CPSC 447, Foundations 3

Updated 2025-02-11 7:45pm (updated text in Q3.1 in red).

Submit through Gradescope, as a file in PDF format, using this Google Doc answer template.

1 Multiple Views (7%)

For each of the multi-view visualizations below, what type of view coordination is used? Remember it is possible to use more than one of these choices simultaneously in an interface.

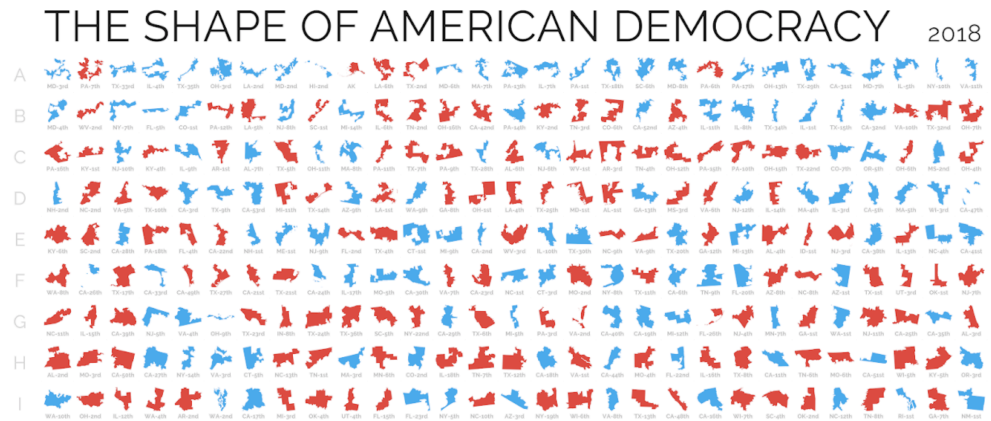
- A: Multiform
- B: Overview/Detail, Same Form
- C: Overview/Detail, Multiform
- D: Small Multiples
- E: No Linkage
- F: Fully Redundant

Justify your choice, according to two criteria:

- Is visual encoding shared between views (choose only one)-:
  - G: Same
  - H: Different
- Is data shared between views (choose only one):
  - I: Same data in both
  - J: Subset of the data in one vs another
  - K: Disjoint data partitioned between the views

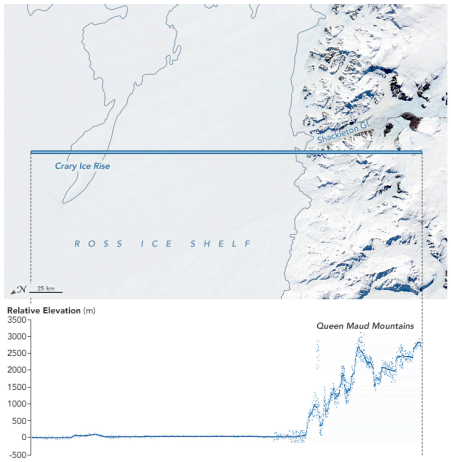
Rubric: 1% each example, 0.5% answer and .25% encoding choice and 0.25% data choice.

1.1 Gerrymandering



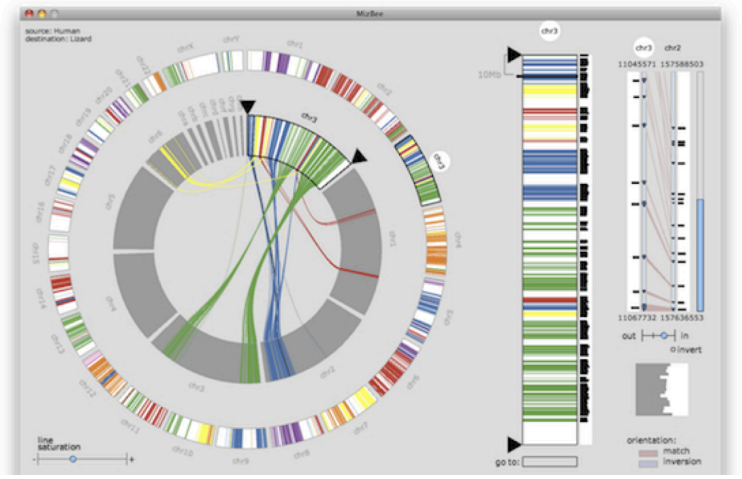
http://www.statsmapsnpix.com/2018/05/the-shape-of-american-democracy-v10.html

