Alina Mawji
CS 416 - Data Visualization

Breast Cancer Incidence and Survival Factors

- **Messaging**. What is the message you are trying to communicate with the narrative visualization?
    - This narrative visualization aims to present various factors that play into the likelihood of contracting and surviving breast cancer, including age, cancer stage, tumor size, race, and marital status, through graphical representations. The primary message of the visualization is that (1) breast cancer mostly occurs in women who are 40 years of age or older, (2) tumor size alone does not have a correlation with estimated months of survival, and (3) race and marital status show differences in survival likelihood.
- **Narrative Structure.** Which structure was your narrative visualization designed to follow (martini glass, interactive slide show or drop-down story)? How does your narrative visualization follow that structure? (All of these structures can include the opportunity to "drill-down" and explore. The difference is where that opportunity happens in the structure.)
    - My narrative visualization is designed to follow the interactive slideshow format. It follows that structure since it uses separate pages for each scene of information, and viewers can navigate through the data-driven story in a linear fashion, gaining additional insights with each slide. Within each slide, users can drill down into specific details to interact with and learn more about the data.
- **Visual Structure.** What visual structure is used for each scene? How does it ensure the viewer can understand the data and navigate the scene? How does it highlight to urge the viewer to focus on the important parts of the data in each scene? How does it help the viewer transition to other scenes, to understand how the data connects to the data in other scenes?
    - In each scene of the visualization, a distinct visual structure is employed to effectively convey different aspects of the data. Each scene is displayed on a slide with a guiding question as the title and accompanying supporting information. These aspects of the slide help the viewer understand the purpose of each scene and provide context for the data. The graphs themselves each contain labeled horizontal and vertical axes with specific units, along with the consistently-colored key takeaways written above to help the user understand the data points, understand important parts of the data, and answer the question. The graph also

contains consistently-colored and sized annotations that help highlight important parts of each graph. Transparent colors are used to encourage the viewer to mouse over and interact with the visual to further understand the data. The contrast in colors direct the viewer's attention to important parts of the data, such as the larger incidence of T1 and T2 cancers over T3 and T4. The viewer is also provided with consistently-placed "Previous" and "Next" buttons at the bottom that help in navigating between scenes, and use a standard header and card formatting that help standardize the visual structure. Similar colors and font sizes are used for cancer stages, key takeaways, and annotations between scenes to help the viewer transition and connect data and insights between scenes.

- **Scenes.** What are the scenes of your narrative visualization? How are the scenes ordered, and why?
    - The scenes for the narrative visualization are, in order, (1) comparing age and incidence of breast cancer cases, segmented by cancer stage, (2) tumor size and estimated months of survival, segmented by cancer stage, and (3) race or marital status by rate of survival. This sequence is structured to lead the viewer through a logical progression of understanding breast cancer risk factors according to stages of the disease. Scene 1 starts with incidence comparisons to establish what point in life the disease may develop. Scene 2 builds upon this by exploring risk factors for survival (tumor size and cancer stage) and their influence on estimated months of survival, deepening the viewer's understanding of disease progression. Finally, Scene 3 introduces demographic factors and their impact on survival rates, providing broader insights. This ordered approach ensures that viewers can understand the various factors that play a role in the development of breast cancer, starting with incidence and progressing to estimated survival months and finally, survival likelihood.
- **Annotations.** What template was followed for the annotations, and why that template? How are the annotations used to support the messaging? Do the annotations change within a single scene, and if so, how and why?
    - The annotations in each slide follow a consistent template of gray dashed lines with small gray text, strategically placed to point out specific details like data points or bars within the charts. This template was chosen for its muted yet clear distinction from the main content of the graph. This ensures that they are able to communicate messages without overwhelming the viewer. They are used to support the messaging by highlighting key insights within the data, guiding the viewer's attention to significant findings or correlations. In addition, they are used to inform the

user of the tooltip capability. On the third slide, the annotations change in content to address the different states of the data. This ensures that all parts of the graph contain annotations that communicate the takeaways of the data.

- **Parameters.** What are the parameters of the narrative visualization? What are the states of the narrative visualization? How are the parameters used to define the state and each scene?
    - In the tumor size scene, the parameter is the cancer stage. These parameters are visually represented using buttons with corresponding colors to distinguish between stages on the graph. Thus, the various states of the narrative visualization include each stage of cancer, and all of them: T1, T2, T3, T4, and All. Each scene is defined by a selected state when a user clicks one of the parameter buttons, as it triggers a filter on the data through a JavaScript function. This results in a re-rendering of the data displayed by the user's selection, illustrating how tumor size impacts survival months through color-coded dots.
    - In the demographic scene, the demographic grouping is the parameter, and the states are either race or marital status. Using a dropdown box, users can define which state they are interested in, and this triggers a filter on the data through a function in the JavaScript.
- **Triggers.** What are the triggers that connect user actions to changes of state in the narrative visualization? What affordances are provided to the user to communicate to them what options are available to them in the narrative visualization?
    - In the tumor size scene, the triggers connecting user actions to changes in state are the selection of cancer stages through clickable legend buttons. Each click on a button dynamically updates the graph to display data according to the selected stage(s), altering the color and position of data points accordingly. Affordances provided to users include visually distinct, shaded boxes around each cancer stage, which are highlighted further when a user hovers their mouse over a button. This setup communicates the available options for exploration.
    - In the demographic scene, the trigger is a selection within a dropdown box for either race or marital status. Each click on a dropdown option updates the graph dynamically to display data according to either grouping. Affordances provided to users include text prompting the user to "Select grouping." This setup communicates the ability to change the data displayed.