

CSCI 4550/6550 Interactive Visualization

Sample Quiz 1 Problems

from a prior term, where we may have read different papers

The problems on this quiz are worth a total of 50 points.

This quiz is closed book & closed notes except for the (optional) notes page you uploaded to Submittity, which is attached at the back. You may separate this page from your packet if desired.

Please state clearly any assumptions that you made in interpreting a question.

Write your answer in the box provided below each question. Be sure to write neatly. You are encouraged to use the provided colored pencils, crayons, or markers.

1 Eenie, Meenie, Minie, Moe [/6]

Match each dataset or application below with the letter of an appropriate chart type (each letter should be used exactly once).

- (A) line graph
- (B) word cloud
- (C) pie chart
- (D) stacked vertical bar chart
- (E) horizontal bars
- (F) scatter plot

Number of people who develop different types of cancer in each country.

Breakdown of the population by health-care coverage: e.g., Medicare, Medicaid, Military/Veteran, Public Option, Employer Insurance, Private Insurance, Uninsured.

Presenting the results of a phone survey on why people don't go to the doctor.

Trying to determine if there is a positive or negative correlation between eating a vegetarian diet and hair loss.

A child's weight from birth to 5 years of age.

Ranking of the total health case expenditures per country.

2 Brewing Color Schemes (It's Crayon Time!) [/9]

Below are three diagrams of the desks in our classroom. Sketch a reasonable estimate of the data for its current occupants and carefully label the legend. On each diagram mark yourself with an asterisk, '*'. Also indicate whether your chosen color scheme is *sequential*, *qualitative*, or *diverging* (each type should be used exactly once).

