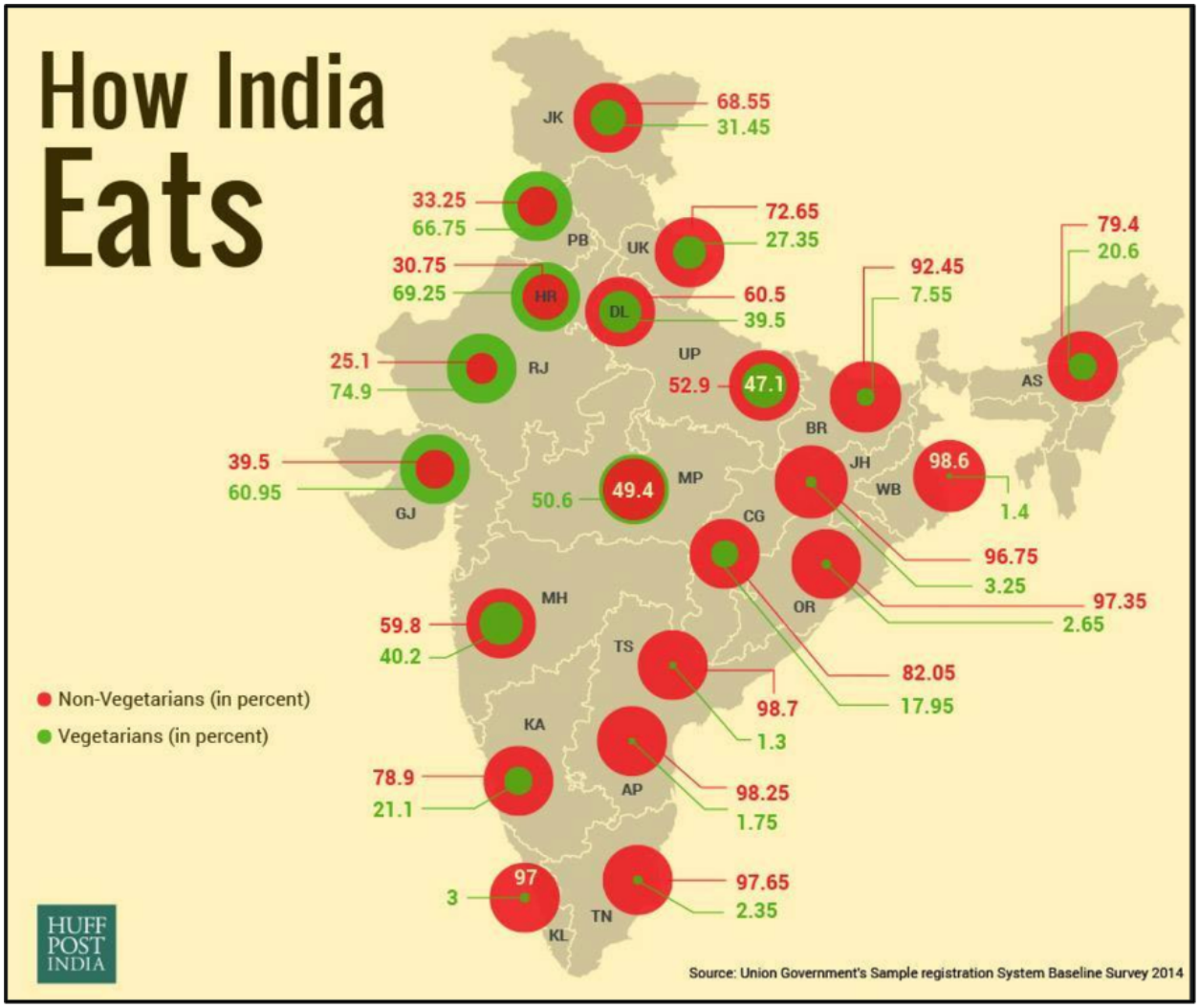
1) Design Critique



Source: http://viz.wtf/image/162284459410

Task 1.1: Critique

Write a design critique of the visualization above (‘How India Eats’). Your critique should also address the following questions:

- Who is the audience?
- What is the objective of the design? What questions does the visualization answer or try to answer?
- What elements of the design are related to that objective?

- Are those elements effective in achieving the objective? Why or why not?
- How is the visualization a success or a failure? List 3 positive and 3 negative aspects.
- Why do you like/dislike this visualization?