1.2 Taking Measure of Antarctic Terrain



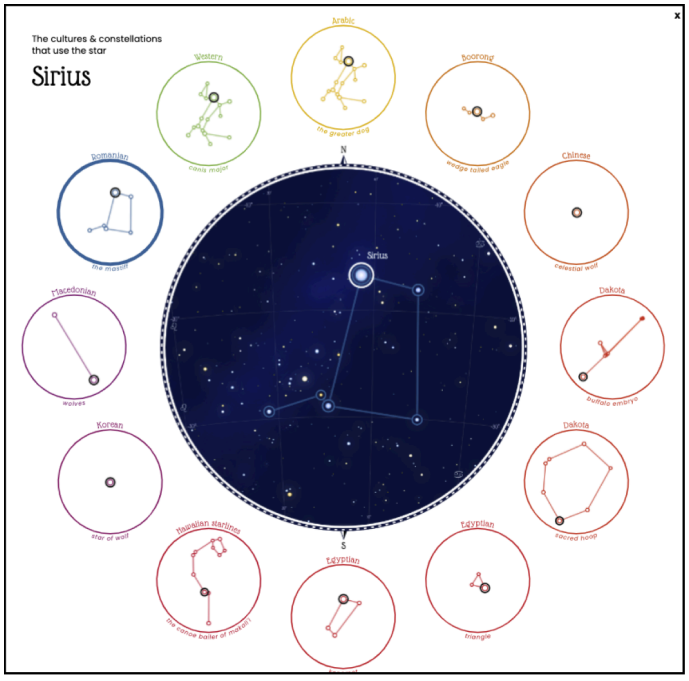
https://earthobservatory.nasa.gov/images/144367/taking-measure-of-antarctic-terrain

1.3 MizBee



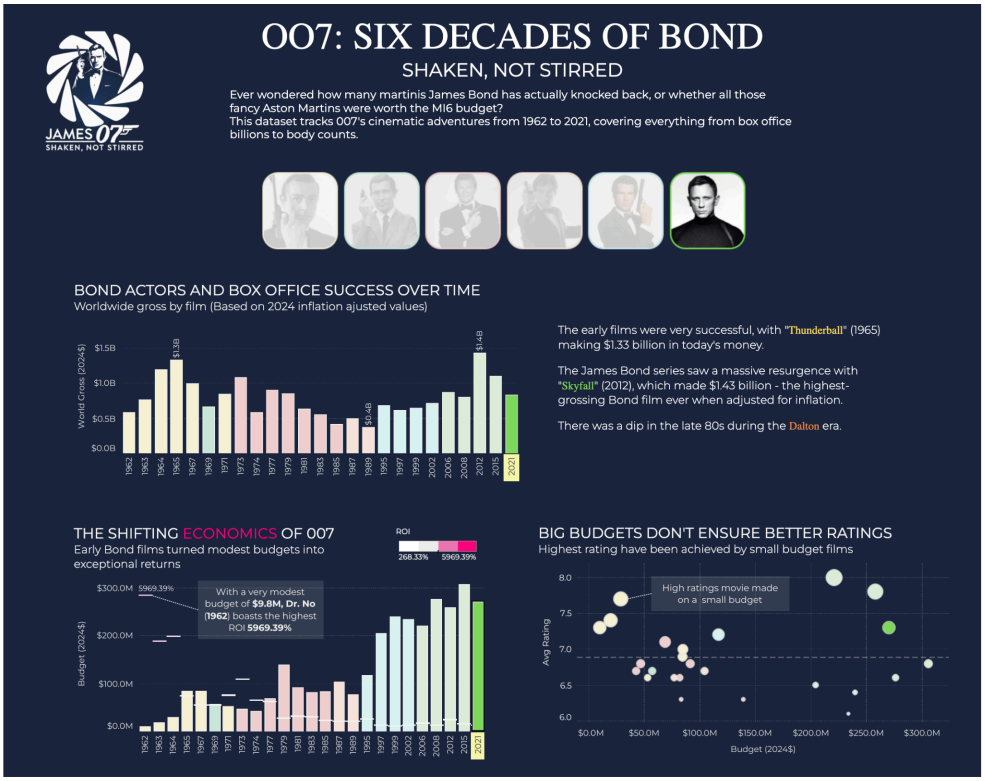
https://www.youtube.com/watch?v=86p7brwuz2g

