FINAL PROJECT – DATA VISUALIZATION

Course - CS7250

Submitted By – Shriyansh Singh

Heart Health Analysis - *A tale of 400K adults surveyed in the US in 2022*

Heart Attacks Reported

13,435

In 2022 CDC Conducted a health survey across United States with 400k+ adults participating. Encapsulating the key highlights on heart health which surpassed fatalities from Covid-19 and Cancer.

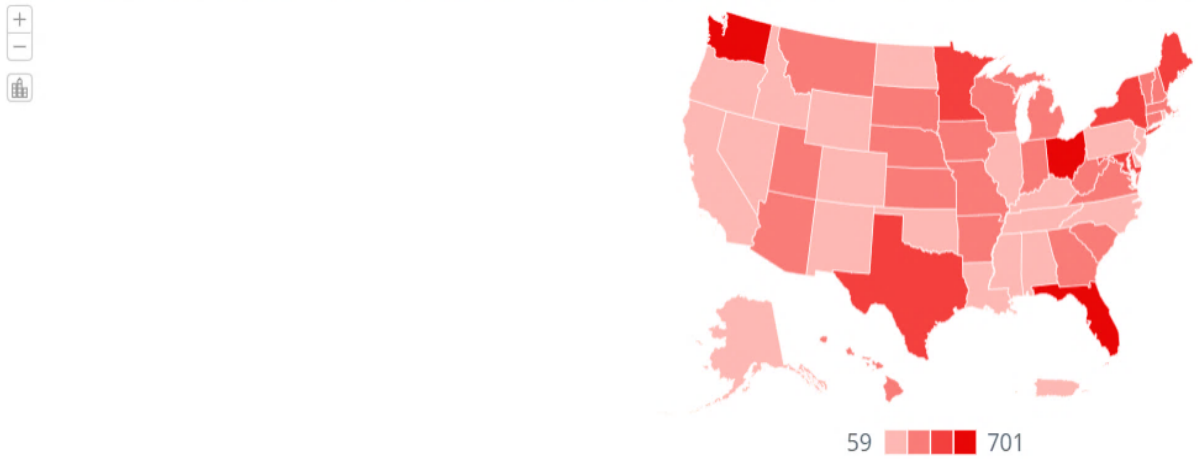
Adults Drinking Alcohol
40%

Adults with Smoking
18%

Adults with Diabetes
35%

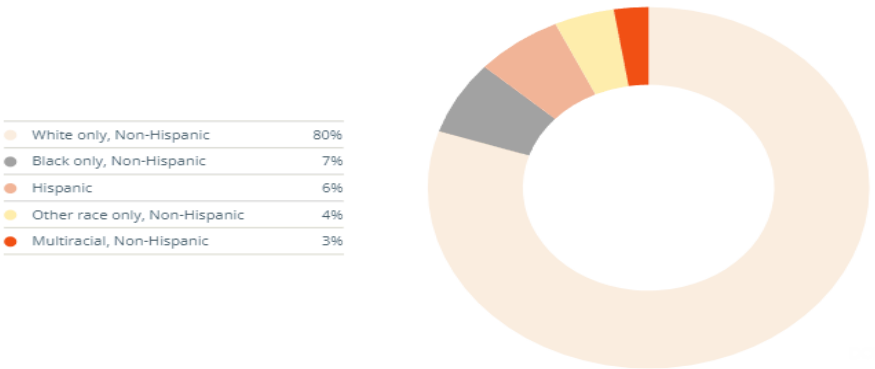
Heart Attack Cases Reported Across USA

The chart displays the distribution of heart attack cases across USA in 2022, as reported in the survey conducted among 400K + individuals

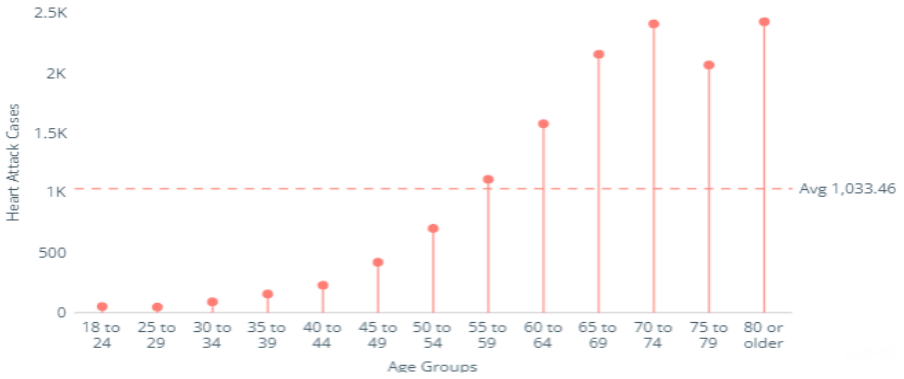


Heart Attacks by Groups - A deeper dive investigating this cardiovascular disorder

