# PROJECT 3

Samuel Montesinos, Erik Arbelo Nguyen, Kelly Le

INTRODUCTION TO HEART DISEASE

Heart disease is a term that refers to several different types of heart conditions. It is one of the leading causes of death worldwide, and it affects millions of people each year. Heart disease can be caused by various factors, such as genetics, lifestyle choices, and underlying medical conditions. The significance of heart disease in public health is that it has a significant impact on both individuals and society as a whole. The economic burden of heart disease, in terms of healthcare costs and lost productivity, is substantial.

The purpose of this presentation is to demonstrate how we can use data to gain a better understanding of heart disease indicators and how we can visualize them effectively. By doing so, we can identify patterns and correlations that may help us develop more effective interventions and prevention strategies. We will explore the sources of data used in the project, the data wrangling techniques used to prepare the data for analysis, and the coding approach used to analyze the data. Finally, we will showcase a final visualization that effectively represents the heart disease indicators analyzed in the project.
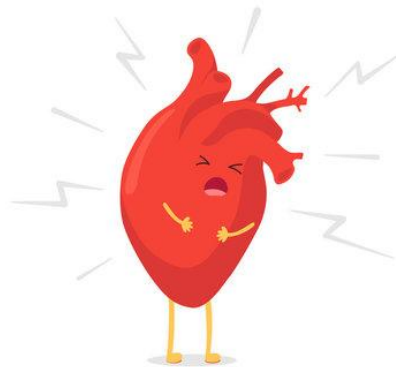
DATA SOURCES UTILIZED

- We used kaggle to wrangle with a survey from 2022 containing 300,000+ participants that answered various healthcare questions, such as their age, any past diseases, sleep schedule, if they're a smoker, etc.
- The dataset gave us a CSV to work with which we modified in many ways to optimize our code without affecting the core data.
- https://www.kaggle.com/datasets/ambujdevsingh/key-indicators-of-heart-disease
- Link is dead as the dataset is frequently updated. Unfortunately, we do not have a link to an updated dataset yet.

Data Wrangling

- As previously mentioned, we had to modify the CSV using Data Normalization.
- The next step was to create a database that would serve as the core of our research. We went through many stages to find the correct SQL format that would work before ultimately settling on SQLite
- We also utilized a Python API to help create Graphics to easily visualize the data.

CODING APPROACH

- To create the database, we utilized a new tool called DB Browser.
- For our API, as mentioned before we utilized Python
- To create our graphics, we utilized html, and javascript to plot and arrange a dashboard from which we can zoom in on individual data, as well as download the graphics for use in our presentation.
- For our interactive quiz, we utilized a combination of html, css, and javascript to create a quiz that will provide a score based on response that will reflect your likelihood for heart disease (please note this is just a quiz for personal fun use. There is no need to take any indications seriously as heart disease occurrence can vary on many variables that are not limited to the ones used in the survey)

HEART DISEASE INDICATORS AND VISUALS

Below is the list of heart disease indicators we are looking at today:

- BMI
- Smoking
- Alcohol Drinking
- Stroke
- DiffWalking
- Sex
- Age Category
- Race
- Diabetic
- Physical Activity outside of work
- GenHealth
- SleepTime
- Asthma
- Kidney Disease
- Skin Cancer

Indicator: Stroke Count

Below we see stroke victims have a 36% chance of having heart disease compared to a 7% chance of having heart disease without a history of a stroke.