1.4 Figures in the Sky



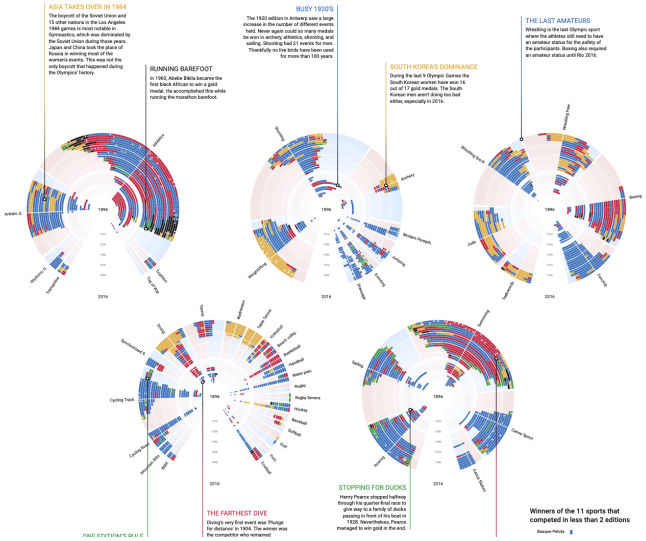
https://figuresinthesky.visualcinnamon.com/

1.5 Bond



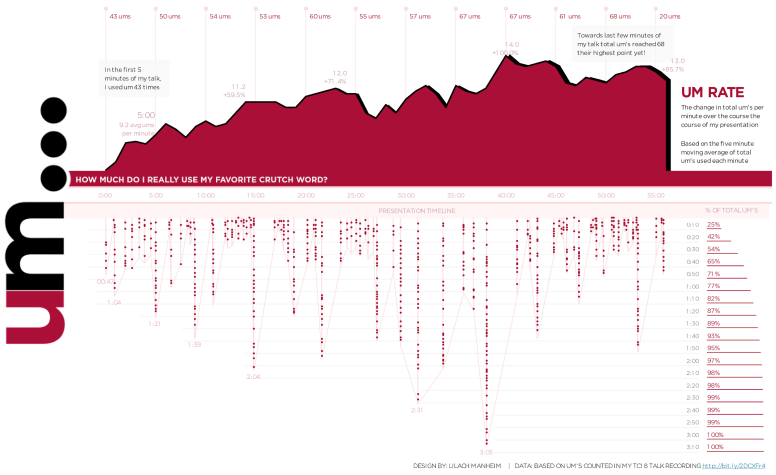
https://public.tableau.com/app/profile/fredfery/viz/007SixDecadesofBondShakenNotStirred/007SixDecadesofBondShakenNotStirred

1.6 Olympic Feathers



https://olympicfeathers.visualcinnamon.com/

1.7 How Much Do I Really Use My Favorite Crutch Word



https://public.tableau.com/profile/lilach.manheim#!/vizhome/AnUm____Analysis/Um____

