HEART DISEASE DATA ANALYSIS

INTRODUCTION

Heart disease remains a significant public health concern globally, impacting millions of lives each year. To better understand the patterns, risk factors, and possible interventions associated with heart disease, this data analysis project aims to delve into a comprehensive dataset encompassing various aspects related to cardiovascular health. We study on impacts of certain lifestyle diseases and other general aspects that are meant to be influencing heart diseases.

Data Analysis is carried out using Python and its data analysis libraries such as Pandas, NumPy, Matplotlib, Seaborn, etc.

Fields considered in the analysis

Age of the patient

Sex of the patient

Exercise-induced angina

Number of major vessels

Chest pain(1.Typical,2. Atypical,3. Non-anginal,4. Asymptomatic)

Resting blood pressure

Cholesterol in mg/dl

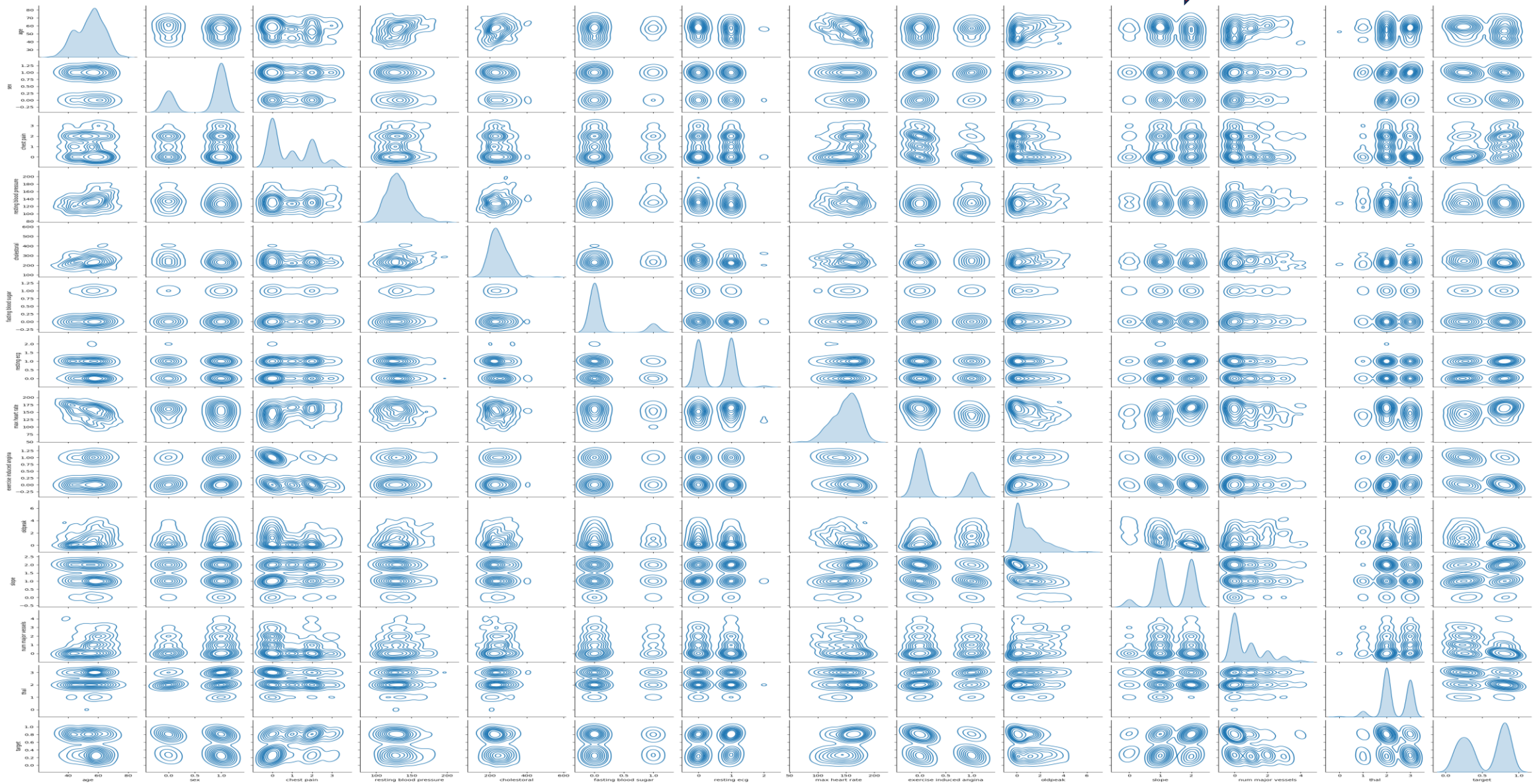
Fasting blood sugar

Resting ECG

Maximum heart rate

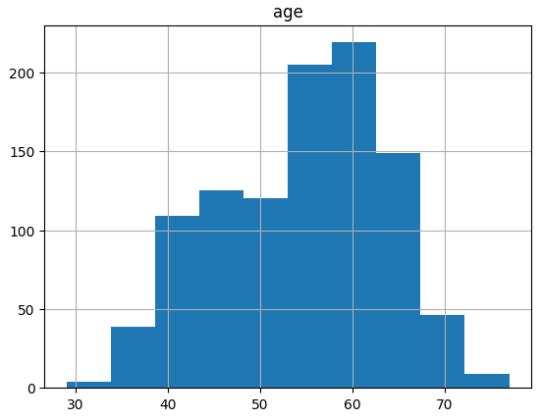
Target (0 less chance of heart attack and 1 more chance of heart attack)

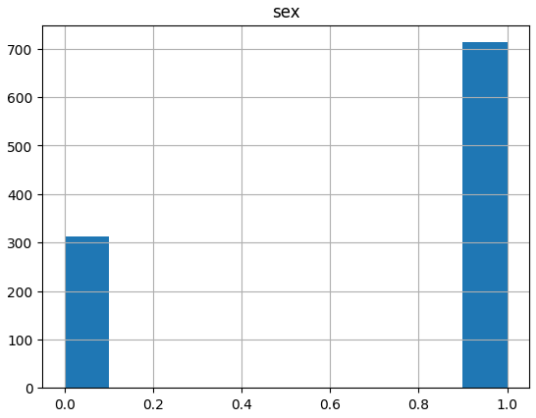
Pair plot representation of the fields under study

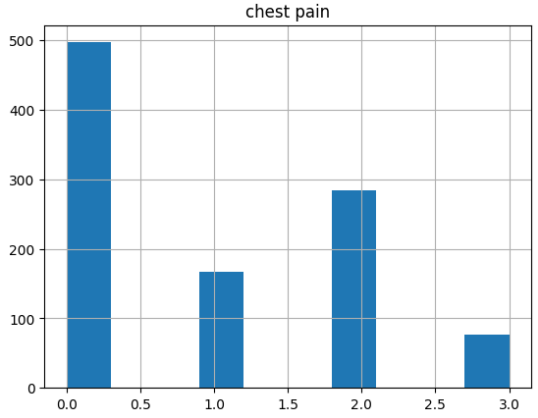


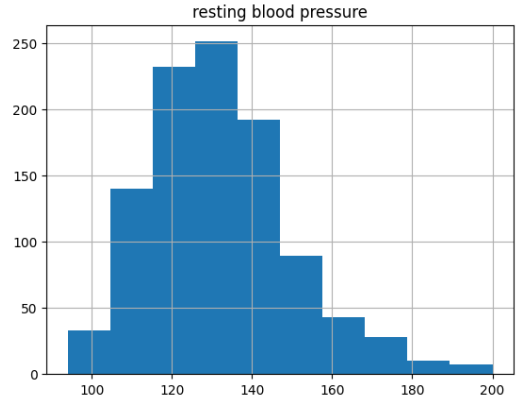
Categorical analysis of each field of study

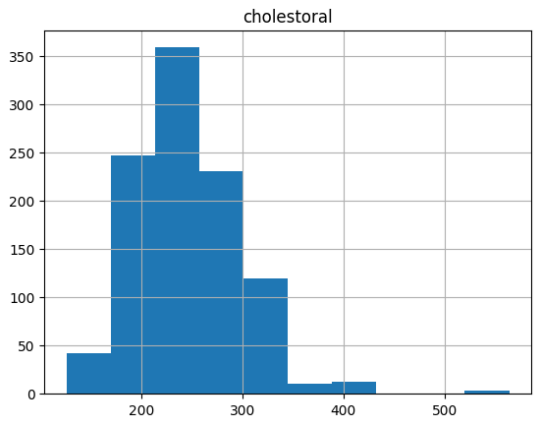
Category analyzing the distribution of the data set of study using histograms.

