

Midterm exam - open book (papers), open notes, no discussion

1) Design a visualization to show the changes between two months in a 3-level tree visualization of employment categories where leaf nodes are the number of people employed in that job category for a given month. Draw and describe your design. Discuss your choice of the way to enable users to see (assume no deletions):

- a) the change in value from month to month (10 points)
- b) the appearance of new leaf nodes (10 points)
- c) the movement of nodes to a new subtree (10 points)

2) Make a list of three important tasks to be carried out in a social network of friendships among UM Computer Science graduate students, where each node has a name (and other attributes such as gender, country of origin, age, etc.) and each edge has a value in $[0,1]$ indicating strength of friendship. Assume a design goal of 300 students. For each of the three tasks indicate whether a node-link or matrix style interface would be more effective and why. (20 points)

3a) In the context of exploring multivariate data sets, MacEachren et al. make a strong point of using entropy as opposed to correlation coefficient for analysis of 2D scattergrams. Explain the distinction and draw example 2D distributions that clarify the distinction. (10 points)

b) The Hierarchical Clustering Explorer allows for entropy (uniformity), correlation coefficient and other criteria. Propose a useful novel criteria that is not described in the Seo & Shneiderman paper. (10 points)

4) Describe situations in which the slice-and-dice layout treemap is superior to the squarified layout? When is the squarified layout more useful? Why was the strip treemap a useful innovation? (15 points)

5) The GRID principles were designed to guide explorers of Multi-Variate data sets. Propose a similar set of principles for real valued temporal data, such as stock prices of thousands of companies over hundreds of days. (15 points) (this is the tough creative question)