Maximum of 400 words

Answer for Task 1.1:

The objective of the visualization is to compare the percentage of vegetarians to the percentage of nonvegetarians in different regions of India, so it is designed for any Huffpost India audience who is curious about dietary habits in India. The main question the visualization tries to answer is, “How many people in each region of India are vegetarian?” It also addresses the following question: “How do the percentages of vegetarians and nonvegetarians differ in each region?” To answer these questions, the visualization pastes several variations of two concentric circles on a map of India to illustrate rates of vegetarians and nonvegetarians in each region. The outer circle corresponds to the larger rate in the binary comparison. In addition to size, position, and shape, the visualization uses color. Green corresponds to the regional rate of vegetarians while red represents the regional rate of nonvegetarians. The elements in the graph somewhat achieve the objective. However, this visualization is more so a failure than a success because the author’s methods sacrifice clarity and graphical integrity. First, a positive aspect of the visualization is the relevance of the base geography to the graph since it only portrays India and divides India into states and union territories. However, the author should use a basic choropleth graph instead of concentric circles to represent the binary comparison between vegetarians and nonvegetarians in each state/territory. The concentric circles are superfluous because only one of the rates (e.g. the percent of vegetarians in a state/territory) and a color gradient are needed to communicate the same information. Second, another positive aspect is that each circle is labeled by a numeric rate. However, the rates are rounded inconsistently and not always accurate. For instance, the percentage of vegetarians in Gujarat (i.e. GJ) is 60.95% while the percentage of nonvegetarians in Gujarat is 39.5%. This labeling implies that the percentage of people in Gujarat is 100.45% instead of 100%, which jeopardizes graphical integrity. Third, a positive aspect is the use of a legend and specification of the data source. However, given any two concentric circles, the size of the outer circle always seems to be the same regardless of whether it represents a rate of 52.9% or 92.45%. Also, the placement of the circles makes it seem as though some states/territories on the map have rates assigned to them while other states/territories lack rates entirely. All in all, I like the visualization because it is creative and thought-provoking. However, I dislike the visualization because it is unnecessarily complicated and sometimes portrays incomplete/incorrect data.

2) Graphical Integrity, Gestalt Principles, C.R.A.P. Principles

Task 2.1: Graphical integrity

Evaluate the visualization below (‘The shrinking family doctor’), based on Tufte’s ideas of…

- a) graphical integrity
- b) lie factor
- c) data-ink ratio
- d) chart junk

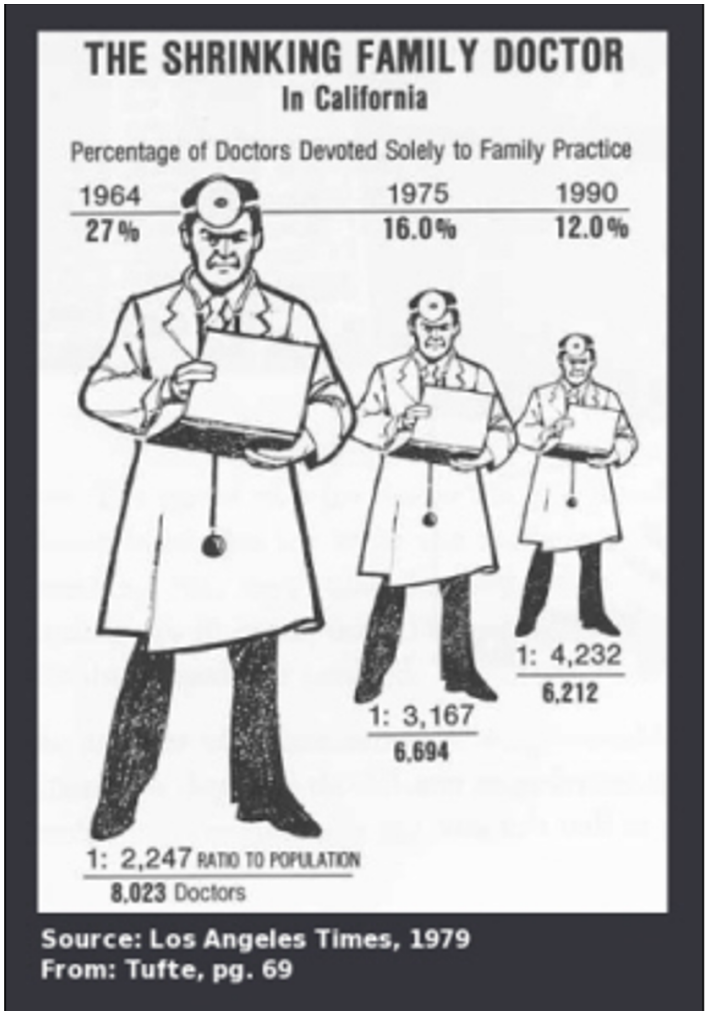
In your response, push yourself to analyze how you might improve the data-ink ratio.