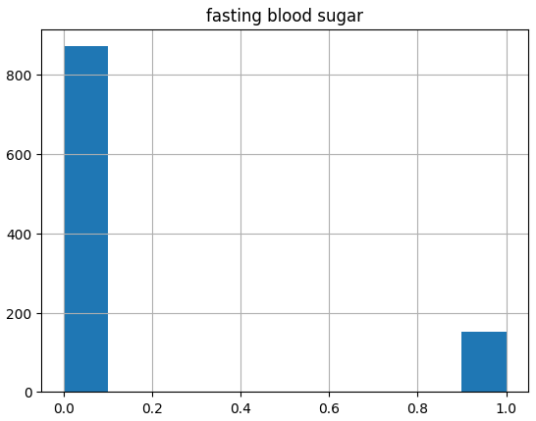


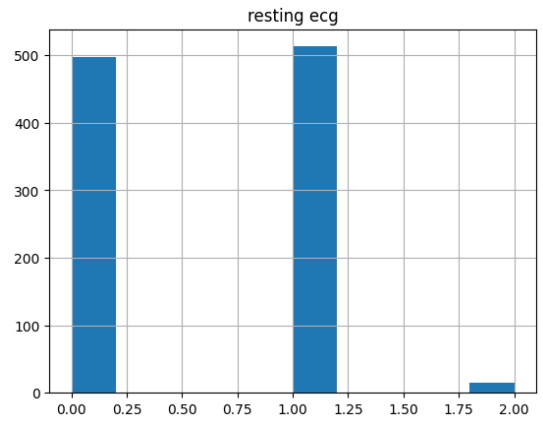


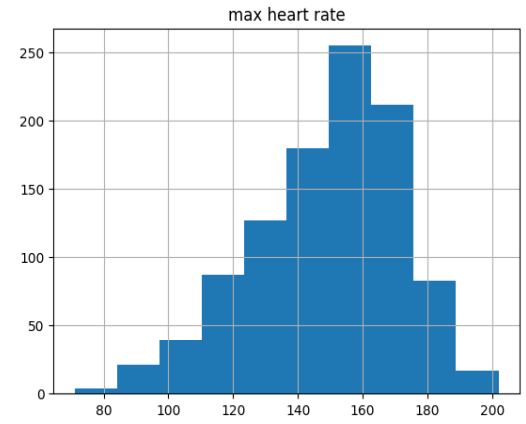


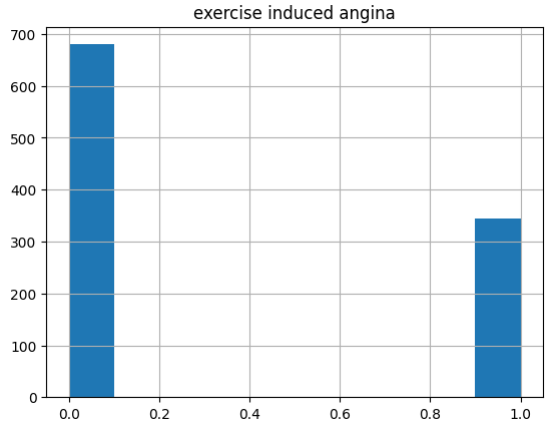


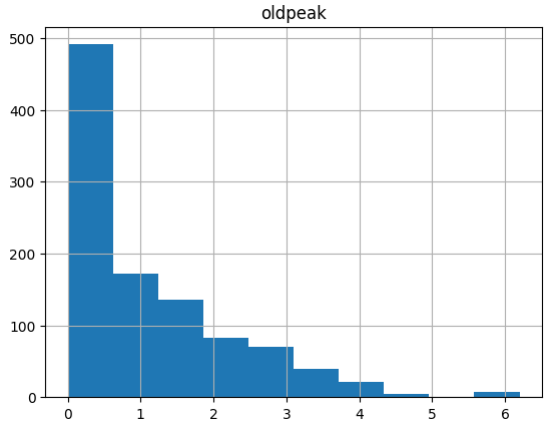


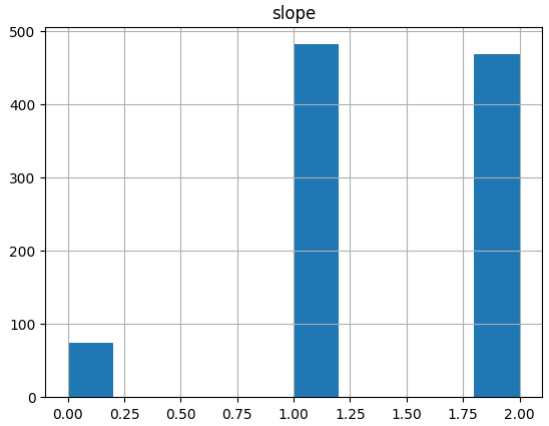


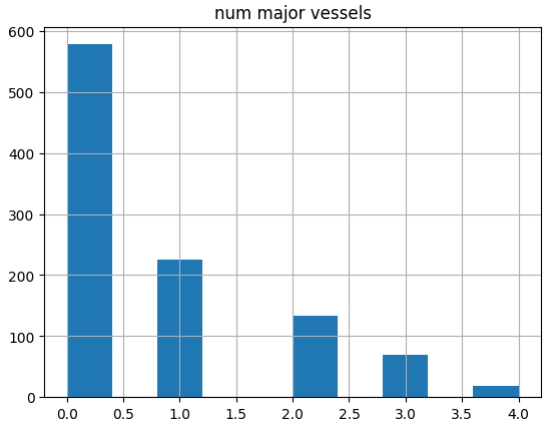


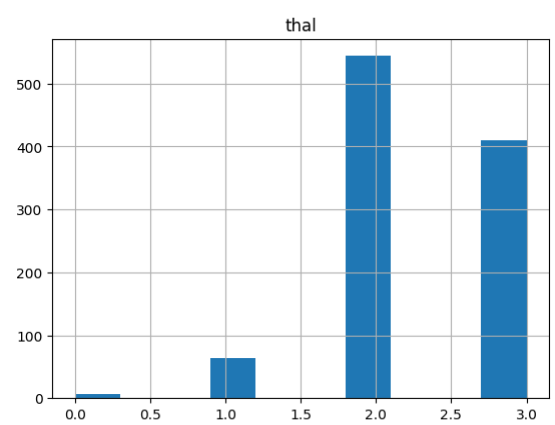


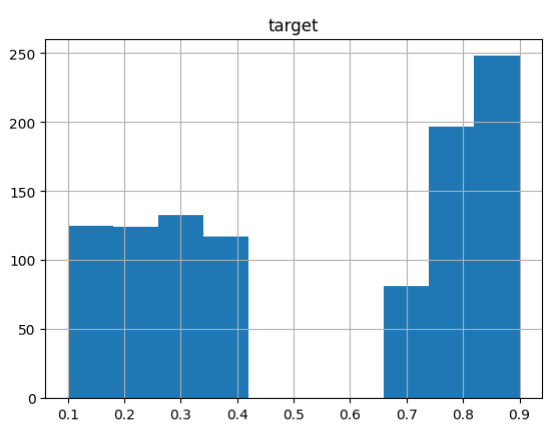












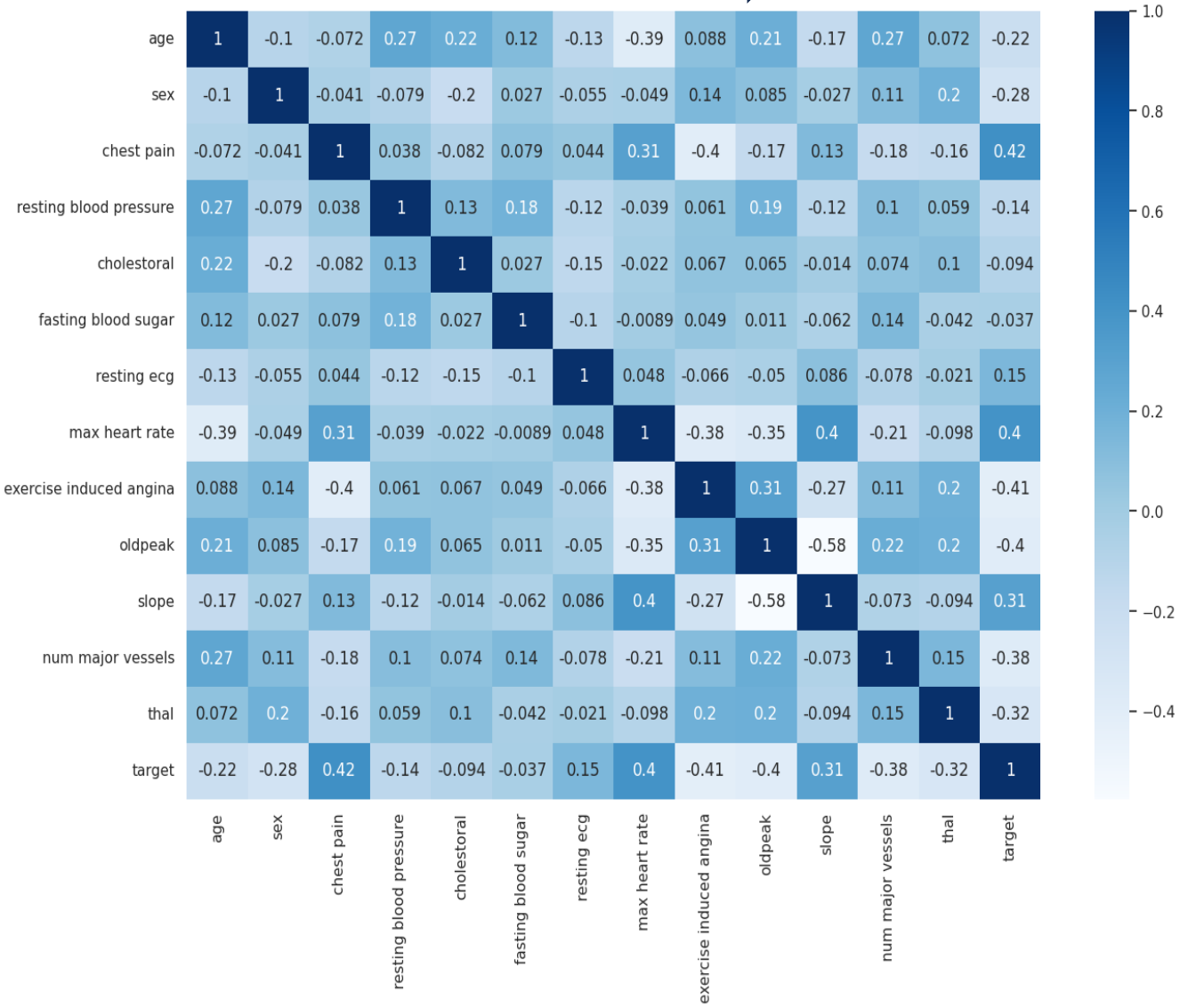
This shows the overall distribution of the data collected in each field that is under study and gives an overall idea about each field.

Correlation Study on Data Using Heatmap

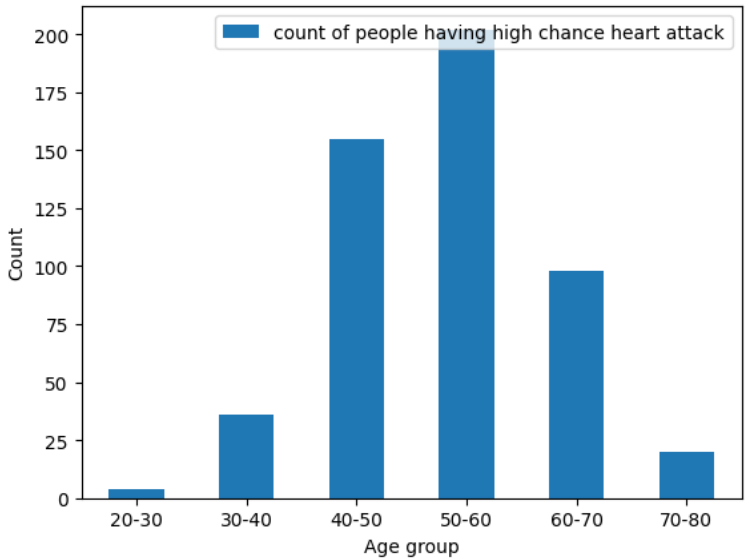
A correlation heatmap is a graphical representation of the relationships between different variables in a dataset.

