Information Visualization

Homework 2

Sourabh Shetty

1. Detailed Data Schema of the data depicted by the S&P Map Visualization:

Dataset Types:

In the default view, while the visualization is presented in a seemingly structured tabular form, the dataset type is actually network. The major categories (sectors) are arranged in a way so as to group the closest related ones together. Similiarly, there are subcategories (industries) within each category that are also grouped together based on how closely related they are.

The sectors, industries, and each individual stock are the nodes of this network, and the mere existance of adjacency between any two nodes acts as a link between them.

Hovering over an industry or individual stock displays another visualization of data. This shows the performance of the stock hovered upon against all the other stocks in that industry. This dataset type is a table, since it plots a key against its values in a simple tabular format.

Data Attributes and Types:

Sector – Categorical

Industry – Categorical

Name of Company – Categorical

Market Cap – Ordered, Quantitative

% Change – Ordered, Quantitative

Stock Price – Ordered, Quantitative

Stock Color: Ordered, Ordinal

2. Ans 2

1. Retrieve Value: What was GOOGL's stock price on the day?
   Using Map: 942.02. This required hovering over GOOGL and then it was listed on top of the table that popped up.
   Using Bubble: 942.02. This also required hovering, but in the dialog that popped up, it's a bit more difficult to identify the actual stock price.

2. Filter: Which companies had a positive growth rate?
   Using the map, it's a simple matter of finding the names of all the green colored stocks.
   In the bubble view, it becomes a little more diffcult, since many of the stocks are overlap over each other, some even completely obscured.

3. Compute Derived Value: What was the number of stocks that grew over 4%?
   Using the map, again, it's difficult to identify stocks that grew over 4% since all stocks that grew beyond 3% are given the same green color.
   Using bubbles, it becomes a lot easier to track down the companies with a growth rate higher than 4%.

4. Find Extremum: Which stock had the highest growth rate?
   Using the map, it takes very long to figure out which stock had the best growth. Any growth over 3% is given the same color, so a lot of manual checking is involved.
   Using the bubble version, it becomes very easy to determine that GPS had the best growth, since the data is plotted against the growth rate itself, so using growth rate as key, it becomes very easy to determine which stock grew the most.

5. Sort: Rank the technology stocks by market cap
   Using the map, we can easily identify rankings since it's visually clear.
   In the bubble version, while the sizes are equally proportionate as the map, the fact that some are overlayed on top of each other makes it difficult to parse.

6. Determine Range: What is the range of stock prices for the information technology provider industry?
   In the map, on hovering over an industry we get to see all the stock prices in that industry in a neat tabular format which makes it easier to calculate the range, i.e, 45.69 – 943.72, however, in the bubble version, it's significantly harder to calculate, since there is no easy comparison table, and we would have to hover over each stock bubble individually.

7. Characterize Distribution: What is the growth distribution?

8. Find Anomalies: Are there any anomalies in the basic materials – oil & gas industry?

In the map, we can immediately identify that two stocks, APA and CHK have positive growth, while the rest are all negative.
In the bubble version, we can look either above or below the 0% line.
The map version is still more convenient however.

9. Cluster: Are there groups of stocks that have done well?
   In this question, the bubble version is superior in that we can immediately identify clusters as opposed to the map version.

10. Correlate: Is there a correlation between sector and growth?
    From the bubble version, we can see clearly that the entire healthcare sector has had a lot more growth than any other sector. There is a definite correlation.
    In the map we see a lot of green, but we cannot identify a definitely correlation.

There were some tasks to which the bubble visualization was more suitable, especially in identifying clusters and trends, whereas more direct questions were more easily answered by using the map visualization.