Visualization #1

1. What is the main “story”?

The two graphs in the image show the changes in the average ADA score of House and Senate Democrats and Republicans over time. The ADA score is a measure of how liberal a politician's voting record is, with 0 being the most conservative and 100 being the most liberal. The x-axis of the graphs shows the years from 1950 to 2010, and the y-axis shows the ADA score.

The visualization's primary message is that Democrats and Republicans have become more parted over time. This is demonstrated by the fact that the two lines on the graph are going wider apart. In 1950, the average ADA score for House Democrats was about 35, while the average ADA score for House Republicans was about 20. By 2010, House Democrats' average ADA score climbed to almost 70, while House Republicans' average ADA score had dropped to about 15.

The same pattern may be found in the Senate. In 1950, the average ADA score of Senate Democrats was about 40, whereas the average ADA score of Senate Republicans was about 25. By 2010, Senate Democrats' average ADA score climbed to almost 75, while Senate Republicans' average ADA score had dropped to about 20.

2. What task(s) does the visualization enable?

The Visualization enables us to analyze and compare the trends and ADA scores of House and Senate’s.

Analyzing trends over time: The visualization shows that the average ADA score of both the House and the Senate has dropped over time.
Comparing the House and Senate's ADA scores: The visualization shows that the House's average ADA score is generally greater than the Senate's. This suggests that the House is typically more liberal than the Senate.

3. What data is represented in this visualization? How is each piece of data visually encoded?

The visualization shows two line graphs, one for the House of Representatives and one for the Senate, which track the ADA score from 1950 to 2010.

- **X-axis:** Represents the year, ranging from 1950 to 2010, The further to the right a point is, the later the year.

- **Y-axis:** Represents the ADA score, ranging from 0 to 100.The higher the point, the higher the average ADA score. A higher score indicates more liberal positions.
- **Colored bands:** Represent the spread of scores within each party. A narrower band indicates a tighter clustering of scores, while a wider band indicates more diverse scores within the party.
- **Color of the bands:** Blue represents Democrats, red represents Republicans.

4. How would you evaluate data-ink and non-data-ink in this visualization?

Data-ink

- **What works:**
    - The lines representing the average ADA score for each party are clearly visible and easy to follow.
    - The colored bands effectively show the spread of scores within each party, allowing viewers to compare the diversity of opinions within each group.
    - The gridlines help viewers estimate values on the axes.
- **What doesn't work:**
    - The data points for each individual legislator could be added to show more granular detail, but this might clutter the visualization if not implemented carefully.
    - Shading or transparency could be used within the bands to show the distribution of scores, not just the average.

Non-data-ink

- **What works:**
    - The labels for the axes and the title are clear and concise.
    - The color coding (blue for Democrats, red for Republicans) is a common convention that most viewers will understand.
- **What doesn't work:**
    - The grey gridlines could be lighter or thinner to reduce their visual dominance, especially in areas with dense data points.

- The source of the data and the specific bills used in the ADA scoring could be mentioned in a caption or footnote to provide context.

5. How do the color choices address the above?

- The lines representing the average ADA score for each party (Democrat and Republican) are clearly visible in dark blue and red, following the data-ink principle.
- The spread of scores within each party is shown by colored bands in lighter shades of blue and red.
- The gridlines, x and y-axis labels, and title use grey tones, following the non-data-ink principle and minimizing their visual dominance.
- This color choice aligns with the "main story" of the visualization, which is to depict the increasing partisan divide between the two parties. The stark contrast between blue and red emphasizes the separation and reinforces the narrative of polarization.

Possible Improvements:

- While the color coding is effective in terms of highlighting the partisan divide, it might be worth considering a color palette that is also accessible to viewers with color vision deficiencies. This could involve using high-contrast alternatives or additional visual cues to differentiate the lines and bands.

6. What design principles to enhance cognition are at work in this visualization? Be specific

Principles of Graphical Integrity:

- **Clarity and Conciseness:** The visualization focuses on two clear variables: average ADA score and score spread within each party. Unnecessary elements like decorative details are minimized, reducing clutter and aiding comprehension.
- **Accuracy and Completeness:** The graph accurately represents the data within the chosen timeframe. While it focuses on averages and spreads, providing access to individual data points could further enhance completeness.
- **Consistent Encoding:** Colors (blue for Democrats, red for Republicans) and line thickness (potentially representing the number of legislators contributing to the average) are used consistently throughout the visualization, promoting understanding and reducing cognitive load.