The study on the heatmap provides the idea that,

- **Chest Pain, Maximum Heart rate, and Slope of the peak exercise ST segment show a good positive correlation with the target which is the probability of having heart disease.**
- **Old peak(ST depression induced by exercise relative to rest), exercise-induced angina, number of major vessels, sex, and age show a good negative correlation with the target.**
- **Fasting blood sugar, cholesterol, and resting ECG have a low correlation with our target.**



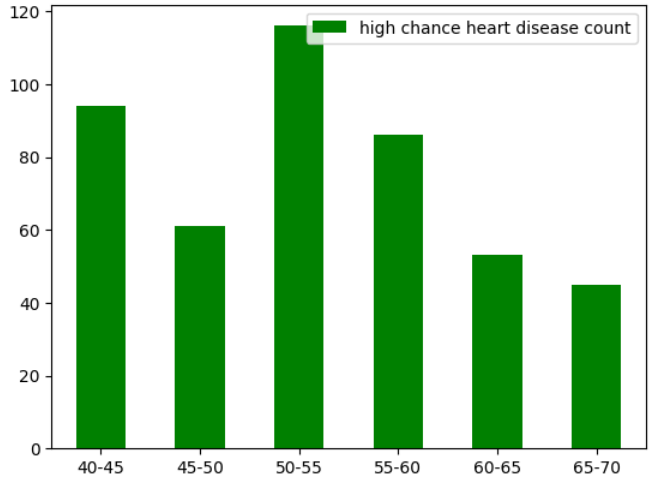
Study on age and heart attack.



Grouping the age distribution in the range of 10 and studying the distribution of people in each age group.