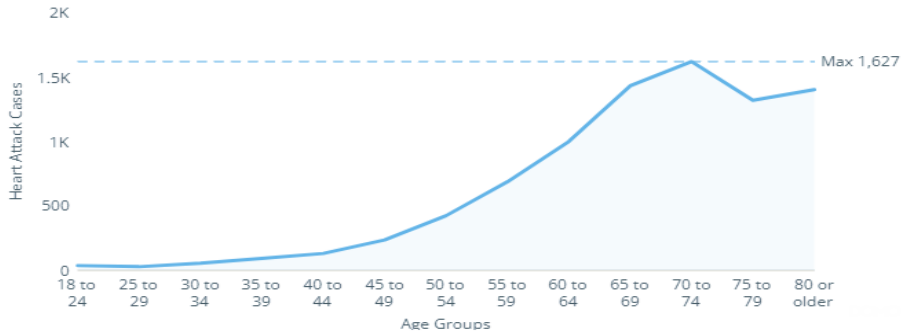
Distribution of Heart Attack - By Race



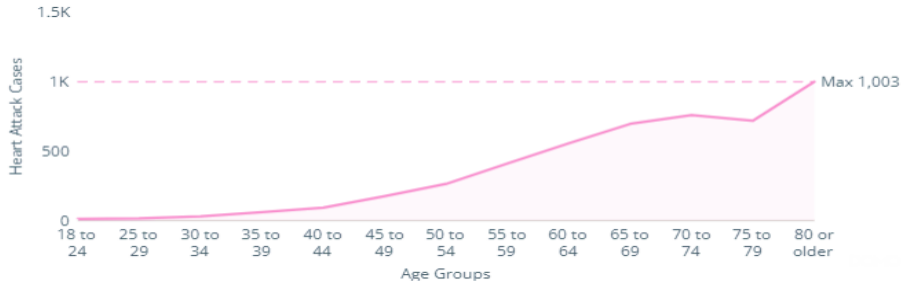
Heart Attacks by Age Groups



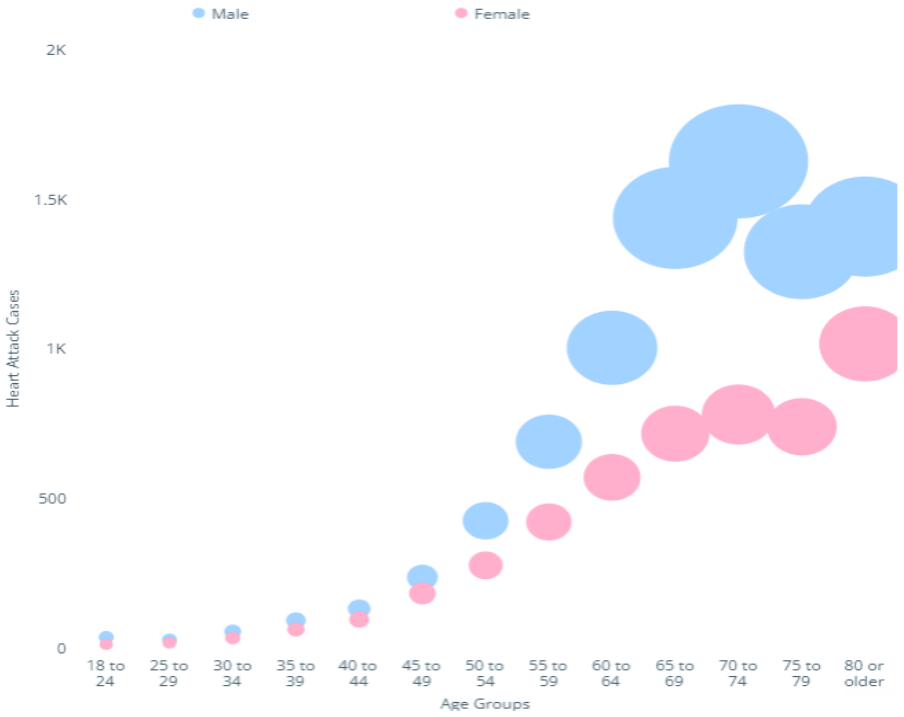
Heart Attack in Males



Heart Attacks in Females



Heart Attack Cases in Males and Females

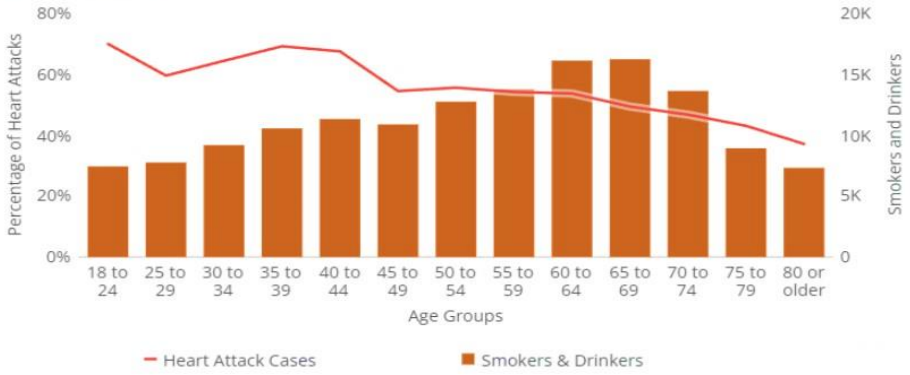


Three Key C's of Cardiovascular Disorder - Causes , Comorbidities and Cautions