Maximum of 300 words

Answer for Task 2.1:

This visualization uses changes in the size of a two-dimensional image to represent changes in the measure of a one-dimensional value. The widths and heights of the images are both proportional to the number of doctors, which makes the visual differences between the images much greater than the actual differences in the number of doctors. As such, the data is misrepresented, which damages graphical integrity. If using a series of two-dimensional images to represent one-dimensional values was necessary, then the areas of those images should be proportional to the one-dimensional values. A lie factor of 1 in a visualization is ideal because it implies that

the size of the effect shown in the graphic corresponds to the size of the effect in the data. This visualization has a lie factor that significantly exceeds 1 because it uses area to show changes in one-dimensional data. Hence, its lie factor is far from ideal. Furthermore, the data-ink ratio is very low because the total ink used to present the data is far less than the total amount of ink used in the entire visual. The fastest way to increase the data-ink ratio would be to use a line chart in place of the above visual. The x-axis could represent the years 1960-1990 while the y-axis could represent the percentage of doctors devoted to family practice. The images themselves are chartjunk because they do not convey any new, accurate information about the data and are not necessary for comprehending the information represented on the graph. Since the images of the doctor are not essential to the graph, removing them will not eliminate any of the data portrayed in the visual. Instead, deleting them will remove a lot of chart junk, thereby improving clarity by reducing redundancies.



Task 2.2: Gestalt principles

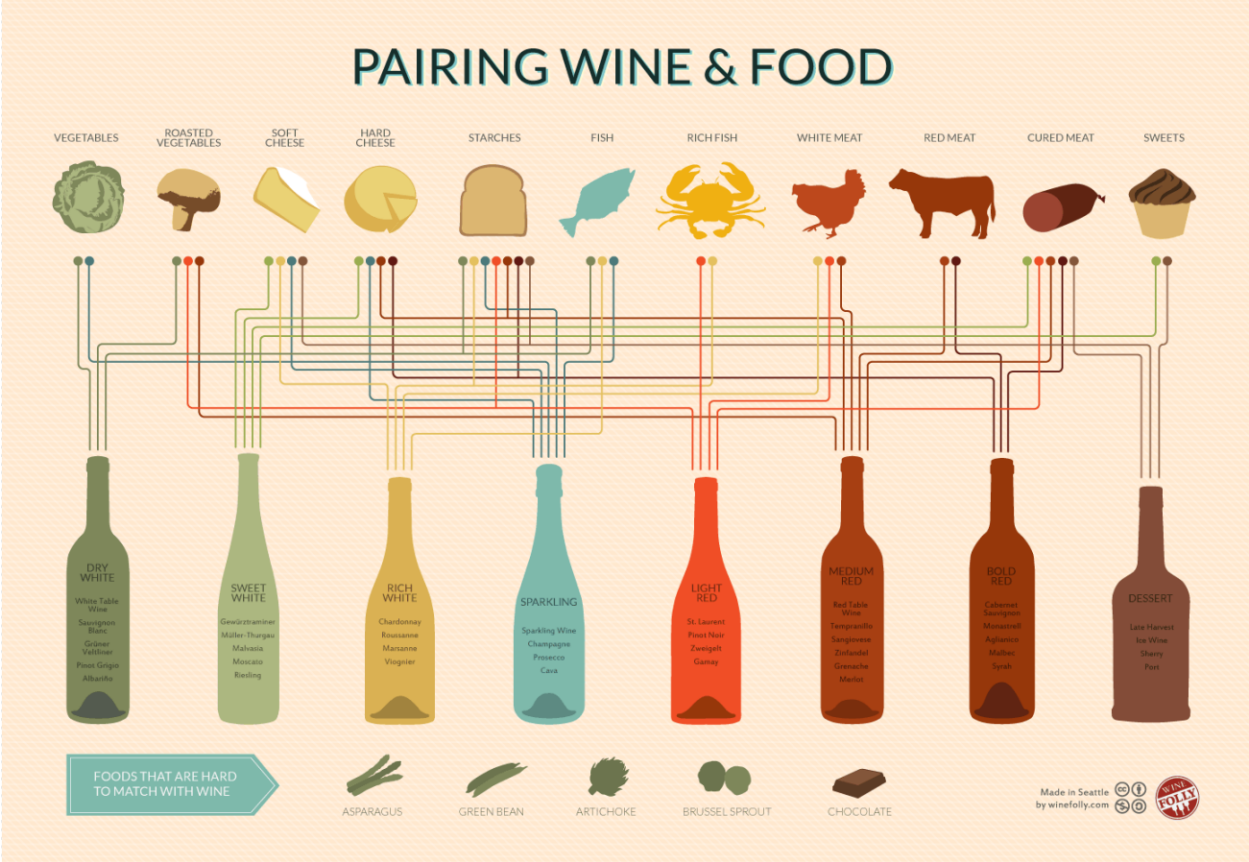
Look at the visualization below (‘Pairing wine & food’) and complete the following tasks…