Stroke Counts with Heart Disease

Stroke Counts without Heart Disease
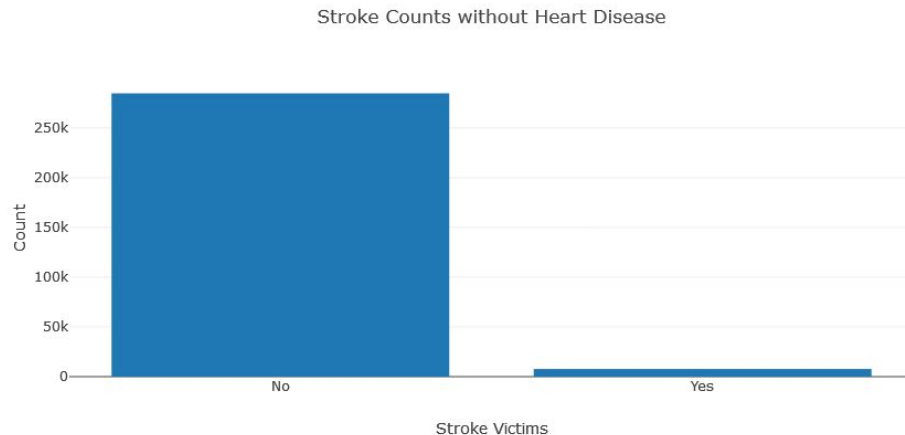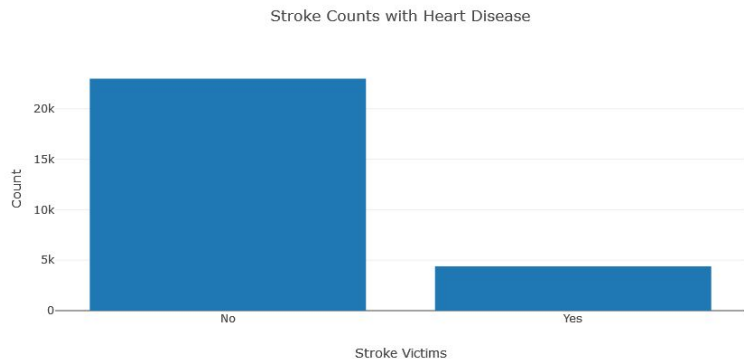
Indicator: Difficulty Walking

Below we see patients with difficulty walking have a 22.6% chance of having heart disease compared to a 6.3% chance of having heart disease without a history difficulty walking.

Individuals that have Difficulty Walking with Heart Disease
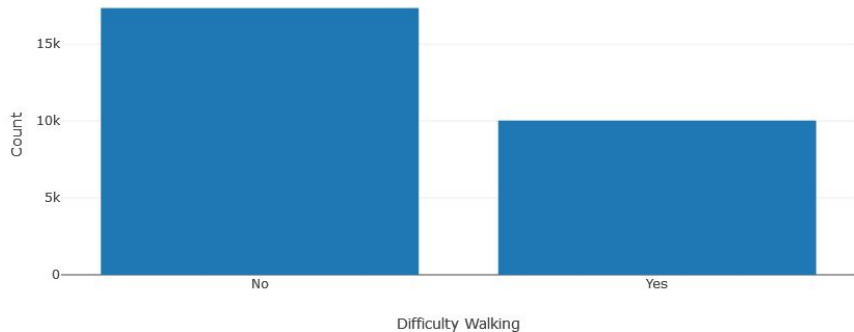
Individuals that have Difficulty Walking without Heart Disease
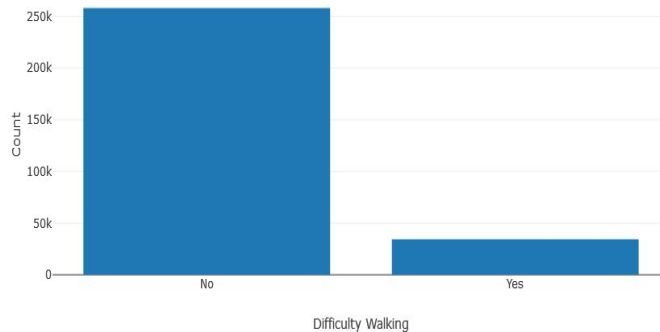
Indicator: Age

When looking at the below graphs, we can see a strong rise in heart disease as the participants increased in age. This trend really begins to show itself beginning at a minimum of 45 years of age.

