

Sample Exam: Data Visualization
Time: x:xxPM ~ x:xx PM (x hours)
(x Pages, x points)
(You can answer either in English or Chinese)

Name: _____ Student ID: _____

Visual Design (?pts)

Q1: Data attribute: (?pts)

[Check color maps on the screen] What **types** of attribute are the following? What are the better **visual channel** to encode each attribute? What are the best **color map** for each attribute if you want to use color to encode them? *Circle* your answer.

1. Name (e.g. Kevin, Ben, David)

Type:	categorical	ordinal	quantitative
Visual channel:	magnitude channel		identity channel
Color map:	(A) (B) (C)	(D) (E)	(F)
2. Survey Response (e.g. strongly disagree, agree, neutral)

Type:	categorical	ordinal	quantitative
Visual channel:	magnitude channel		identity channel
Color map:	(A) (B) (C)	(D) (E)	(F)

Q2: True/False: (?pts) Circle either T (True) or F (False)

1. T F Color saturation is a much more effective visual channel than length..
2. T F Bar chart is a better choice than line chart if the key attribute is quantitative.
3. T F If human can separate more different values/steps of a visual channel, it means the visual channel has a better pop out effect.
4. T F Reordering the rows and columns in a (matrix) heatmap could be helpful to show special pattern inside the dataset.

Q3: Single-choice question. (?pts) Write down your answer at the left hand side of each question.

1. If you have a dataset with 50 data items, each item has 4 attributes, each attribute is ordered, and you want to design a visualization on a traditional newspaper for readers to observe the correlation of any 2 attributes, which of the following visualization could be the best choice?
(A). Bar chart
(B). Parallel coordinate plot
(C). Dual-axis line chart
(D). Scatterplot Matrix
2. Which map visualization allows users to recognize a region by the region's size most easily?
(A). grid cartogram
(B). dorling cartogram
(C). contiguous cartogram
(D). choropleth map
3. If you have a network data structure and you want to find a path from node A to node B, which one is the best choice?
(A). Adjacency matrix
(B). Treemap

- (C). Node-link diagram
- (D). Parallel coordinate plot

Q4: Short-text-answer question (?pts)

1. [Check the extra color A4 paper] Among (A), (B) and (C), please map (A), (B) and (C) to hue, saturation and luminance?
2. If you have a dataset that describes car sales of each city in Taiwan, when (for what task) should you select bar chart, instead of a choropleth map, to visualize the data? Explain your reason in detail.

Q5: Design (?pts)

You work in a design agency which needs to develop a Facebook Monitoring application for the president election in Taiwan. The goal of the application is to show the sentiment of posts across candidates, time, and space. Sentiment analysis is the interpretation of the emotional meaning of text, as positive (supportive, e.g. happy) or negative (unsupportive, e.g. angry) or neutral.

Data: You are provided with data extracted by a data processing engine that monitors Facebook and produces the following information at one-day intervals:

Attribute	Attribute Type	Cardinality/Range
Time	Quantitative	C: 2023-01-01 to 2023-12-31 R: 365
Candidate	Categorical	C: 3
City/County	Categorical	C: 22
Average sentiment of posts (over interval since previous timestamp)	Quantitative	R: -1 to 1
Volume (# of posts) (over interval since previous timestamp)	Quantitative	R: 0 to 500k

The time interval is one year. There are 36.5K (365*3*22) items in a tabular dataset. There are also shapefiles for the geographic boundaries of Taiwan.

A: Design a visualization that shows all of this information on one screen.

The visualization can be interactive and will probably contain multiple views. It should allow the viewer to answer all of the following questions about each candidate:

- Q1: How has the sentiment for candidate X changed over time in each city/county?
- Q2: At time point T, compare the sentiments between both candidates in each city/county.
- Q3: In which city/county has the sentiment changed so that the most positively discussed candidate flips from one to the other candidate, between time points T1 and T2?
- Q4: What is the distribution of the total volume of posts across Taiwan at time point T? and overall time?
- Q5: Is the sentiment correlated with the volume of posts, for either candidate?
- Q6: Are there geographic patterns to the sentiment distributions for the candidates?

Do not forget to include titles, axis labels or legends as needed! Your sketch can be hand-drawn. It can also be mocked up using Powerpoint, a graphics editor, or wireframe tools.

B: Rationale for your design choices

Provide a rigorous rationale for your design decisions.

Document the visual encodings you used and why they are appropriate for the data and tasks that the visualization should support. Specifically:

- For each view, what marks and channels did you choose, and why?
- For each view, what questions are easy vs difficult to answer?

Scientific Visualization (?pts)

Q1: True/False: (?pts) Circle either T (True) or F (False)

1. T F To store a scientific dataset, the domain structure (grids and cells) defines where to sample and store data values.
2. T F Isocontour computation is an inverse problem of data value interpolation in a cell.
3. T F To detect a boundary in a scientific dataset, we can calculate the data value gradient. If the gradient is larger at a point, the location could be where a boundary is.

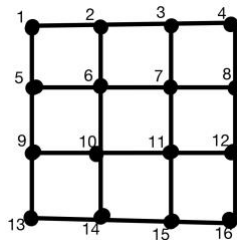
Q2: Single-choice question. (?pts) Write down your answer at the left hand side of each question.

1. When discussing direct volume rendering technique, what information is defined by a 1D transfer function?
(A). The mapping from data values to grid point locations, and from data values to color
(B). The mapping from data values to opacity, and from data values to grid point locations
(C). The mapping from data values to opacity, and from data values to color
(D). The mapping from data values to gradient, and from data values to color
2. When discussing grid structures to store scientific dataset, we have different grid structure choices, such as cartesian mesh, relinear mesh, or curvilinear mesh. What is the advantage of using a complex grid structure (e.g. curvilinear grid) ?
(A). Quickly access the data value at a grid point when visualizing the dataset.
(B). To store less data attributes for a multi-variant scientific dataset (e.g. a data with temperature, pressure...) to reduce dataset file size
(C). Have a better data representation quality when spending same amount of storage size?
(D). None of above
3. For a tetrahedral cell, how many unique isosurface topological cases does it have?
(A). 2
(B). 3
(C). 4
(D). 5



Q3: Short-text-answer question (?pts)

1. The figure is a 2D dataset with 4x4 grid points. The number next to each grid point is the data value of the grid point. Please draw the iso-contour with data value "7.5". You have to correctly draw which edge passed by the iso-contour, but you do not have to draw what location on an edge intersected with the iso-contour precisely.



2. The figure is a cell with 4 grid points. The number next to a grid point is the data value of the grid point. The (x, y) next to a grid point is the 2D coordinate of the grid point. Please calculate the data value at $(1.25, 3.5)$ by bi-linear interpolation.

