Obesity Visualization Documentation

Design Process:

- **Data Source**: The dataset used for this visualization is sourced from Kaggle. It comprises information on estimating obesity levels among individuals in Mexico, Peru, and Colombia, focusing on dietary patterns and physical health. With 17 attributes and 2111 records, it offers an overall view of obesity-related factors.
- **Understanding the Data:** An in-depth analysis of the dataset was conducted to comprehend variables and distributions, extracting key insights. Due to constraints, a subset of four variables was selected for visualization.
- **Iterative Prototyping:** D3.js was employed to prototype various visualizations, experimenting with chart types, colors, and interactivity. These prototypes were strategically positioned on the page for optimal impact and engagement.
- **Finalizing the Design**: After several iterations, a design was refined to effectively convey insights, encourage data exploration, and clearly communicate information. Charts include titles, named axes, and clear markers.

Questions Explored:

1. **Does obesity run in their family?**
   To explore the prevalence of a family history of obesity, indicating potential genetic influences.

2. **What is the age distribution of the dataset?**
   To explore the demographic composition and identify age-related obesity patterns.

3. **What is the obesity level per gender?**
   To analyze obesity prevalence across genders to discern contributing factors in further evaluation.
   Obesity levels:
   - **Normal Weight:** Individuals with a body mass index (BMI) within the range considered normal for their height and weight.
   - **Overweight Level I:** Individuals with a BMI that falls within the overweight range, typically defined as a BMI between 25 and 29.9.
   - **Overweight Level II:** Individuals with a higher BMI than Overweight Level I, indicating a more severe degree of overweight, often associated with increased health risks.
   - **Obesity Type I:** Individuals with a BMI that falls within the obese range, typically defined as a BMI between 30 and 34.9. This is the first level of obesity and carries increased health risks compared to being overweight.
   - **Insufficient Weight:** Individuals with a BMI below the normal weight range, indicating that they may be underweight, which can also have negative health consequences.

- **Obesity Type II:** Individuals with a BMI that falls within the obese range, typically defined as a BMI between 35 and 39.9. This is a more severe level of obesity than Type I and is associated with even greater health risks.
- **Obesity Type III (Morbid Obesity):** Individuals with a BMI of 40 or higher. This is the most severe level of obesity and is associated with significant health risks and complications, including cardiovascular disease, diabetes, and reduced life expectancy.

4. **Do you eat high-caloric food frequently?**
   To examine the frequency of high-caloric food consumption to understand its potential impact on dietary patterns and health outcomes.

Rationale of Design Choices:

- **Color Encoding:** Utilized color to differentiate categorical variables, facilitating comparison while considering color blindness to ensure universal accessibility and comprehension of the data.
- **Chart Types**:
  - Pie chart for family history
  - Histogram chart for age distribution
  - Bar chart for obesity levels per gender
  - Bar chart for high-caloric food frequency
- **Layout**: Vertical arrangement to highlight focus on each graph.
- **Interactivity:** Tooltips for enhanced engagement and understanding.

Conclusion:

The data visualization has provided invaluable insights into the obesity dataset, shedding light on significant patterns. By effectively representing key variables such as family history, age distribution, and gender disparities in obesity prevalence, alongside the consumption of high-caloric food, the visualizations serve as a tool for data exploration and understanding.

The identification of a substantial portion of individuals with a family history of obesity suggests a potential genetic component influencing obesity risk. Moreover, the balanced age distribution with a slight skew towards younger age groups underscores the importance of considering age-related factors in obesity analysis and intervention strategies. The observed differences in obesity prevalence between genders highlight the need for targeted interventions and further investigation into the factors contributing to these disparities. Plus, the identification of a significant proportion of individuals regularly consuming high-caloric food underscores the prevalence of dietary habits potentially contributing to obesity trends within the studied population.

In summary, the data visualization offers valuable insights into obesity trends, revealing the interaction of various factors such as family history, age distribution, gender disparities, and

dietary habits. These findings underscore the complexity of obesity and emphasize the importance of comprehensive approaches to effectively address this public health concern. Moving forward, targeted interventions and further research are essential to understand and mitigate the contributing factors to obesity prevalence.