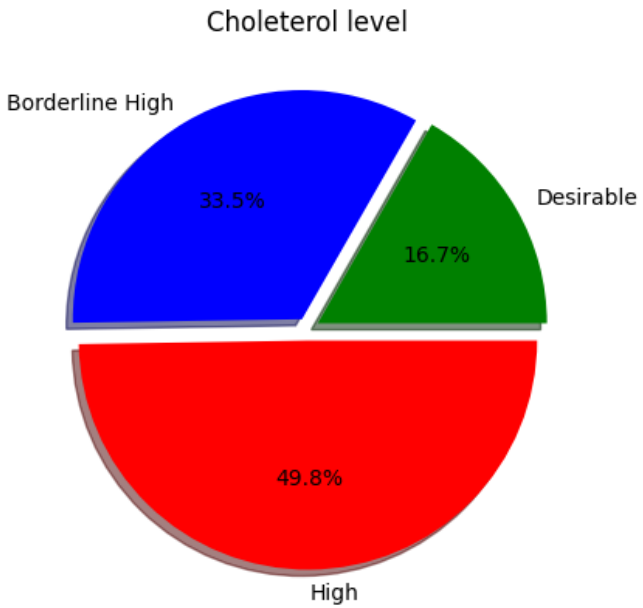
- **The age group 50-60 has the highest count of people having a high chance of heart attack.**
- **The 40-50 age group is second highest and 60-70 is third.**
- **About more than 80% of people have a high chance of heart attack lying in 40-70.**

Let's do further study in this age group to predict the age for which heart attack is common,



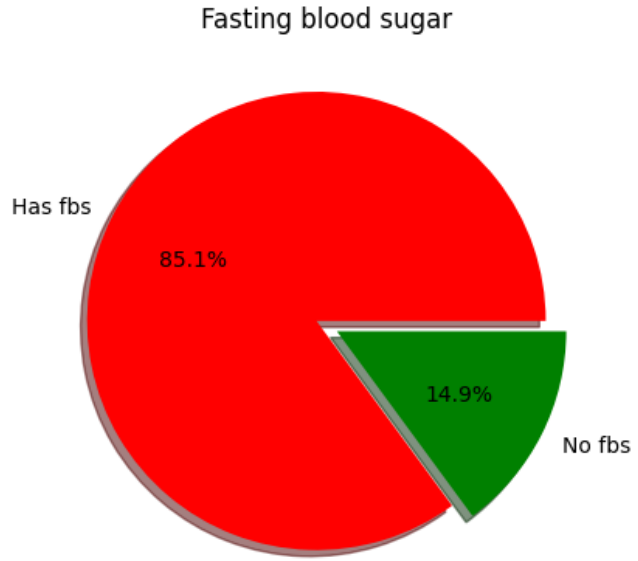
Further division of the age group into the range of 5 gives more idea that the age group 50-55 is most common. The group 40-45 is the second highest and 55-60 is the third highest.

Cholesterol and Blood sugar on heart disease



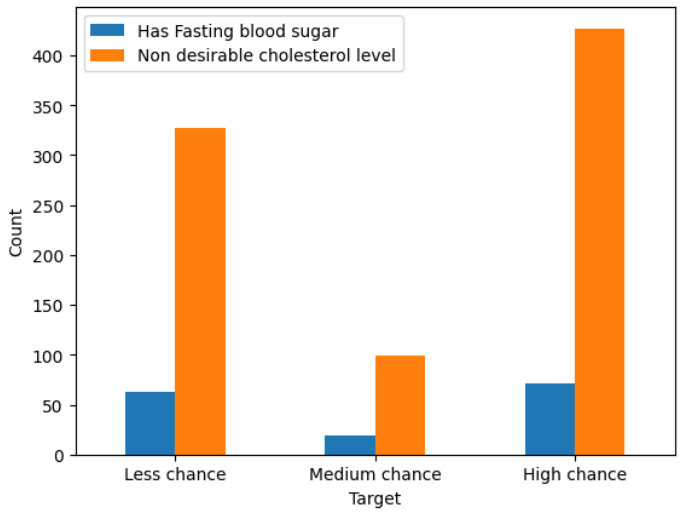
- About 50% of the people who have been analyzed for heart diseases have high cholesterol levels and 33.5% have borderline which is also greater than normal levels.
- Thus, more than 83% of individuals with heart disease display cholesterol levels exceeding the desirable limit.

From the data, we can infer that Cholesterol level has a significant role in the case of heart disease.



- More than 85% of people who are inspected for heart disease are having Fasting blood sugar.

From the visualization, we can infer that Fasting blood sugar has a significant role in the case of heart disease.



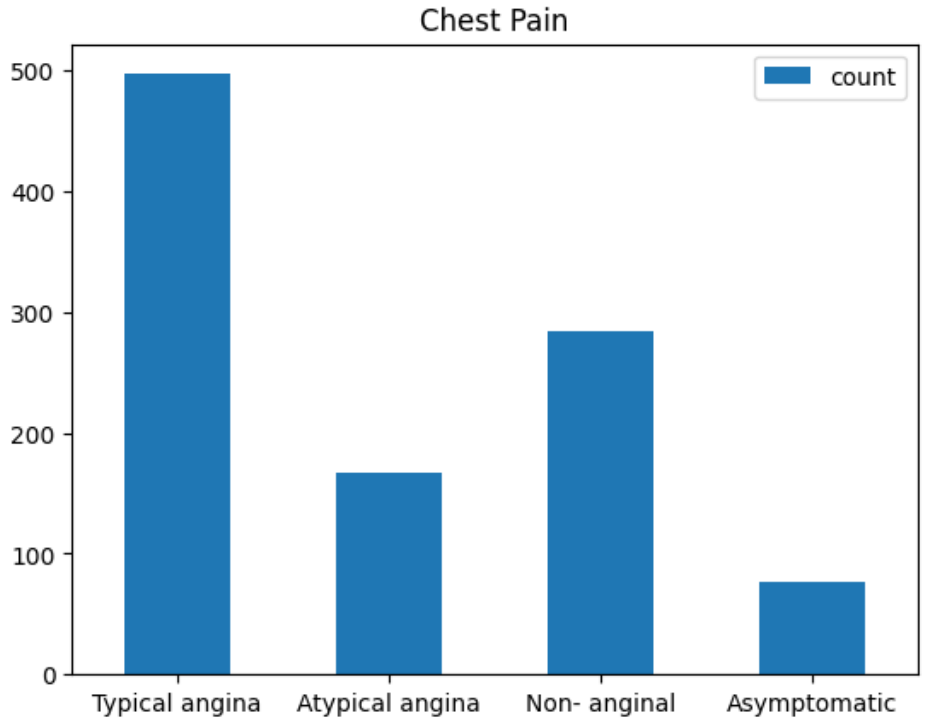
- **Bar plot represents a count of patients having non-desirable cholesterol levels (both borderline and high levels) (orange bars) and Fasting blood sugar (blue bars) against the target(Chance of heart attack).**
- **More than 50% of the patients investigated who have Non-desirable levels of Cholesterol have a higher chance of Heart attack.**
- **About 50% of overall patients investigated having Fasting blood sugar have a higher chance of Heart attack.**

Recommendation based on analysis of Cholesterol and Fasting Blood Sugar

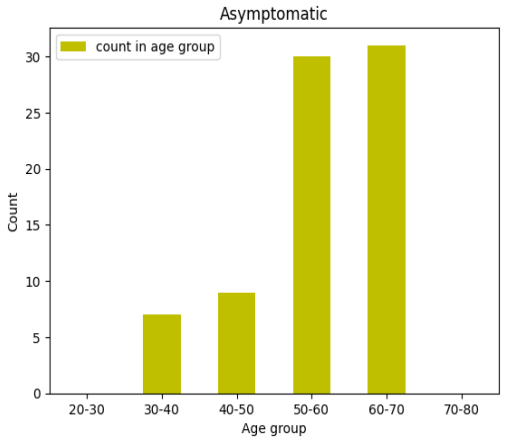
- **Managing heart disease necessitates focused efforts on cholesterol and fasting blood sugar control due to their strong ties to cardiovascular health. Over 83% of heart disease patients have high cholesterol, while 85% exhibit elevated fasting blood sugar levels, underlining the need for dietary adjustments, exercise, and, if necessary, medication.**
- **Understanding the data presented, it's clear that managing cholesterol levels and fasting blood sugar is critical in reducing the risk of heart attacks.**

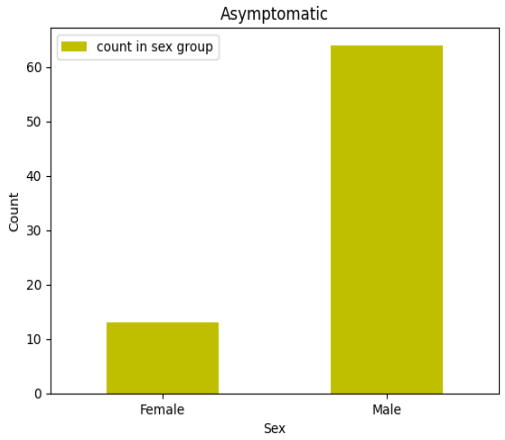
Analysis on types of Chest pain and effect on Heart attack

- Overall analysis of the four types of Chest Pains are plotted in the bar plot.
- Typical anginal patients are the most in the count among all the four types.
- Second most is Non-Anginal pain comparatively high rather than Atypical and Asymptomatic.
- Even though Atypical and Asymptomatic anginas are comparatively less count among people, need more care since it has a significant effect on heart attacks.