- List at least 4 Gestalt principles being used in the visualization.
- Describe each principle you listed in context.

Maximum of 300 words

Answer for Task 2.2:

One Gestalt principle applied in the visual is proximity. Objects arranged close together are perceived as more related than those placed farther apart. Wines that are more similar in nature (e.g. white wines or red wines) are placed more closely together at the bottom of the visual to convey their resemblance. In addition, the close spacing between the food icons placed together at the top of the visual designates that the food types represent one set of categorical data. Another of the Gestalt principles used is similarity. When objects share similar attributes (such as color, direction, or shape), they are perceived as being part of a group. Each type of wine (e.g. dry white, medium red, etc.) is portrayed via a similar icon shape of a bottle. The shared bottle shape indicates that the wine descriptions contained in the bottles constitute a set of related and comparable data. One other Gestalt principle at work in the visual is connection. Objects that are connected are perceived as part of a group. Each wine pairing is illustrated via a line that connects a type of wine to a type of food. For instance, vegetables and dry white are connected by a line to convey that dry white wine is a good pairing for vegetables. This linkage via a line communicates a connection between two different sets of categorical data. Continuity is another Gestalt principle employed in the visual. Marks/objects that follow a path appear to be grouped together. Consider again the pairing between vegetables and dry white wine. This pairing is represented by four disconnected components. From top to bottom, these components are the title “Vegetables,” a lettuce icon, a small point attached to a line, and a green bottle titled “Dry White.” These components are not explicitly connected. However, the Gestalt principle of continuity makes the viewer more likely to see a continuous and smooth flowing path that extends from “Vegetables” to “Dry White.” Once the viewer’s eyes begin to follow something, they will continue to travel in that direction until they encounter another item or object. The viewer’s mind fills in the blanks (i.e. the disconnected parts) to create a continuous path.



Source: https://laurenadhav.wordpress.com/2014/02/15/data-visualization-critique/

To learn more about Gestalt principles, please refer to material from our class on October 12th.

Task 2.3: C.R.A.P. design principles

Look at the visualization below (‘OECD well-being’) and complete the following tasks:

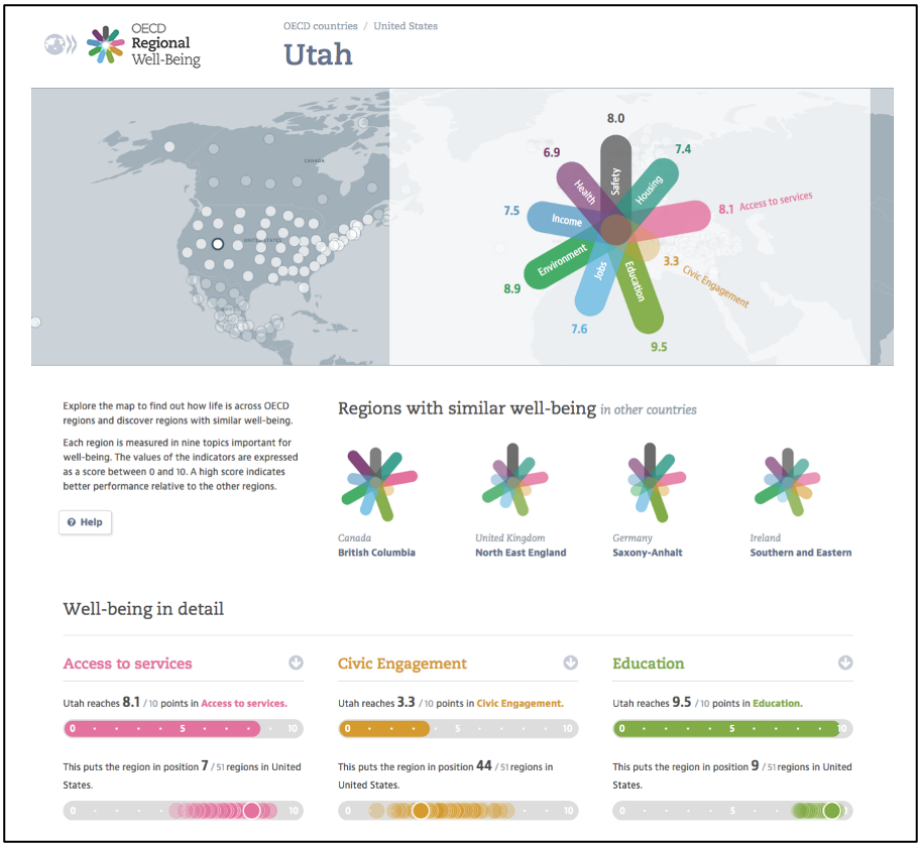
- List all of the C.R.A.P. principles
- Discuss how they are applied in this website design