**height,
relative to
the average**

type: _____

hair length

type: _____

**undergraduate
graduate student,
ta, other,
unknown**

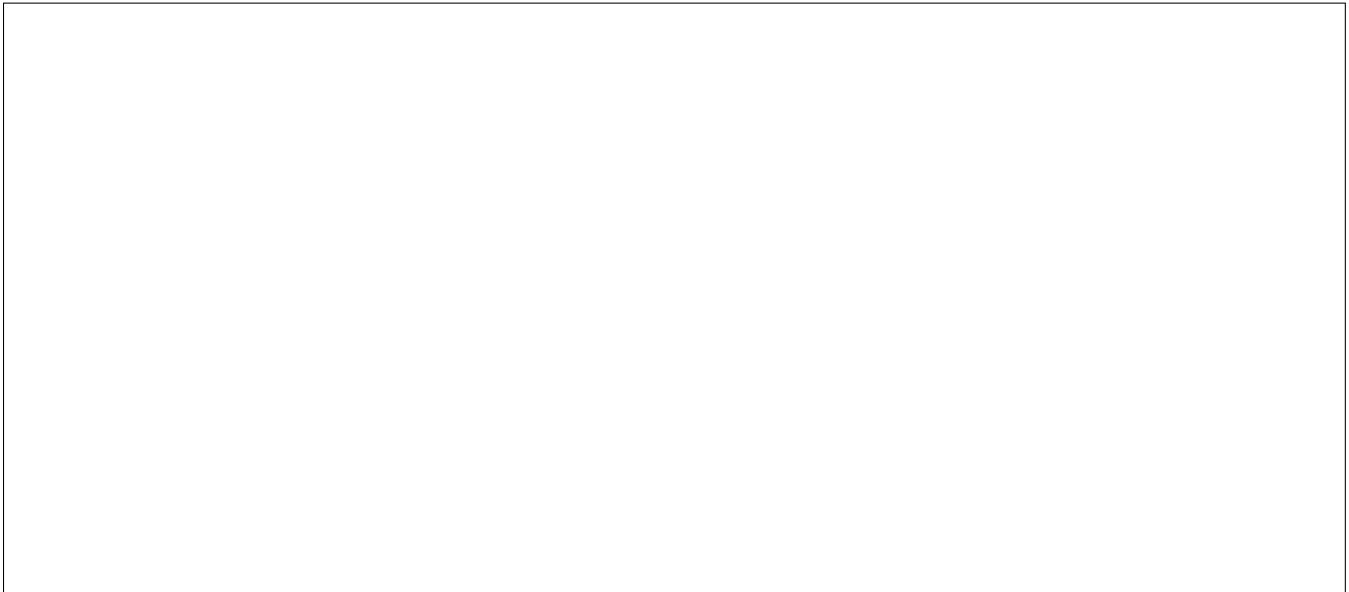
type: _____

3 Parallel Coordinates [/9]

Design a (non-interactive) parallel coordinates visualization for sports with five axes: cost for an individual to participate in the sport, popularity of sport in the U.S.A., your probable skill level if you were to participate in the sport, number of individuals needed to participate (e.g., play a match/event), and how much you like watching other people play the sport. Think carefully about conclusions you expect the viewer to draw from this data. Neatly sketch this visualization with four different sports (your choice). Clearly label the data values and axes on this visualization.



Write 3-4 sentences describing the choices you made in designing this visualization. Explain why these choices are better than other options, ensuring that the viewer accurately interprets the data.



4 Short Answer [/18]

4.1 k -Means Clustering & Convex Hulls [/5]

Plot 10 TV shows or movies on a graph with 2 axes: how much you liked the show/movie, and (approximate) release date. Neatly label the movies/shows and axes. Run k -means clustering on this data, with $k=3$. Then sketch the convex hulls for each of the 3 groups.



4.2 Node & Edge Graph Drawing Vocabulary [/4]

Write the common graph term in the box next to each definition below.

Vertices can be separated into two groups. No edge connects two vertices from the same group.

A collection of vertices such that an edge connects every pair of vertices within the collection.

Number of edges incident on the vertex.

A 2D drawing where no edges cross.

A path along edges in the graph, starting and ending with the same vertex.

4.3 A User Study about Chart Junk [/5]

We all read the paper: “Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts” by Bateman et al. that described the results of a user study. First, write 1-2 sentences *summarizing* the user study. Then, critique the user study by describing (in 1-2 sentences) a *strength* and a *weakness* of the study methodology.

4.4 Stream Graphs vs. Stacked Bar Graphs [/4]

The paper “Stacked Graphs – Geometry & Aesthetics”, by Byron and Wattenberg included feedback from New York Times readers about their novel visualization. Write 1-2 sentences describe one of the common positive reactions to the streamgraph, and 1-2 sentences describing one of the common negative reactions to the streamgraph.

5 Truthiness [/8]

Almost all of the statements that follow are false. Identify each statement as false or true, and correct each false statement so that it is true (but still informative and relevant).

5.1 GraphVis for Automatic Node-Edge Graph Layout [/2]

True or False For a complex network with cycles and some redundant connections, you should try the `dot` layout algorithm; however, if you have a hierarchical graph with directional edges, you should try `neato` to minimize spring energy. (The `circo` layout algorithm is completely useless.)

5.2 Visualizing Uncertainty [/2]

True or False We read a couple papers that compared gradient plots, violin plots, error bars, and other options for encoding data uncertainty, in which the authors concluded it is impossible to properly convey statistics and probability to the novice reader using visualizations.

5.3 Importance of Interactivity [/2]

True or False The paper “LineUp: Visual Analysis of Multi-Attribute Rankings”, by Gratzl et al., allows a skilled user to inspect and adjust the relative importance of different attributes in a complex dataset to pick the best college, most nutritious food, or *best-value-for-my-dollar* computer.

5.4 Ethics and Visualization [/2]

True or False According to the authors of “Adaptive Privacy-Preserving Visualization Using Parallel Coordinates” and “Agile Ethics for Massified Research and Visualization”, the most important ethical consideration for visualization professionals is avoiding financial conflicts of interest.