7. What might you do differently to better support understanding and exploring the data?

- **Consider color accessibility:** While the blue and red color coding is effective for many viewers, it's important to consider those with color vision deficiencies. You could use a colorblind-friendly palette or add additional visual cues to differentiate the lines and bands.
- **Highlight key trends:** The visualization currently shows the average score and the spread of scores for each party. It might be helpful to add a line representing the overall average ADA score across both parties, making it easier to see the overall trend of polarization.

Visualization #2:

1. **What is the main "story"?**
    - The main story of the visualization is that there is a big difference between the amount of money donated to different diseases and the number of deaths they cause.
    - Diseases with more public awareness or personal connections tend to get more funding. This could explain why cancer charities receive more money than others and the heart disease is the top killer but receives less in donations than other diseases.

2. What task(s) does the visualization enable?

- Comparing the amount of money raised by different charities to the number of deaths caused by the corresponding diseases.
- Identifying which diseases are underfunded or overfunded.
- Raising awareness of the issue of disease burden and philanthropy.

3. What data is represented in this visualization? How is each piece of data visually encoded?

- In the visualization the diseases are given in individual color coding's and the charity names below the diseases in the color coding legend itself and later the visualization is divided into 2 parts left side with money raised and right side with deaths in circles of size with according to the numerical data respectively and the numerical data is given in the area beside the color circle.

- **Colors Encoded:** Dark blue: Heart disease, Beige: suicide, Pink: Breast Cancer, Green: Diabetes, maroon: HIV/AIDS, Sea-green: Prostate Cancer, Purple: Motor Neuron Disease, Blue: Chronic Obstructive Pulmonary Disease.

4. How would you evaluate data-ink and non-data-ink in this visualization?

Data-ink:

- **Effective:**
    - The areas of the colored squares are the primary focus of the visualization, effectively encoding both the amount of money raised and deaths caused. This allows for direct comparison between these two key pieces of data.The use of color helps to differentiate between diseases and makes the chart visually appealing.
- **Ineffective:**
    - The dollar amounts and death tolls are written in small text, making them difficult to read, especially for viewers with visual impairments.

Non-data-ink:

- **Effective:**
    - The overall design is clean and uncluttered, with minimal gridlines and borders, which helps to focus attention on the data. The use of white space around the squares helps to separate them visually and improve readability. The inclusion of disease names and a clear title helps viewers understand the context of the data.
- **Ineffective:**
    - The legend using different shades of the same color to represent different numbers of charities is not very effective, as it's difficult to distinguish between the minimal change of color. A legend with distinct colors or symbols would be more helpful.

5. How do the color choices address the above?

- **color and Disease Association:** The colors don't appear to have a strong association with the specific diseases they represent. Using a color scheme that evokes specific disease categories (e.g., reds for heart disease, blues for water-borne diseases) could enhance the visualization's storytelling ability.

- **Limited Emotional Impact:** The color palette leans heavily on calming and neutral tones, which might not effectively convey the urgency or emotional weight of the topic. Consider incorporating bolder colors or accents to draw attention to specific data points or emphasize the human impact of the diseases. This dual use can be confusing, as viewers might misinterpret color variations as representing different things. Ideally, separate visual cues should be used for distinct data points.

6. What design principles to enhance cognition are at work in this visualization? Be specific.

Principles of Graphical Integrity:

Clarity and Conciseness:

- Encoding two data points (money raised, deaths) with the same visual element creates confusion.
- Small text size for data labels reduces readability.

Accuracy and Completeness:

- Data seems accurate based on the source (2011 data might not reflect current state).
- Title and labels convey intended message.
- Encoding two data points in one element potentially misrepresents information.

Consistent Encoding:

- Same visual element encodes both money raised and deaths, leading to misinterpretation.
- Color intensity within the same color for squares represents different data points, adding confusion.

7. What might you do differently to better support understanding and exploring the data?

- Instead of using the same circles to represent both money raised and deaths caused, This could be separate bars, different shapes or additional elements within
- Increase the font size of data labels (money raised and deaths) for better readability. Ensuring all labels are clear, consistent, and well-aligned as the font size is reduced for smaller circles.

- Using a color palette that effectively differentiates diseases and avoids using the same shade color to encode different data points. Considering using colors that have some association with the disease categories they represent.
- Redesign the legend to use distinct colors or symbols that clearly represent the different numbers of charities associated with each disease.