Lack of Heart Disease based on Age
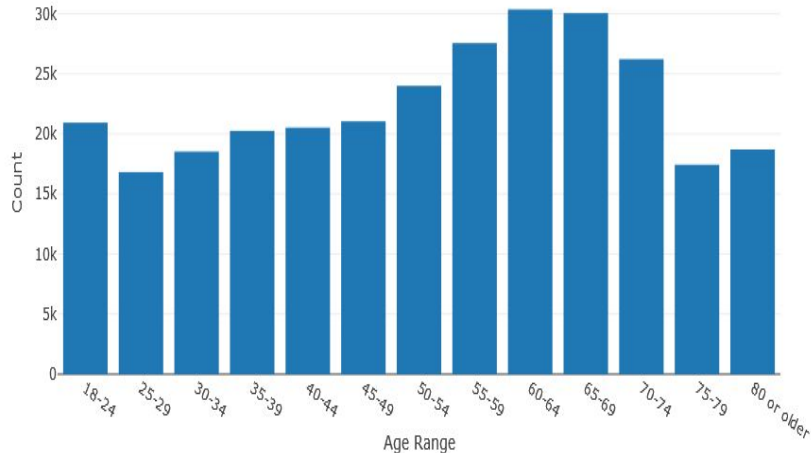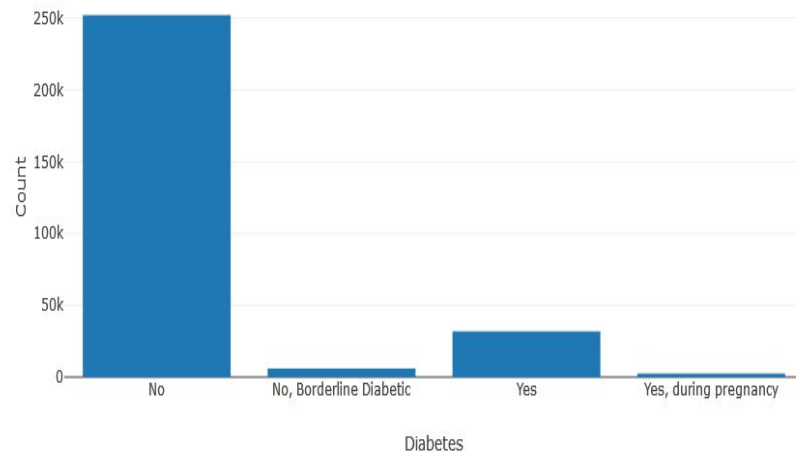

Occurences of Heart Disease based on Age
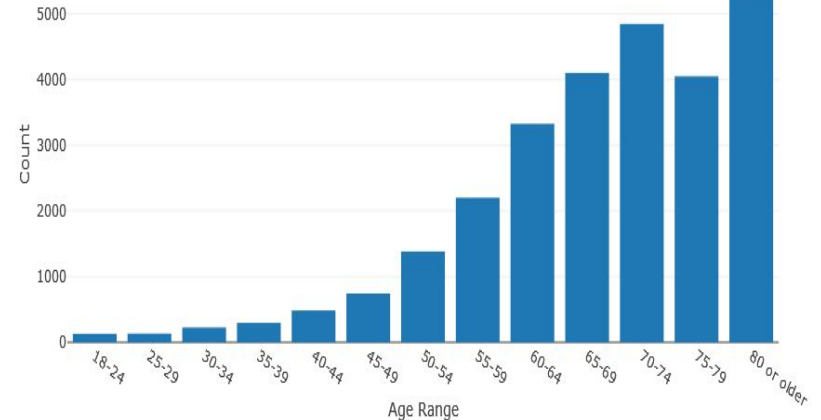
Indicator: Diabetes

Below we see patients without diabetes have a 6.4% chance of having heart disease , borderline diabetics have a 11.6% chance, diabetics have a 22% chance, and those that were diabetic during pregnancy have a 4.2% chance.


Occurences of Heart Disease based on Diabetes
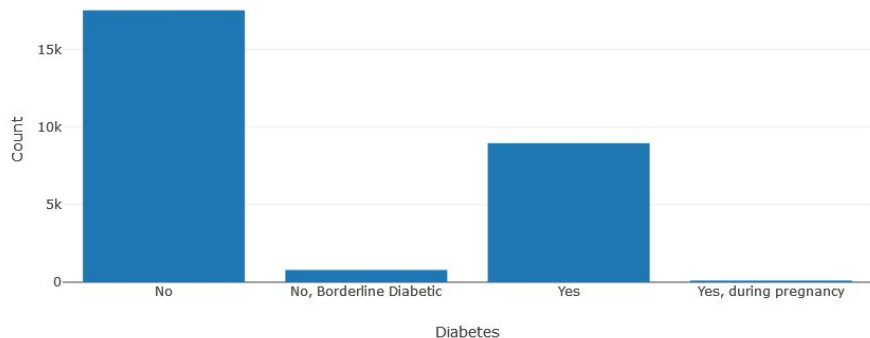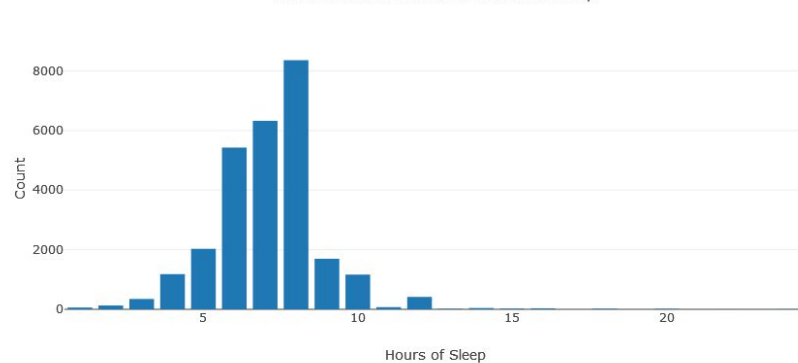

