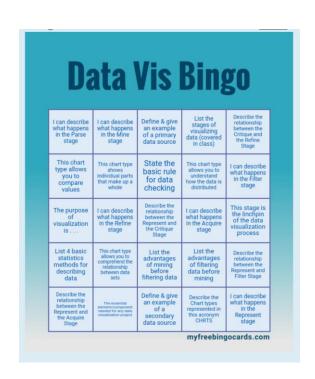Lab 06: Midterm Review Bingo

CGT 270 – 007

Megan Jacobs

Bingo Card:

Questions

1. **Define & give an example of a primary data source**
   a. An example of a primary data source is the top 100 songs listened to on Spotify. And the data has to come from Spotify to be a primary source because it needs to come from the original researcher at Spotify. Some more examples of primary data include:
      i. A student records the amount of times they miss class throughout the semester
      ii. Purdue records the amount of students graduating from Purdue each year
2. **List the stages of visualizing data (covered in class)**
   a. Acquire, Parse, Mine, Filter, Represent, Critique, Refine
3. **Describe the relationship between the Critique and Refine stage**
   a. The critique and refine stage's relationship is very direct and what fuels each other. Once the person has created the visualization, they will receive feedback or critiques from user's, peers and other participants and then refine their visualization. And then it will start over with critiques again. This looping relationship could happen many times until the author and the users are satisfied with the final visualization.
4. **This chart type allows you to compare values**
   a. Categorical

5. **This chart type shows individual parts that make up a whole**
   a. Hierarchical
6. **I can describe what happens in the filter stage**
   a. In the filter stage, what we are trying to accomplish is taking the data that we have and taking out parts of it to showcase a smaller subset of the data. We usually do this because it either fits the purpose or answers a question better than using all of the data and it may showcase patterns that were otherwise hidden.
7. **I can describe what happens in the Refine Stage**
   a. In the refine stage what we are doing is, taking the visualization we had create in the Represent stage and using the feedback we got in the Critiquing stage and making changes to better the visualization that we have created. This can range from creating a whole new chart type, to adding more labels and annotations to even just changing the color scheme of the visualization.
8. **Describe the relationship between the represent and the critique stage**
   a. The represent stage is the stage directly before the critique stage, and the visualization created in the represent page is the object that is being critique. So without the representation we would not be able to critique and better the visualization.
9. **List 4 basic statistics methods for describing data**
   a. Average, Mode, Median, Maximum, Minimum, Range and Length
10. **List the advantages of mining before filtering data**
    a. Some advantages to mining before filtering data is that patterns that we may have thought we hadn't seen any patterns in the data but with mining before filtering we may see those patterns. Also, it may guide us to filtering subsets of the data we hadn't previously thought to.
11. **List the advantages of filtering before mining the data**
    a. Some advantages of filtering before mining data is that you would not have to mine data that you will not use later on. By filtering we are only using a subset of data and that way it saves us some time mining.
12. **Describe the relationship between the represent and the filter stage**
    a. The filter stage allows the author to have the data that is directly answering the question they are trying to answer and create a visualization to answer the question. Without the filter stage we could see visualizations that don't have a purpose or don't answer a question in a clear manner. It may confuse the audience when the data is filtered to see patterns or to be able to answer questions.
13. **Describe the relationship between the represent and acquire stage**
    a. The represent stage is taking the data we gathered in the acquire stage and presenting findings in the data into a visualization. If we did not start with acquiring the data in the first place we would not be able to move onto later stages such as the represent stage.
14. **Define & give an example of a secondary data source**
    a. A secondary data source is any data source that is collected by anyone else that is not the researcher. So this data comes from things like surveys, questionnaires or experiments. Some secondary data sources can be websites, research journals, and books/textbooks.