Asymptomatic





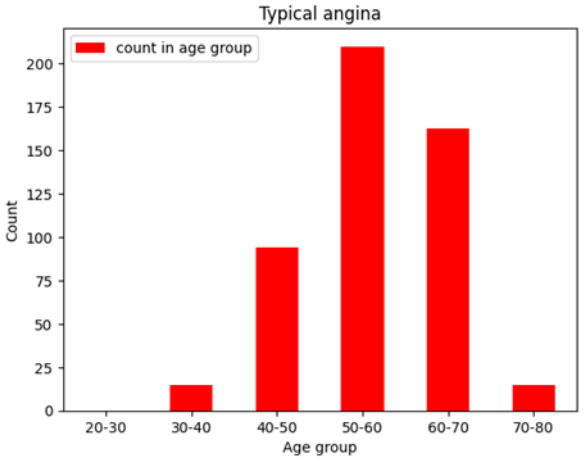
- Asymptomatic anginal pain typically manifests as chest discomfort, pressure, or a squeezing sensation, usually occurring when the heart muscle doesn't receive enough blood flow, often due to narrowed or blocked coronary arteries.
- Mostly found in the 60-70 age group and also the 50-60 group.
- The affected counts are comparatively very small which shows it is rare.
- Very low count of female patients.

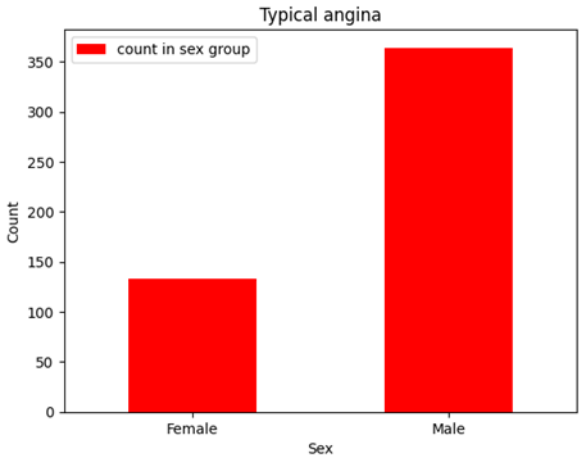
Typical angina

Typical angina is often characterized by symptoms such as chest pain or discomfort that occurs with physical activity and goes away with rest.

The most commonly found typical angina is mostly found in the 50-60 age group.

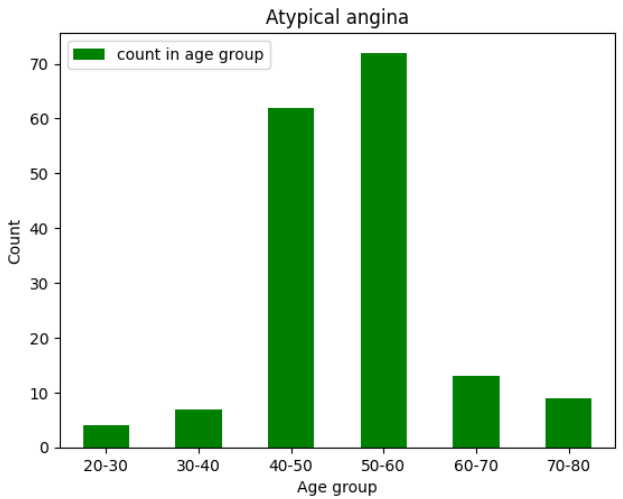
Comparing among sexes, the male has a significantly more count than female.

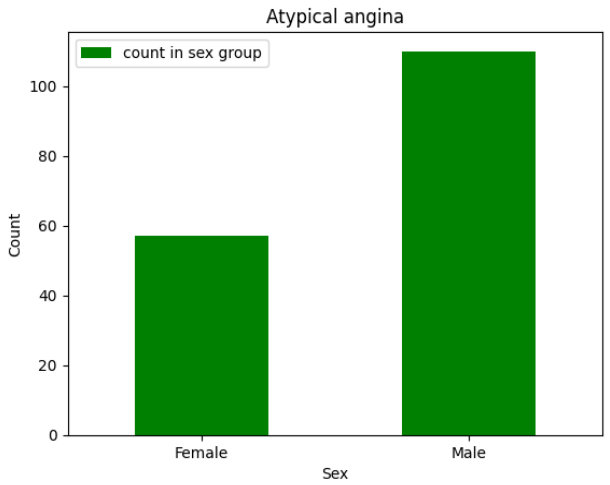




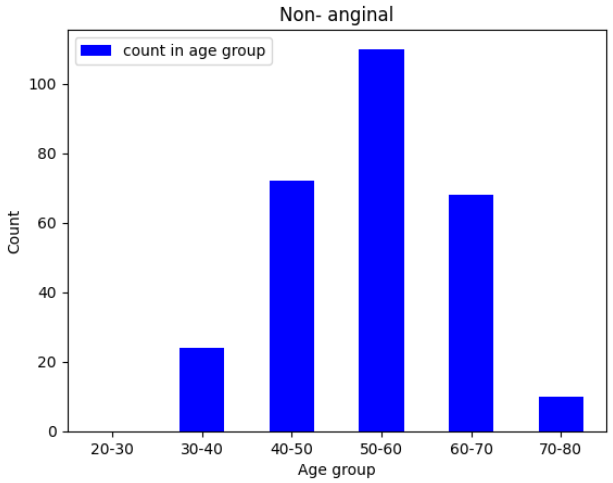
Atypical Angina

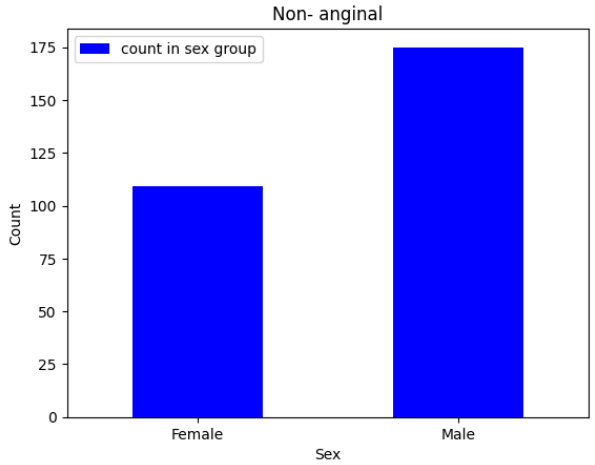
- Atypical angina describes symptoms that do not fit the typical pattern of angina.
- More than 80% of Atypical angina patients lie in the 40-60 age group.
- Comparing among sexes, the male has a significantly more count than female.