Lack of Heart Disease based on Diabetes

Indicator: Hours of Sleep

Below we see a downward trend where individuals with less sleep have a higher chance of heart disease. This is especially prevalent in 5 hours or less of sleep where individuals have an increased chance of as much as 20% for heart disease..

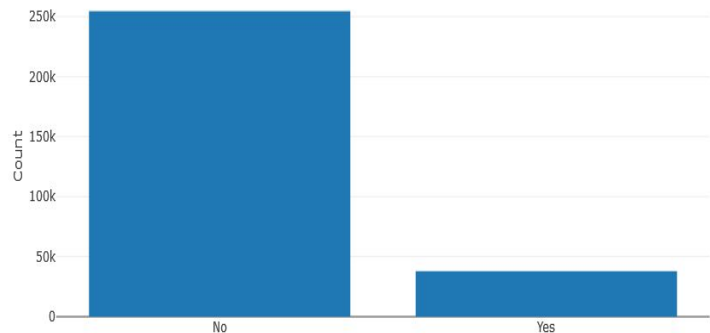Lack of Heart Disease based on Hours of Sleep
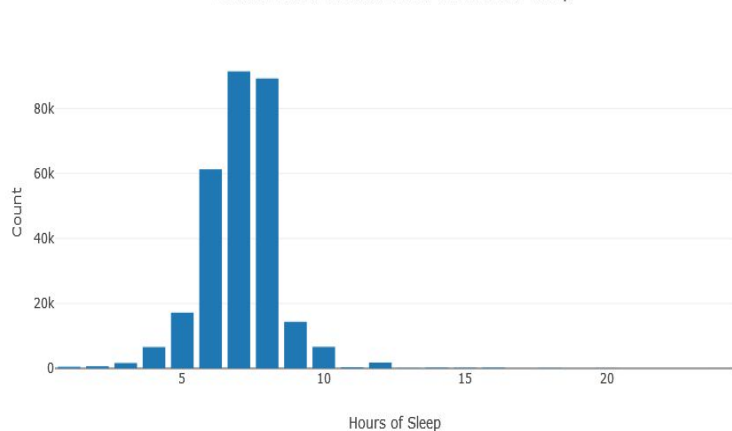
Heart Disease based on Hours of Sleep
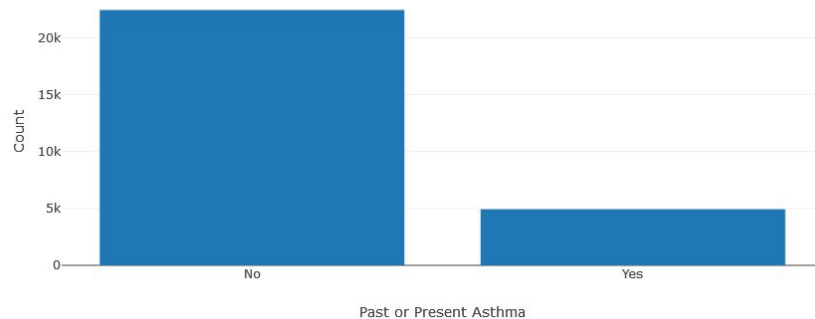
Indicator: Asthma

Below we see those with asthma have a 11.5% chance of heart disease compared to an 8% chance of heart disease without asthma
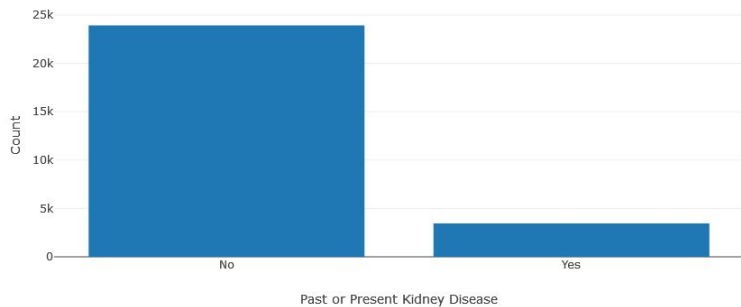