Visualization #3:

1. **What is the main "story"?**

The graph shows the comparisons between elo ratings of two teams and it is an interactive graph and the comparison between Denver Broncos and San Francisco is

Denver Broncos: The Broncos' Elo rating has shown a generally upward trend since 1960, starting from around 1000 and reaching 1750 by 2020. This indicates a consistent improvement in their Elo rating over time.
San Francisco 49ers: The 49ers' Elo rating has fluctuated more than the Broncos'. They started around 1200 in 1960, dipped to around 1000 in the 1970s and 1980s, but then increased to around 1550 by 2020. This suggests periods of both improvement and decline.

2. What task(s) does the visualization enable?

- Enables the Compare the Denver Broncos with any other NFL Teams or any comparisons between two teams and status of single teams without comparisons
- Tracking and comparing the Elo rating of the Denver Broncos over time with others.
- Explore specific data points, By hovering over points on the graph, users can see the exact Elo rating for each year. This interactivity enables a closer look at specific points in time.
- Identify periods of improvement or decline: By analyzing the slopes of the lines and their relative positions, users can gain insights into when each team performed better or worse.

3. What data is represented in this visualization? How is each piece of data visually encoded?

- The Data is represented in an Interactive Line Graph
- Time is represented on the x-axis as years, ranging from 1960 to 2020.
- Elo Rating:

Denver Broncos: Shown by the orange line.
San Francisco 49ers: Shown by the beige line.

- Specific Elo rating values: Not directly displayed on the graph, but users can likely see them by hovering over data points.
- Super bowl wins for Denver Broncos are Represented by orange circles with the number of wins displayed inside each circle.

4. How would you evaluate data-ink and non-data-ink in this visualization?

Data-ink:

- Clear lines: The orange and beige lines representing the Elo ratings for each team are easily distinguishable and trackable throughout the graph.
- Super Bowl wins: The circles clearly indicate the years when each team won the Super Bowl, adding relevant information without overwhelming the visual focus.

Ineffective:

- No values on lines: The absence of numerical values directly on the lines makes it cumbersome to compare specific Elo ratings at different points in time. Hovering above is not always easy.
- Limited data display: The visualization only shows two teams, which limits its scope for broader comparisons across the league.

Non-data-ink:

- Simple background: The white background avoids distracting clutter and keeps the focus on the data.

Ineffective:

- Gridlines and axes: The gridlines and axes are slightly more prominent than necessary, potentially drawing attention away from the data lines themselves.
- Missing title and labels: The absence of a title and clear labels for the axes and lines makes it difficult to fully understand the visualization without additional context.

5. How do the color choices address the above?

- Basic Differentiation: The different colors (orange and beige) help distinguish the Denver Broncos and San Francisco 49ers' Elo rating lines, improving visual separation.

- Accessibility Issues: The color scheme is not ideal for people with color blindness, potentially hindering their ability to interpret the data accurately.
- Limited Symbolic Meaning: The colors don't seem to have a strong connection to the "story" or task of the visualization. They simply differentiate the teams without reinforcing the trends or data.

6. What design principles to enhance cognition are at work in this visualization? Be specific.

Principles of Graphical Integrity:

Clarity and Conciseness:

- The visualization focuses on the main data points: Elo rating and Super Bowl wins for each team.
- Lack of data values directly on the lines makes it difficult to compare specific Elo ratings at different points.
- Only two teams are shown, limiting broader comparisons.
- Missing title and axis labels reduce understanding without additional context.

Accuracy and Completeness:

- The data seems to be presented accurately, but the source is unknown for verification.
- Completeness is questionable as it only shows two teams and lacks information like average league Elo rating.

Consistent Encoding:

- Colors (orange and beige) consistently represent the respective teams throughout the visualization.
- Super Bowl wins are consistently encoded as circles above the corresponding team's line.
- The color scheme is not ideal for accessibility, potentially misleading viewers with color blindness.

7. What might you do differently to better support understanding and exploring the data?

- Include data values directly on the lines: This would eliminate the need to guesstimate Elo ratings from the line positions and make it easier to compare values across different points in time.

- adding more teams: While focusing on two teams can be helpful for a specific comparison, including more teams from the league would provide a broader context and allow viewers to compare the Broncos' and 49ers' performance against a wider range of competitors.
- Adding a clear title and labels: A concise title summarizing the main finding and clear labels for axes and lines would provide essential context and eliminate the need for external explanations.
- Considering adding filtering options: Allowing viewers to filter by specific eras, playoff appearances, or other criteria could enable more focused exploration and discovery of interesting trends.

Visualization #4:

1. **What is the main "story"?**

The visualization is based on data from the Tobacco Atlas, which is a project of the Institute for Health Metrics and Evaluation (IHME) at the University of Washington.