Maximum of 300 words

Answer for Task 2.3:

The C.R.A.P. principles are contrast, repetition, alignment, and proximity. Due to word-count limitations, the following response will focus on a selection of the many applications of these principles in the website design. First, making elements different via contrast increases understanding. On the webpage, contrast is created through a

variety of colors: yellow represents civic engagement, light green represents education, sky blue represents jobs, etc. All other parts of the webpage are gray. The contrast between the colors and the background information makes the colored elements stand out as important components of the data. Second, repeating visual elements helps create strong unity. The enlarged asterisk-like shape in the top right corner is repeated four times in the section directly below it. Each repetition of the asterisk-like shape—where a given pill-shape color correlates to a fixed component of well-being and the length of each pill shape is proportional to the number of points that component of well-being received—indicates that the same method of analyzing/measuring well-being is being individually applied to five different regions. Third, alignment (as opposed to arbitrary placement) also helps illustrate relationships between elements. Careful alignment of the webpage's visual components separates the page into three rows, which correspond to three sections (namely, the selected region, the regions with similar well-being along with other descriptive information, and the details about well-being in the selected region). In each row, columns created by alignment further group the data into related subsections. Fourth, proximity is used in the visual where related elements are placed together. In the top row, the visual of the asterisk-like shape is placed on top of the world map where the only outlined circle corresponds to Utah. This proximity in the two images indicates that the well-being metrics displayed on the right side of that row measure well-being in Utah.



Source: http://www.oecdregionalwellbeing.org/

To learn more about C.R.A.P. design principles, please refer to material from our class on October 12th.

3) Visual Vocabulary, Visual Channels

Task 3.1: Sugar content in foods over time

Scenario: Visualize the change in the amount of sugar contained in different food groups (cereal, bread, etc) from 1950 to today.

Look at the visual vocabulary (https://bit.ly/visvocab or https://bit.ly/miro-visual-vocabulary) and pick the top two visual encodings you would choose to visualize the data of the above scenario.

For each of your two visual encodings, describe (in max 300 words):

- what questions you would answer,
- what data you (assume) you have,
- why you would choose the visual encoding, and
- how you would encode the data (i.e., which marks and channels would you use to encode which attributes of the data).

Please also include a sketch for each of your top two visual encodings that show how you map the scenario's data to the visualization.

Answer for Task 3.1:

A line graph (specifically, a multi-line graph) and a circle timeline are the top two visual encodings that I would choose to display the aforementioned data. The main question my visualization would answer is, "How does the amount of sugar contained in different food groups change over time from 1950 to today?" It would also answer the following questions: "How does the sugar content increase or decrease over time for a given food group?"; "Which food groups are being considered?"; and "How does the sugar content in each food group in 1950 compare to the sugar content in each food group in 2022?" I would assume that my data includes a list of food groups and the amount of sugar in each food group measured annually (or at some other discrete interval) from 1950 to 2022. In the multi-line graph, each line would represent a different food group, and each food group would correspond to a differently-colored line. There would be a key that assigns a color to each food group. The x-axis would represent time and the y-axis

would represent sugar content in grams. In the circle timeline, each horizontal timeline would correspond to one food group. The x-axis would represent time. The magnitude of the circles would be proportional to the sugar content in each food group at a given point in time. The title for both graphs would be “Change in the Amount of Sugar Contained in Different Food Groups from 1950 to Today.” I have included two Tableau sketches of my visual encodings below. (I know that sketches are typically hand-drawn because they are meant to be quick and easily disposed of, but I found that using Tableau as a sketching medium was most convenient for me in this instance. The data in both sketches was contrived arbitrarily, so there is no data source.)