Non-Anginal pain

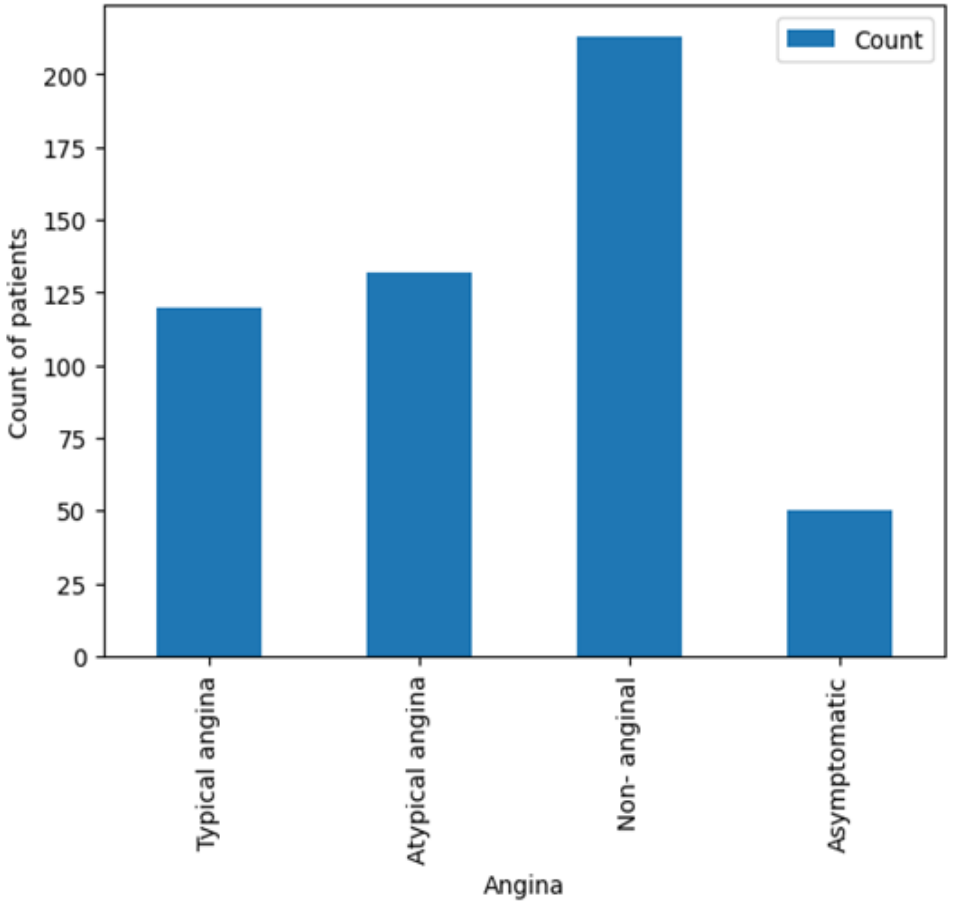




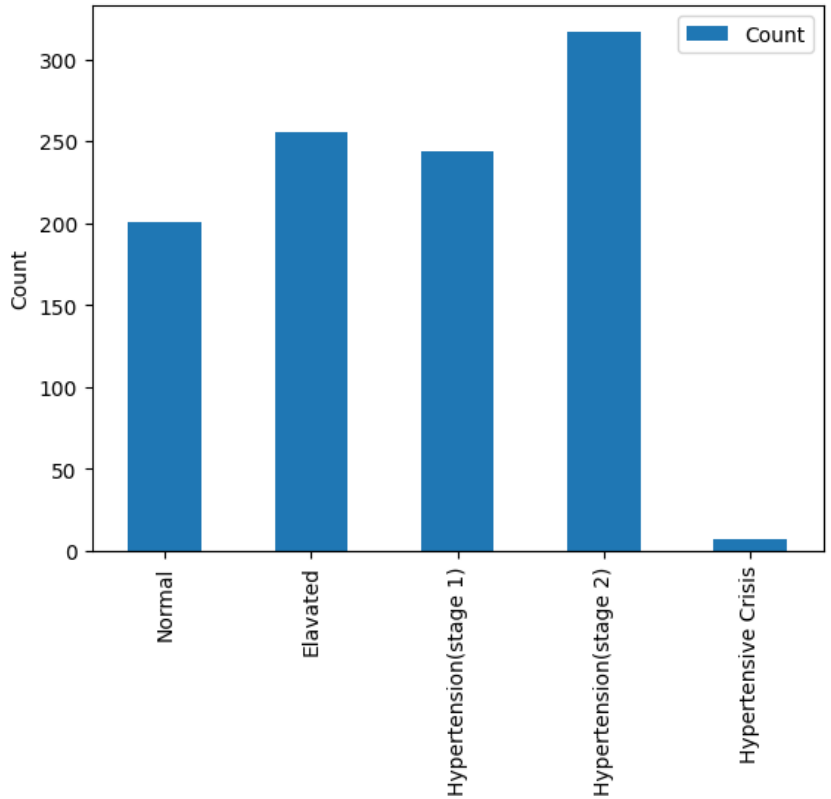
- Non-anginal pain refers to chest pain that is not caused by coronary artery disease.
- Apart 50-60 age group 40-50 and 60-70 age groups has also significant count of patients.
- Compared to other anginas male-female ratio on the counts is higher.

Chest pain Vs Chance of heart attack

- **The bar diagram represents the count of patients having a higher chance of a Heart Attack who have suffered different types of anginal pain.**
- **Even though Typical angina is most common, patients who suffer from Non-Anginal pain have a significantly higher chance of Heart attack.**
- **Atypical angina and Typical angina has also comparatively higher chance than Asymptomatic anginal pain.**

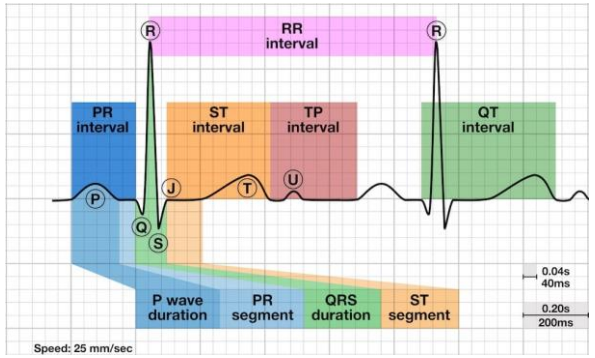


Analysis on resting blood pressure of heart disease patients



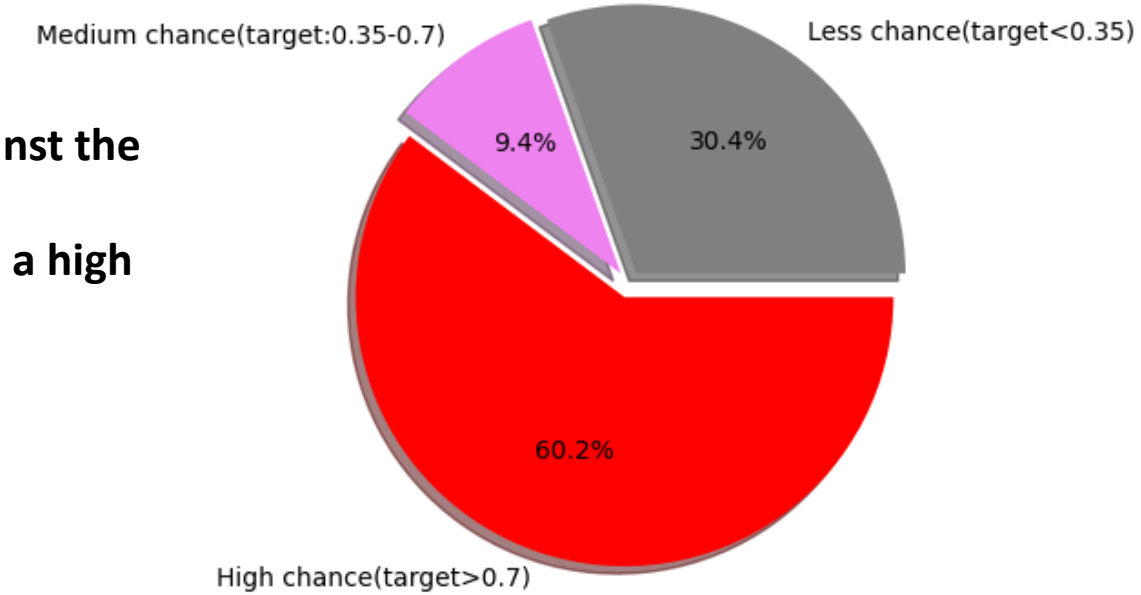
- Analyzing patients having heart disease from the data set to come up with a plot that represents the distribution of the Resting blood pressure among patients.
- Here we infer the Systolic pressure(in mm Hg)which measures the pressure in your arteries when your heart beats.
- From the figure we can infer that the resting blood pressure of more than 75% of people is higher than normal.
- More people who have some kind of heart disease is suffering from Hypertension (stage 2 and stage 1).