- The chart shows that the percentage of people who smoke daily has increased from 36% in 1980 to 59% in 2015.
- The map shows that smoking prevalence is highest in Central Europe, Eastern Europe, and Central Asia, and lowest in Sub-Saharan Africa.

2. What task(s) does the visualization enable?

- Explore global smoking patterns: The visualization allows users to see how the prevalence of daily smoking varies from country to country.
- Compare smoking prevalence over time: The visualization allows users to see how smoking prevalence has changed over time in different regions of the world.
- Identify areas with high smoking prevalence: The visualization can help to identify areas with high smoking prevalence.

3. What data is represented in this visualization? How is each piece of data visually encoded?

- Smoking prevalence: This is the main piece of data that is being visualized. It is represented by the pie chart, where the different slices of the pie represent different levels of smoking prevalence. The slices are color-coded, with darker colors representing higher levels of prevalence. The specific colors used are not specified in the legend, but they

appear to range from yellow (for the lowest prevalence) to red (for the highest prevalence).

- Super-regions: The pie chart is divided into six slices, which represent different super-regions of the world. These super-regions are labeled in the legend.
- Year: The data for smoking prevalence is shown for five different years: 1980, 1990, 2000, 2010, and 2015. These years are listed along the bottom of the pie chart.
- Age: The data for smoking prevalence is shown for all ages. This is indicated by the label "All ages" at the bottom of the pie chart.
- Sex: The data for smoking prevalence is shown for both sexes (males and females). This is indicated by the label "Both sexes" at the bottom of the pie chart.

4. How would you evaluate data-ink and non-data-ink in this visualization?

Data-ink:

- The pie chart itself is the primary data-ink element, representing the smoking prevalence data. It effectively encodes the proportions of smokers in different super-regions.
- The color gradient used in the pie chart is also data-ink, as it helps to visually represent the different levels of smoking prevalence. However, the lack of a legend makes it difficult to interpret the specific values represented by each color.
- The labels for the super-regions and years are also data-ink, as they are necessary for understanding the data.

Non-data-ink:

- The background grid lines and chart border are non-data-ink elements. While they may provide some visual structure.
- The title and source information are also non-data-ink, but they are necessary for understanding the context of the visualization. However, they could be made smaller or less visually prominent to reduce their impact on the data-ink ratio.
- The legend for some parts of the data is missing, which is a significant issue. It is essential for interpreting the color coding used in the pie chart and thus understanding the data it represents. Without a legend, a large portion of the data-ink becomes meaningless.

5. How do the color choices address the above?

- The Legend significantly enhances data comprehension by clarifying the specific values associated with each color.
- The visualization Effectively utilizes a sequential gradient (light yellow to dark red) to represent increasing smoking prevalence.
- In the non-data ink elements, using even more neutral tones (e.g., light gray) for background elements like grid lines could further improve data-ink focus.

6. What design principles to enhance cognition are at work in this visualization? Be specific.

Principles of Graphical Integrity:

Clarity and Conciseness:

- Data is divided into six super-regions, reducing cognitive load.
- Super-regions are ordered geographically, aiding mental connections.
- Users can highlight specific locations for focused exploration.
- Hovering over regions reveals detailed data, enhancing user engagement.

Accuracy and Completeness:

- The Color gradient Represents increasing smoking prevalence, but lacks specific values for precise comparisons.

Consistent Encoding:

- The Color gradient Uses a sequential scheme (yellow to red) aligning with natural associations.
- Clear labels for super-regions and years ensure consistent understanding.

7. What might you do differently to better support understanding and exploring the data?

- While the legend mitigates the need for extreme contrast, consider making gridlines even thinner or removing them entirely to emphasize data-ink elements.
- Although the legend clarifies color meaning, consider slightly adjusting the gradient for more distinct steps between shades, especially for accessibility with visual impairments.

- Allow users to filter the data by year to explore changes in smoking prevalence over time for each region.
- Enable users to select and compare the smoking prevalence of multiple regions side-by-side.

Lastly, give your opinion on which visualization's message came through most clearly, and why?

According to me The Visualization #2 has given the information most clearly as

- The message is clear and concise: The title and labels make it clear what the visualization is about, and the use of color helps to highlight the contrast between the two variables.
- The data is well-represented: The use of bubble size to represent the number of deaths and bubble color to represent the amount of money raised is an effective way to show the relationship between the two variables.
- The design is visually appealing: The use of color and white space makes the visualization easy to look at and understand.
- It avoids common visualization mistakes: The visualization does not use any misleading elements, such as 3D effects or distorted scales.