus we decided not to use that dataset and keep our tool simple.

In terms of technologies, we mainly used simple HTML and CSS as a core skeleton of the website. Then we used D3 to build dynamic visualizations. All the code that has been taken from online sources was referenced in the comments at respective places.

Approach
(Used techniques: filtering, zoom, linking and brushing, tooltip)

The first step we did in making design decisions was to explore the dataset. Since we didn't have the flexibility in terms of the data that we could have for our tool, we had to adjust our vis design and complete tool according to the dataset we had. The dataset had simple structure: Cause, State, Year, Count, Age-Adjusted Rate. Since we had the same data points for different times, we decided to implement a time-line view which would be linked to all the other views and allow filtering according to the selected period of time. Also we decided to show the changes of rates over time in the line chart. As one changes the time period, the line-chart **zooms in and zooms out**, i.e. the scales and axis change according to the period. Since dataset was continuous and chronological, the line-chart was the best design decision in our point of view. Then one might want to see the ranking of the states in some death cause or vice-versa rating of causes in a specific states. Those things were implemented using simple bar-charts which would make it easier to compare between states or causes. Even though we have implemented simple input fields on the header of the website, one could use the bar-charts as **filters** as well. We have **linked** them to other views, which allows user to choose state or cause directly by clicking the according bar in the chart. This makes the exploration more convenient. Users also have the **tooltip** convenience, i.e. when using bar-charts tooltip with info appears when you hover over the bar. The last view we decided to implement was heat-map across the states to give the spacial idea to the users. One can see the state they are exploring or if state is not chosen, user can see where is the specified cause of death more popular.

Implementation

Simple HTML&CSS were used as a backbone for the website. Some snippets of codes were taken from online resources and linked in the code itself. Vis views were mainly implemented using the D3 library. It gave the convenience in handling the dataset and building the views.

The main challenges were putting the views together and making them work as one dashboard. Views, which work perfectly fine on separate pages, ended up being not working, hidden etc. when put together with other views. Also creating the views themselves was different due to the lack of experience in working with d3 library. Working with SVGs can also be tricky if you don't know what you are doing. The way we handled the challenges was mainly by trial and error technique. Spending more time over the problem and getting more familiar with the technologies usually was enough to solve the problems.

Results

Bla-bla..

Discussion

Bla-bla..

Task separation

More bla-bla…

References:

Any?