Homework 2

Provakar Mondal

1 Problem 1

- **Dataset Types:** Cluster. This Visualization shows a stock martket data in which industires are showed in a cluster based on their types like Technology, Telecommunicaiton, Consumar etc.

- **Data Types:**

  - **Items:** Industry. For example, Microsoft, Google, Cisco etc
  - **Attributes:** Total Stock Price, Percentage increment/decrement.
  - **Attributes Types:**
    i) **Quantative:** As the visualization depicts the stock price of the industries and this prices are quantative values.
    ii) **Categorical:** Industry are categorical as they are clustered based on their types.

- **Dataset and Data Sizes:** Dynamic data. And these dataset and data are scalable. New industry can be added if its stock status lies in the range of this visualization. Similaly, any industry can be removed from a cluster. Same condition is also applicable for the Cluster Dataset types.

- **Semantics:** This visualization gives idea about the stock market statistics of different types of industries. The big fish industries are visualized in a larger rectangle. Besides, types of industries who are dominating stock market is also visualized using larger rectangle as cluster size. For example, from this visualization it is understandable that **Technology-Software-Infrastructure, Technology-Consumer-Electronics, Communication Services** types industries are ruling the stock market.

2 Problem 2

Ten Task questions from visualization 1 (finviz Map)

i) Retrieve Value: Find the stock price difference between Danaher Corporation(DHR) and Waters Corporation(WAT).
ii) Filter: Find the industries whose stock price lies in the range 500 - 1500.
iii) Compute Derived Value: Make new attribute in each cluster by subtracting

each industry's stock from the max stock price wihin this cluster.
iv) Find Extremum: Which type of cluster posses the largest volume?
v) Sort: Order the industries in ascending by their stock price change.
vi) Determine Range: What is the range of market stock price?
vii) Characterize Distribution: Waht is the distribution of stock price in Technology-Software-Infrastructure?
viii) Find Anomalies: Is there any anomaly in the stock curve of VeriSign Inc?
ix) Cluster: Is there any cluster in this visualization?
x) Correlate: What is the relationship between cluster size and stock prices of the industires in it?

Ten Task questions from visualization 2 (finviz Bubbles)

i) Retrieve Value: What is the stock change value of PayPal Holding Inc (PYPL)?
ii) Filter: Get the industires within the stock change between -1% to +1%.
iii) Compute Derived Value: How many industires are in this visualization whose average volume is under 1M?
iv) Find Extremum: Find the industry with maximum percentage change of stock.
v) Sort: Sort the industry by their average volume size.
vi) Determine Range: Find the range of stock percentage change.
vii) Characterize Distribution: Find the distribution of bubble size based on stock percentage change.
viii) Find Anomalies: Is there any industry whose stock percentage changes shows exception than others?
ix) Cluster: Is there any group based on Communication Services in this visualization?
x) Correlate: Is there any relationship between bubble size of the industry and their market price?

Table results based on the task questions from 3 users

Map task																			
	Task A		Task B		Task C		Task D		Task E		Task F		Task G		Task H		Task I		Task J	
	time(s)	correct	time(s)	correct	time(s)	correct	time(s)	correct	time(s)	correct	time(s)	correct	time(s)	correct	time(s)	correct	time(s)	correct	time(s)	correct
User 1	57	1	177	1	567	1	12	1	413	0	345	1	543	1	34	0	02	1	08	1
User 2	41	1	282	1	290	0	07	1	519	1	293	1	489	0	76	1	02	1	06	1
User 3	93	1	300	0	731	1	15	1	862	1	306	0	614	1	23	1	01	1	13	1

Bubble task																			
	Task A		Task B		Task C		Task D		Task E		Task F		Task G		Task H		Task I		Task J	
	time(s)	correct	time(s)	correct	time(s)	correct	time(s)	correct	time(s)	correct	time(s)	correct	time(s)	correct	time(s)	correct	time(s)	correct	time(s)	correct
User 1	15	1	371	1	84	1	02	1	810	0	153	0	256	1	22	1	09	1	03	1
User 2	17	1	304	1	121	1	05	1	785	0	432	0	371	1	18	1	05	1	04	1
User 3	21	1	269	1	48	1	08	1	268	0	173	0	218	0	17	1	13	1	08	1

My Hypothesis

- Vis 1 contains more information at a first glance than vis 2.

- Though vis 1 contanins more info at first glace, users find it more difficult to depict information than vis 2.

- Users from who I have collected task answers, everyone thinks vis 2 is more east to retrieve information than vis 1.

- Viz 2 is more friendly in the respect of fetching information. The information in vis 1 is compacted rather than vis 2.

- I think for the 1st time users, vis 2 is better than vis 1 to represent information before them.