Lack of Heart Disease based on Asthma
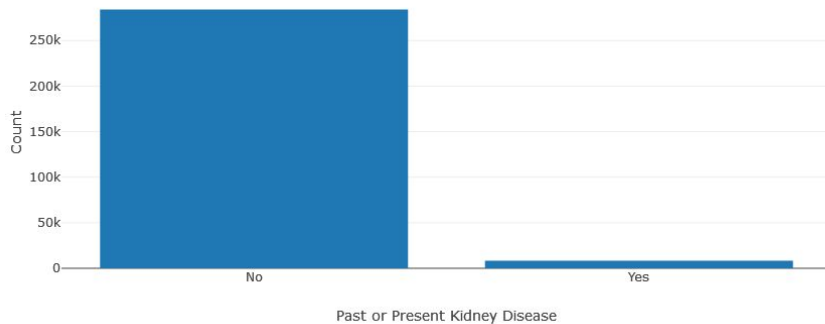

Heart Disease based on Asthma

Indicator: Kidney Disease

Below we see those with Kidney Disease have a 29.3% chance of heart disease compared to a 7.8% chance of heart disease without Kidney Disease

Heart Disease based on previous/present Kidney Disease

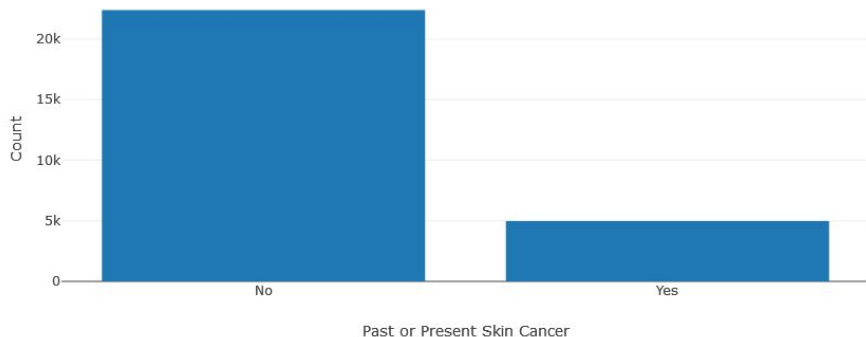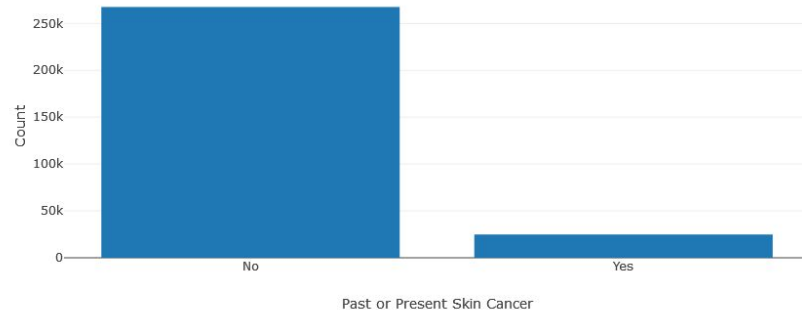Lack of Heart Disease based on previous/present Kidney Disease

Indicator: Skin Cancer

Below we see those with Skin Cancer have a 16.7% chance of heart disease compared to a 7.7% chance of heart disease without Skin Cancer

Heart Disease based on previous/present Skin Cancer

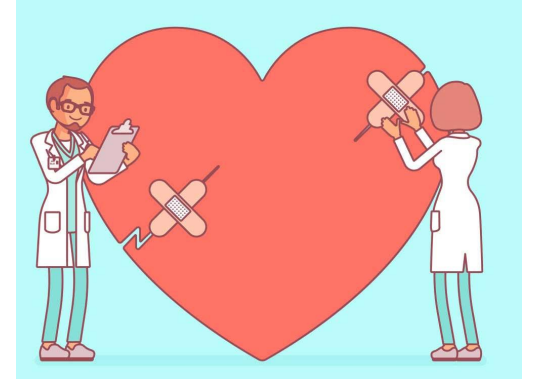Lack of Heart Disease based on previous/present Skin Cancer

INTERACTIVE QUIZ

- kiernguyen.github.io

CONCLUSION

Conclusion

- Summarize the key takeaways from the presentation.
- Reinforce the importance of understanding heart disease indicators and the potential for data to inform public health policies and interventions.

QUESTIONS?

THANK YOU FOR LISTENING!!