Causes - *Highlighting metrics proven to be a major cause of heart attacks*

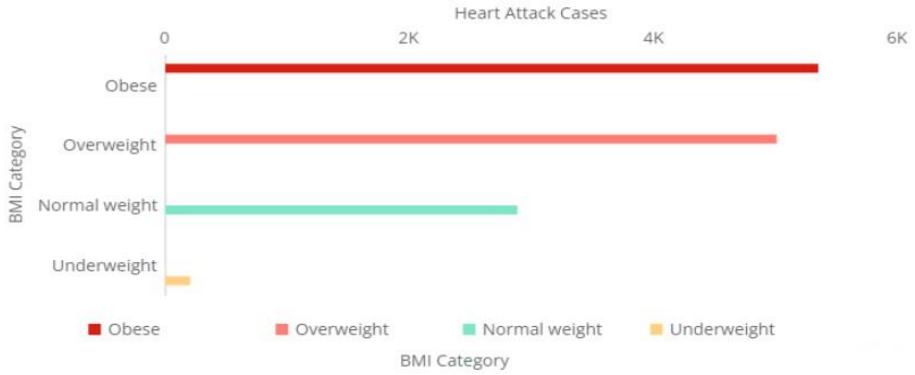
Adults who Smokes and Drinks Alcohol

The chart displays the heart attack cases in adults who reported smoking cigarette / e-cigarette and/or consumed alcohol.



Adults with Obesity

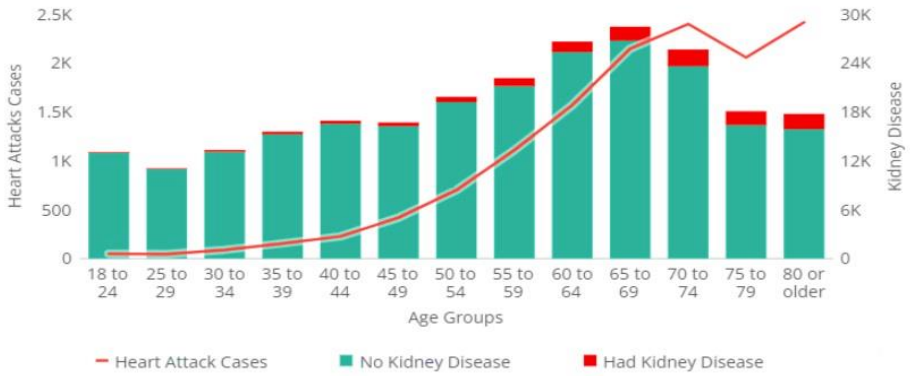
The chart displays the heart attack cases in adults belonging to different BMI categories - underweight, normal, overweight and obese