2 Generate Table Alternatives (12%)

Table A (left)

	Age	Best 100 m	Furthest Jump	Sex
Amy	16	13.2	5.2	F
Basil	18	12.4	4.2	F
Clara	14	14.1	2.5	F
Desmond	22	10.01	6.3	M

Table B (right)

	BPM T1	BPM T2	BPM T3
Amy	90	130	150
Basil	70	110	109
Clara	60	140	141
Desmond	84	100	108

2.1 Sketch a suitable visual encoding for Table A. The sketch should include all of the items and attributes for that table as a static visual encoding: do not assume any interaction will take place. You may use a single view, or multiple views.

Provide rationale justifying your design choices vs alternative possibilities with respect to:

- attribute types and channel choices (in terms of expressiveness and effectiveness),
- semantics of the attributes (an important part of data abstraction), and
- information density.

2.2 Same as above, for Table B. Note that 'BPM' stands for Beats Per Minute, a measure of heart rate.

2.3 Identify (at least one) weakness in your previous choices for Table A and write it down. Create a new design that addresses it, sketch that. Provide rationale justifying your new choices.

2.4 Same as above, for Table B.

Do your best to create reasonable designs that you can justify. However, if your analysis adequately explains why a design is unsuitable, you will receive full credit for both. Sketches do not need to be polished: very rough sketches are fine as long as they are legible. The point is to show the design idea, not to make something polished. (See above, as with Q1.)

Rubric: 3% each of 4 tables: 1% sketch/design & 2% analysis.

3 Aggregation (8%)

3.1 Bugs 1

Design a suitable visualization for the following dataset. You may use multiple views. You may define the intended task yourself, but make sure it's sufficiently complex to justify the need for visualization. Your task can emphasize some attributes more than others, but it should involve all attributes to some extent. Create a simple sketch to illustrate, and describe any interactions or view coordination that you use. Justify the suitability of your choices.

- Bug Species dataset: ~~7~~ 8 attributes, 100 items
  - Life duration: quantitative, 1-10000 (days)
  - Legs: quantitative, 4-1306 (number of legs)
  - Wings: quantitative, 0-6 (number of wings)
  - Weight: quantitative, 1-3000 (milligrams)
  - Speed: quantitative, 0-35 (miles per hour)
  - Length: ordinal, 3 levels (small, medium, large)
  - Diet, categorical, 100 levels
  - Name, categorical (unique key per species)

Rubric: 3% (1% sketch, 2% design suitability / scalability)

3.2 Bugs 2

Same as above, but with a larger dataset where there are **100,000** items
Rubric: 3% (1% sketch, 2% design suitability / scalability)

3.3 Bugs 3, Idiom Comparison

Consider the dataset of Q3.1, and how it could be represented with parallel coordinates (PC) vs scatterplot matrix (SPLOM). Explain one strength/pro, and one weakness/con, for each of those two choices of idiom.
Rubric: 2% (0.5% for each pro/con)

4 Twitter Monitoring Application (63%)

Problem: You work in a design agency which needs to develop a Twitter Monitoring application for the federal election in a country. The goal of the application is to show the sentiment of tweets 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 Twitter and produces the following information at one-day intervals:

Attribute	**Attribute Type**	**Cardinality/Range**
Time	Quantitative key	R: 2020-01-01 to 2020-12-31 C: 365
Candidate	Categorical key	C: 2
State	Categorical key	C: 50

Average sentiment of tweets (over interval since previous timestamp)	Quantitative value	R: -1 to 1
Volume (# of tweets) (over interval since previous timestamp)	Quantitative value	R: 0 to 500K

The time interval is one day. There are 36.5K (365*2*50) items in a tabular dataset. There are also shapefiles for the geographic boundaries of the states.

