CS498dv Narrative Visualization Essay

Messaging:

With narrative visualization, the correlation between various factors and income from the first four scenarios can be investigated. The conclusion could be the strongest factor affecting adult income is education.

From the last scenario, the income distribution across states could be seen. The conclusion could be California, Illinois, Texas have the most well paid people.

Narrative Structure:

Interactive slideshow. The structure is demonstrated using 5 scenarios that are trying to find correlations by telling a story. The visualization guides the user through the slides, with an order that best conveys the message, but still, it gives the user the flexibility to drill down and investigate further in each slide. For instance, the annotations are cleared after 25 seconds so the user gets the chance to look at the visualization and investigate further without distraction. Once the user finishes the investigation, he can use the navigation button to proceed to the next scene. Once the button is clicked, it triggers an action that opens a new page to render the next scene, and changes the parameters that trigger different data and annotations relevant to the scene.

Visual Structure:

Interactive slideshow structure. There are multiple tooltips which show different information of data and the user is able to see the parameters by moving mouse to data points. By labels and points, it highlights to urge the viewers to focus on the variables. The other scenarios have similar and consistent display and view. Data description and graph legend explains the connections between different scenarios data. Because of the multiple slides with navigation buttons, you can navigate backward and forwards by advancing slides. Regarding the structure, a scene template is used for visual consistency, so each scene has one chart and a brief description text that summarize the message that the chart is trying to convey. Also, the transition is used to keep the user oriented as much as possible.

Scenes:

Scene 1: The narrative visualization starts with an overview of the entire data using a scatterplot chart with two dimensions that can best spread out the data, which are "Age" and "Weekly working hours". The adult income, that can be either above 50 thousand dollars (>50K) or below or equal to 50 thousand dollars (<=50K), was highlighted using a discrete color palette. We can notice that the area with the highest income is in the middle close to the average worker adult age (40 years) and average working hours (40). This means that other factors might have a larger effect. So further investigation is needed.

Scene 2: Here we are looking on how Occupation affects the income. We plotted the Number of Adults for vs. each occupation segregated by the income. We can see that Executive managers and specialized professionals have the higher wages among all other occupations. But still there is a higher percentage of those two groups taking wages less than 50K, which means that there might be other stronger factors that we might need to look at.

Scene 3: In this scene, we are looking at the Marital Status factor. Similar to scene 2, we plotted the Number of Adults vs. the Marital Status segregated by the income. We can see that even though the Married people, in general, have a higher percentage of the >50K than any other Marital Status, but still the majority of married people has income less than 50K, which means that the Marital Status probably not the most effective factor.

Scene 4: Now if we look at the education level, "doctorate" for example, we find that education level has a high effect on adult income because the number of >50K group is higher than the <=50K/. the same for other postgraduate education categories, Graduate School, and Masters, which clearly proves that education level is affecting adult income.

Scene 5: High income people distribution across states.

It follows a linear ordering, the scenes are ordered linearly so as to demonstrate the correlations between variables. From least correlation to most and obvious, the last graph is more like an overall view while the first four are specific to the categories.

Annotations:

Moving bubble circle. It is easy and obvious to see, and visually friendly. The annotations are used as one of the messaging techniques to highlight and emphasize some of the interesting aspects of the visualization. For example, in scene 1, the annotation is used to circle the area where highest income people are in the age - work hours distribution, by the definiton of a higher percentage of the >50K income group than the <=50K. Regarding the template, all the annotations follow one template to support visual consistency, it consists of a text, linked with a line(s) to a circle(s), which points to the area of interest. Since each scene has a different chart, the annotations are cleared between scenes for convenience.

Also, to support the "interactive slideshow" narrative structure, the annotations are cleared after 25 seconds to give the user the chance to look at the visualization and investigate further without distraction. The annotations are consistent and do not change over every scene.

Parameters:

Scene numbers are the main parameters that control the narrative visualization. As a response to the user action of advancing to the next scene, the button triggers an action based on the scene number to render the corresponding scene's chart, data, annotations and messaging. Hence, it controls the state machine of the narrative visualization. Also, for the last graph, year is the parameter to control scenes. Choose a year and you can see the corresponding stats and distribution in that year by a slider.

Triggers:

The main trigger is the user event of a "button click", which triggers the scene number parameter, which controls the state machine of the narrative visualization. Another trigger is the user event of placing the mouse on an element, which triggers the tooltip. Those available use events are communicated in the welcome page of the visualization in order to make sure the user is aware of the available triggers. Also, in the last graph, the triggers in this visualization are user interface controls. The user can drag a slider to select

a year. Thus, corresponding stats and distribution would be shown. These triggers are easy to see, click and pass by through the mouse. Switch over scenes, check data points in scenes, slide over years are options to users in the narrative visualization.