Analysis on Resting ECG with S-T wave abnormality with Heart Attack

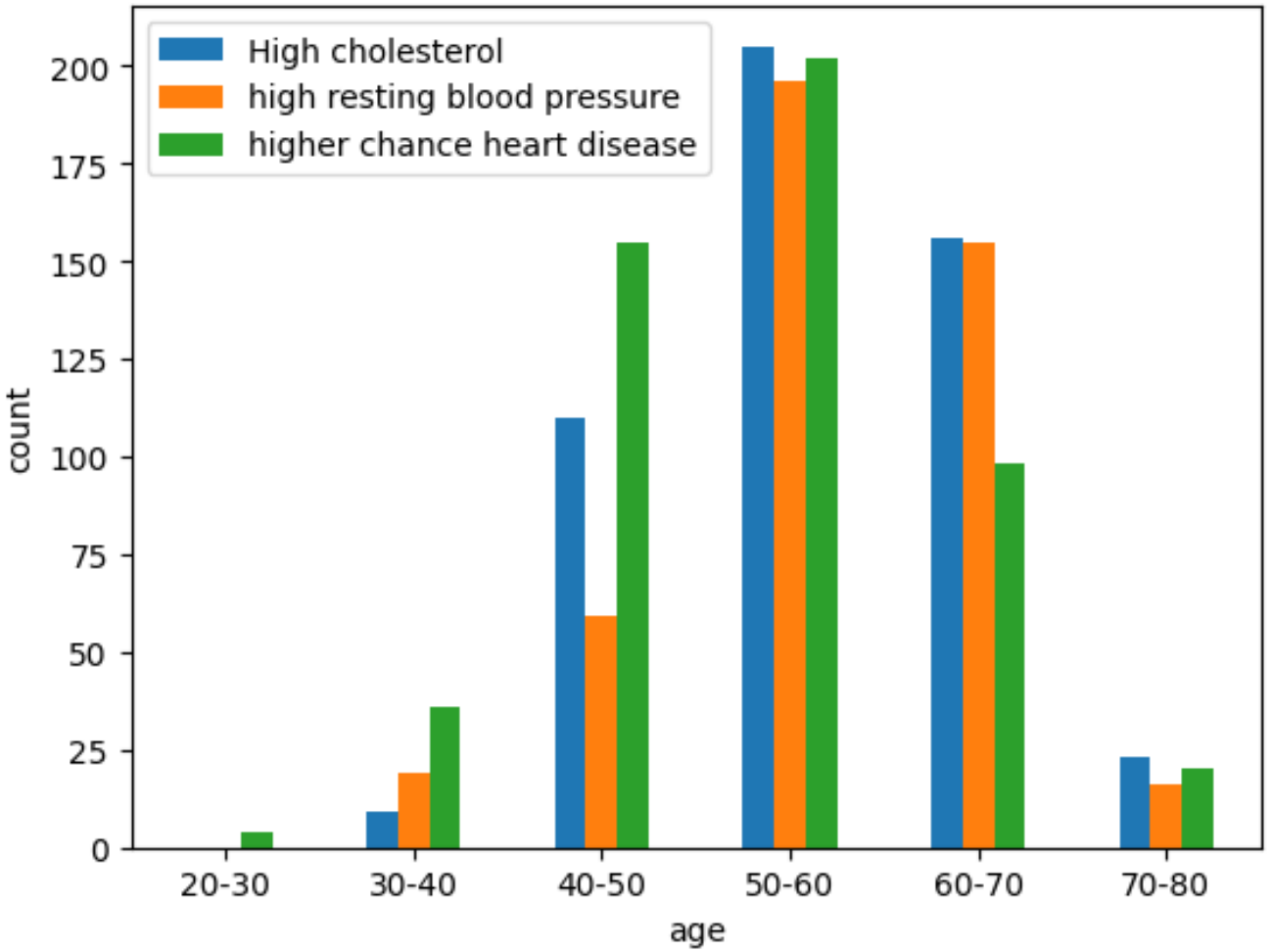


- The ST segment is the flat, isoelectric section of the ECG between the end of the S wave (the J point) and the beginning of the T wave.
- The ST Segment represents the interval between ventricular depolarization and repolarization.

- Analyzed the patients having S-T wave abnormality against the chance of heart attack.
- More than 60% of patients having S-T abnormality have a high chance of Heart attack.

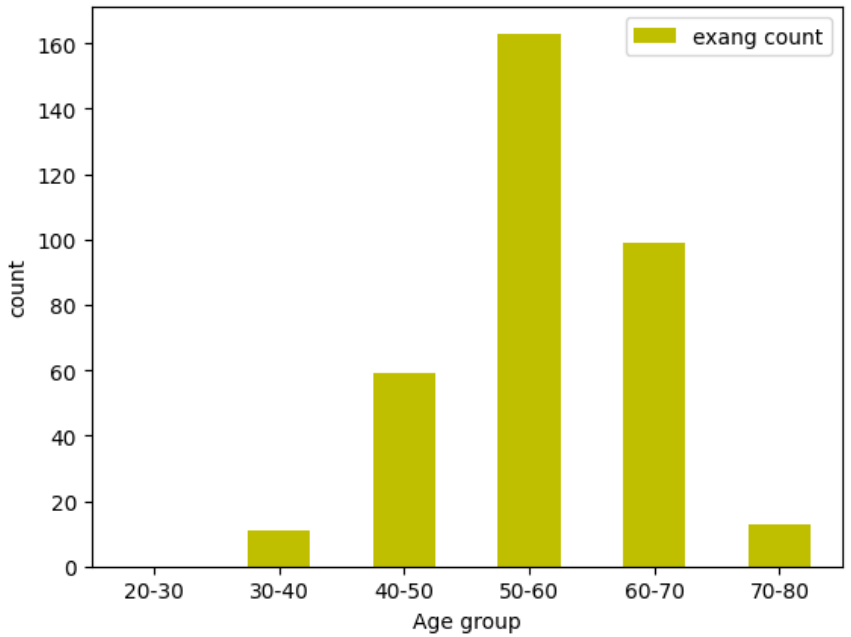


Cholesterol, RBP, Target within age groups



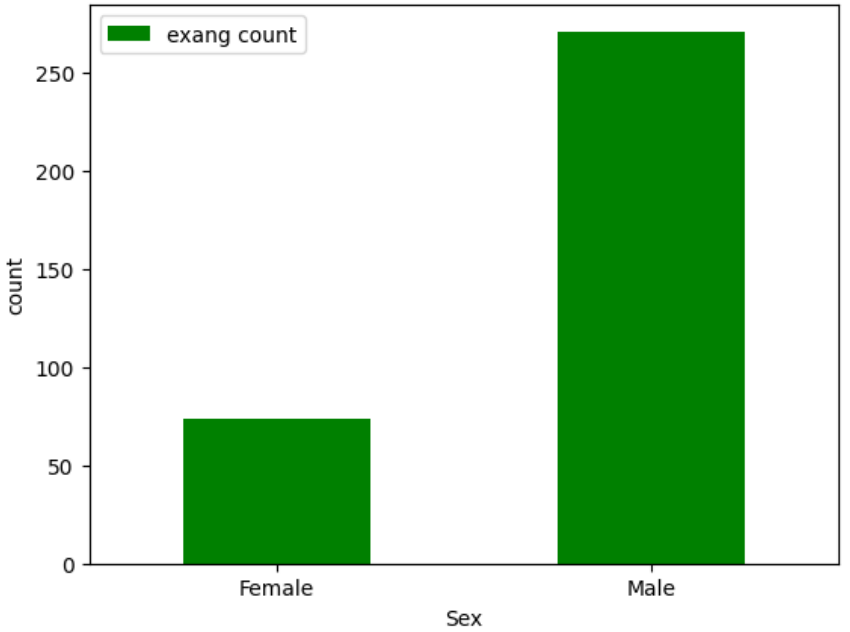
- **Partitioned patients having high cholesterol levels, high resting blood pressure, and a high chance of heart attack in different age groups.**
- **The 50-60 age group patients have cholesterol, blood pressure, chance of heart attack higher than other groups.**