4.1 Design a visualization that shows all of this information through multiple views *on one screen*.
The visualization can be interactive and should contain multiple views side by side. The combination of all views should effectively 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 state?
- Q2: At time point T, compare the sentiments between both candidates in each state.
- Q3: In which states 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 tweets across the states at time point T? Over all time?
- Q5: Is the sentiment correlated with the number of tweets, for either candidate?
- Q6: Are there geographic patterns to the sentiment distributions for the two candidates?

You will need to refer to views by name in the analysis below, so make sure the interface is self-documenting with readable titles for each view. Also use axis labels when appropriate. Include a legend. Your sketch can be hand-drawn. It can also be mocked up using Powerpoint, a graphics editor, or wireframe tools, such as Balsamiq.

4.2 Provide a rigorous rationale for your design decisions.
Explain the connection between views and questions, the coordination between the views, and how the visual encodings and interaction idioms you chose match up with the data and questions.

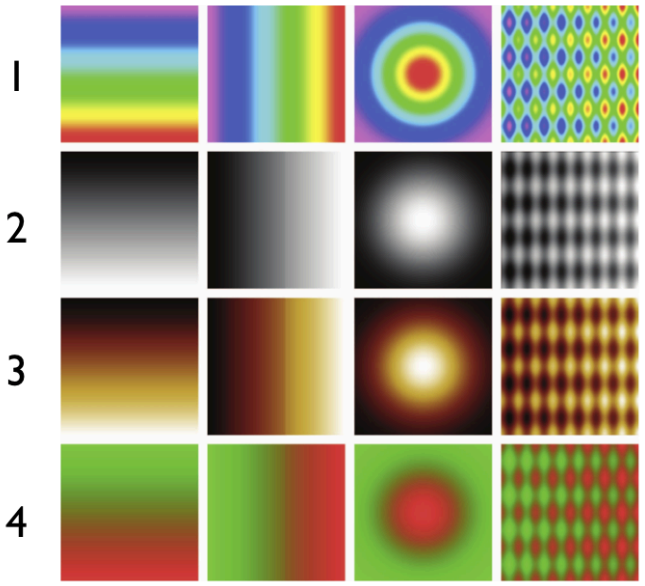
First, describe your overall strategy, with a brief summary of each view including any interaction it supports, and a brief description of the coordination between these views.

Then, for each question, state which views support it. Explain how the marks and channels used within those view(s) match up with the dataset items and attributes necessary to answer that question.

Rubric: 39% Design (6% for each of six questions, if handled in sketch. 3% labelling: title, axis labels, legends, etc). 24% Rationale (6% overall explanation including interaction and view coordination, 18% per-question analysis of marks/channels suitability)

5 Color Channels (2%)

What color channels are used for each row?



Rubric: .5% each example

6 Design Color Palettes (8%)

Use the ColorBrewer (www.colorbrewer.org) tool to help you design a color scheme for each of these four scenarios. For each question, include a screenshot of your result. Your screenshot should show the full ColorBrewer screen, so that both the swatches/legend view and the map view are visible. In writing, briefly explain your choices including diverging vs sequential vs cyclic color palettes for quantitative attributes and the number of distinguishable colors for categorical attributes.

6.1. Design a color scheme that shows as much detail as possible about patterns in consumer spending, where high values show places where people spend more than they make and low values show people living below their means. This visualization will be used in financial literacy classes for high school students.

6.2. Design a color scheme that shows accumulations of toxic waste in that region, to be used in a presentation to policymakers on how to allocate cleanup funds.

6.3. Design a color scheme that allows you to easily distinguish every color in the random section of the map (the lower left). If you have a ten-class map, you should be able to see clearly ten unique colors.

6.4. Design a color scheme to display patterns in preference changes for favorite ice cream flavor. The time series dataset has two different years of data: this year and last year. The flavor attribute has a list of fifteen possibilities for favorite flavor at that time point. Your audience is a group of marketing managers who are particularly interested in places where preference changes, in order to develop strategies for co-marketing ice cream toppings.

Rubric: 2% for each: 1% color palette, 1% rationale