Comorbidities - *Highlighting metrics proven to be a major cause of heart attacks*

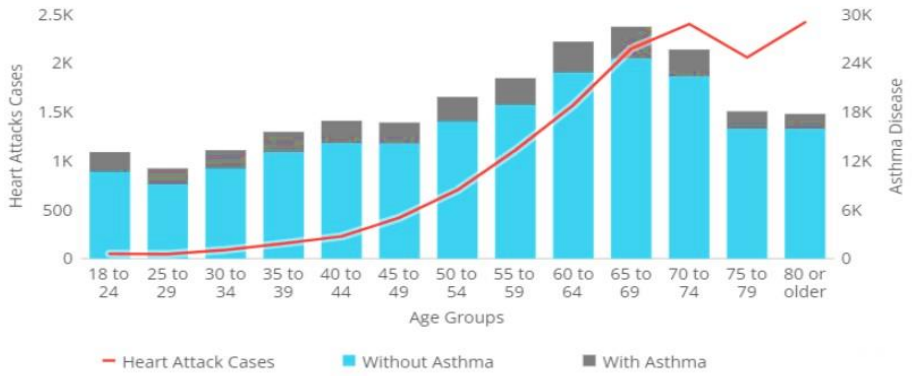
Adults with Kidney Disease

The chart displays the heart attack cases in adults who were diagnosed with kidney disease along with a heart attack.

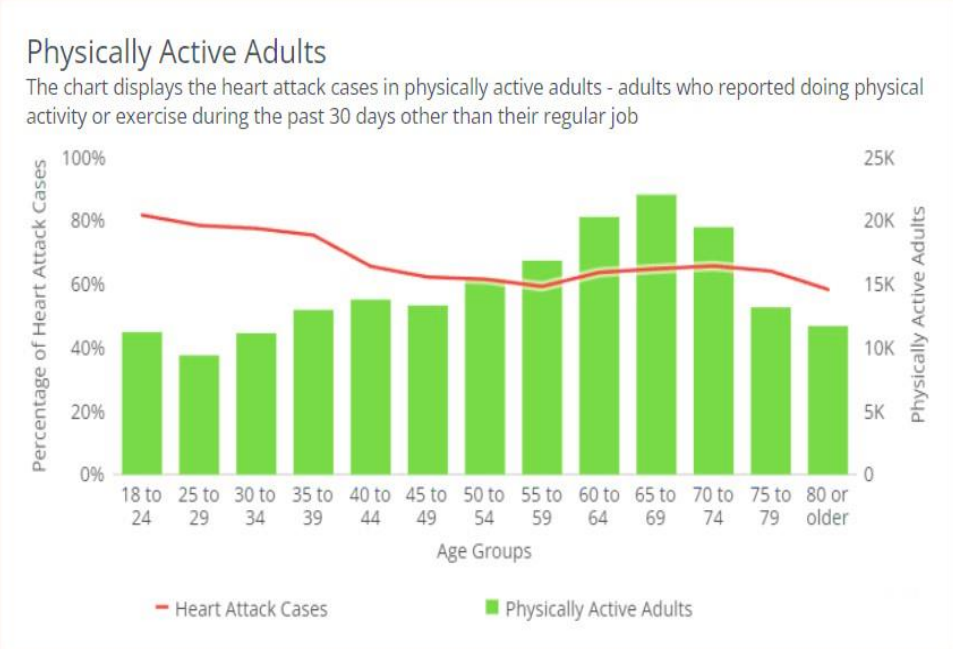


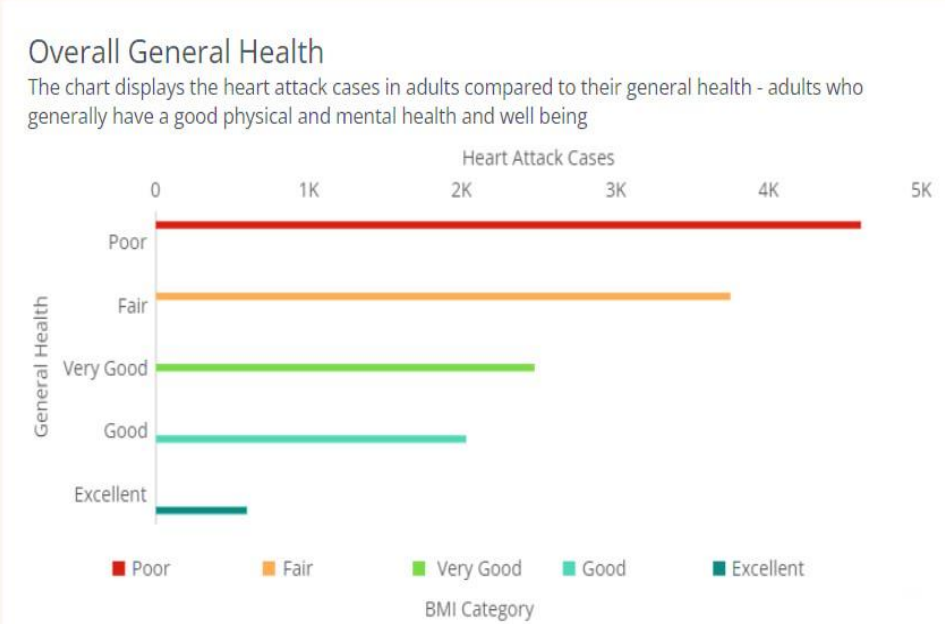
Adults with Asthma Disease

The chart displays the heart attack cases in adults who were diagnosed with Asthma disease along with a heart attack.



Cautions - *What has proven to limit the risks of heart attacks*





Introduction:

Heart diseases, particularly heart attacks, stand as one of the gravest health concerns globally. In the United States, 2022 witnessed a staggering rise in heart-related fatalities, surpassing cancer and even COVID-19. The alarming statistics prompted a profound exploration into the data surrounding heart diseases, emphasizing the urgency for awareness and preventive measures. The annual CDC survey data, "Indicators of Heart Disease," obtained from Kaggle, serves as the backbone of this visualization project, encapsulating health insights from over 400,000 adults in

the USA.

The primary goal is to raise public awareness by presenting easily understandable visualizations that delve into the metrics associated with heart diseases. The choice of a dashboard design, inspired by the flow and coherence found in Dr. Lace Padilla's impactful visualizations, aims to provide a clear, engaging, and informative narrative. The design not only adheres to the principles learned in the course but also responds to valuable feedback emphasizing simplicity and clarity.

The urgency of the matter is accentuated by red, strategically employed to symbolize the severity of the epidemic-like rise in heart diseases. This color choice aligns with the principles of expressiveness and effectiveness discussed in Chapter 5 of the visualization analysis and design textbook and reiterated in Dr. Lace Padilla's lectures.

Dataset Description:

The dataset, originating from the CDC's Behavioral Risk Factor Surveillance System (BRFSS), encapsulates a comprehensive 2022 survey of U.S. residents. With over 400,000 adult interviews annually, BRFSS stands as the world's largest continuously conducted health survey. The dataset encompasses crucial indicators like high blood pressure, high cholesterol, smoking, diabetes status, obesity, physical activity, and alcohol consumption. Preprocessing involved meticulous checks for errors, anomalies, and deeper exploration of metric definitions, ensuring a reliable foundation for visualization.

The journey through the dataset was not merely technical but a profound exploration into the health dynamics of a diverse population. Understanding the intricacies of each metric and its potential impact on heart diseases became a crucial aspect of the dataset exploration. The dataset, therefore, not only serves as a collection of numbers but as a narrative of health patterns and risk factors.

Design Rationale:

The visualization employs carefully chosen visual encodings, aligning with principles discussed in the course. Red, symbolizing urgency and negative implications, serves to highlight the epidemic-like rise of heart diseases. The design decisions are rooted in the principles of expressiveness and effectiveness discussed in Chapter 5 of the

visualization analysis and design textbook and reiterated in Dr. Lace Padilla's lectures.

Considerable thought went into the selection of alternatives, acknowledging the sensitivity of the topic. Prototypes were tested, including monochromatic designs and interactive approaches, with the final choice prioritizing simplicity and ease of understanding for a broad audience.

The iterative design process involved continuous refinement, keeping in mind the delicate balance between comprehensive data representation and viewer comprehension. The selection of the dashboard format itself was a strategic decision, allowing for a structured and cohesive flow of information from top to bottom.

Development Process:

The development process of this project was not just a sequence of technical steps, but a transformative journey marked by thoughtful decision-making and adaptability, guided by invaluable feedback. Despite being a solo endeavor, the project's development was enriched by insights from peer reviews and a keen awareness of the dataset's challenges.

Addressing the dataset's sheer size, a peer wisely cautioned against overwhelming the viewer. This resonated deeply, steering the choice away from a poster design to a more segmented dashboard format. The new design allowed for a logical flow of information from top to bottom, addressing concerns about potential viewer confusion.

In the initial stages, the dataset's exploration went beyond technicalities, mirroring the profound health exploration of a diverse population. This aligns with a peer's suggestion to ensure the metrics are not only numbers but also narratives of health patterns and risk factors. The dataset became a dynamic story waiting to be told, guided by the principle that data isn't just about numbers; it's about weaving a compelling narrative.

Navigating the dataset posed challenges, leading to explorations with D3 and Vega-Lite. This adaptation was prompted by a conscious acknowledgment of the dataset's size, showcasing a willingness to explore alternative tools and approaches. The development process wasn't a rigid adherence to conventional methods but a strategic series of decisions, adjustments, and resource allocations, with an unwavering commitment to creating a robust foundation for the final visualization.

Peer feedback played a crucial role in shaping the development process. Concerns about avoiding clutter and unnecessary complexity, especially given the sensitivity of the topic, became guiding principles. Every visual element was scrutinized for its clarity and purpose in conveying key insights, aligning with a peer's astute observation about the potential impact on viewers and the need for simplicity.

The meticulous adherence to the timeline wasn't just about meeting deadlines but a conscious effort to allocate time strategically. The emphasis on exploratory data analysis was a direct response to a peer's advice to be cautious about the dataset's overwhelming nature and to ensure a strong foundation for subsequent phases.

In summary, the development process was a dynamic journey, adapting to challenges, and incorporating peer insights. It reflects a conscious effort to balance technicality with storytelling, leveraging feedback to refine the approach and create a visualization that not only informs but resonates with its audience.

Feedback Incorporation:

Peer reviews played a pivotal role in shaping the final visualization. Balancing comprehensive data display and viewer comprehension was a key consideration, leading to the adoption of a dashboard design. Feedback urging caution regarding potential misinterpretations guided the avoidance of trends that could be misconstrued as causation. The inclusion of a choropleth map and the articulation of the visualization's purpose were direct responses to insightful peer suggestions, contributing significantly to the refinement of the project.

Specifically, one peer wisely pointed out the challenge of showing all the data without overwhelming the viewer due to the dataset's size. This resonated deeply, leading to the choice of a dashboard design where information is logically segmented, allowing viewers to absorb insights progressively.

Another astute observation highlighted the need to avoid clutter and unnecessary complexity, given the sensitivity of the topic. This feedback became a guiding principle throughout the design process, ensuring that every visual element served a clear purpose in conveying key insights.

The incorporation of a choropleth map was a direct response to a peer's suggestion, adding a geographical layer to the visualization. This enhancement aligns with the acknowledgment that different regions might have varying rates of heart diseases, contributing to a more comprehensive understanding.

Reflecting on the impact of peer insights, it becomes evident that this project is not just an individual endeavor, but a collaborative effort shaped by diverse perspectives. The collective intelligence of the learning community significantly contributed to elevating the quality and effectiveness of visualization.

Final Thoughts:

The project offered profound insights into the intricate process of developing professional visualization. Lectures, guest presentations, textbooks, articles, and peer critiques collectively shaped the outcome. The appreciation for visualizations has transcended mere aesthetics, and a deeper understanding of methodologies and theories has been gained. This project stands as a testament to the transformative impact of visualization knowledge on perspective.

The journey wasn't merely about creating a visualization; it was about learning to see data differently. A heightened awareness of the responsibilities that come with visual storytelling emerged, prompting a thoughtful reflection on the ethical dimensions of data presentation. This project was not just a culmination but a beginning—a beginning of a more nuanced and thoughtful approach to visual communication.

Acknowledgments:

Professor Dr. Lace Padilla: Invaluable guidance and inspiration.
Michael Corell: Insightful guest lecture.
Visualization Analysis & Design textbook: Fundamental learning resource.
Websites: Kaggle for dataset, Coolers for color pallet.
Tools: Domo, Tableau, Power BI, MS PowerPoint.
Dataset Link: https://www.kaggle.com/code/georgyzubkov/heart-disease-exploratory-data-analysis

This project is a culmination of collective knowledge and collaborative efforts, and the acknowledgment extends to the entire learning community.