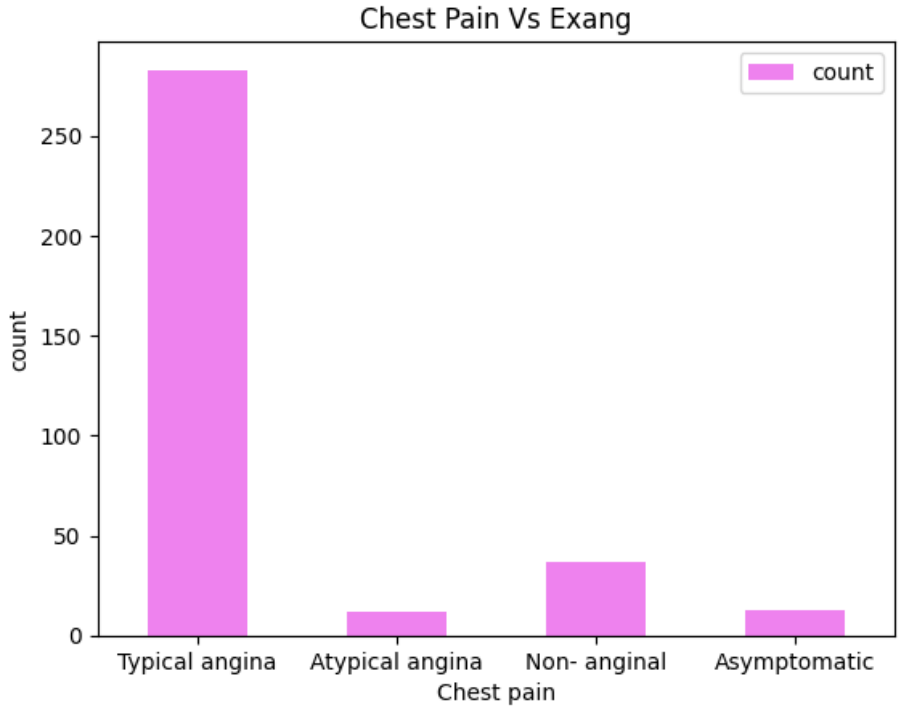
Exercise-induced angina Vs age, sex, chest pain

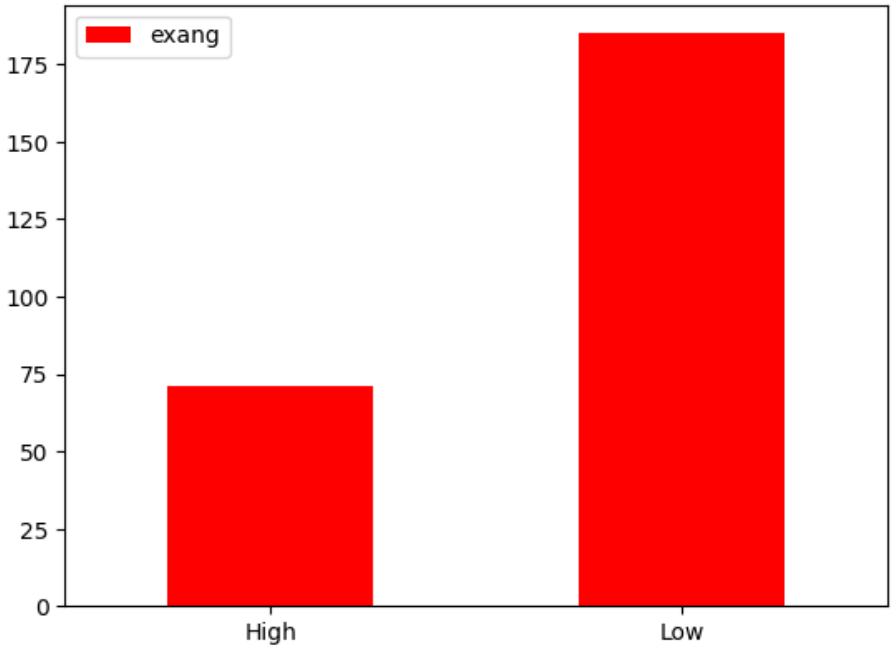


- Exercise-induced Angina is a pain in the chest that comes with exercise, stress, or other things that make the heart work harder.
- Exercise-induced angina is commonly found in the 50-60 age group. And the second highest is in 60-70 group.

- Exercise-induced angina is most commonly found in male heart disease patients.







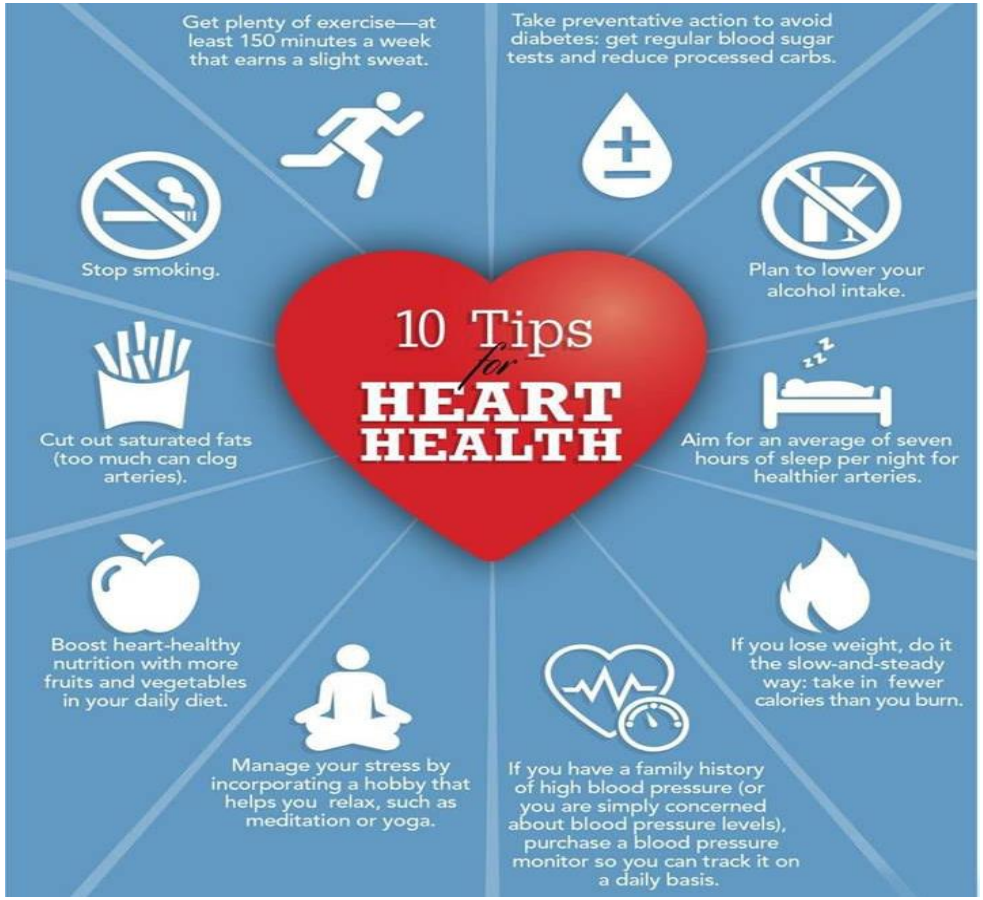
- **Patients having Exercise-induced angina are having Typical angina in more than 80% of cases.**
- **From the second plot, patients with Exercise-induced angina are studied against the chance of heart attack. From this Exercise-induced angina has a lower chance of resulting in a Heart attack.**
- **One might worry that exercising could trigger the symptoms or cause a heart attack, but the risk is low if you.**

Conclusion

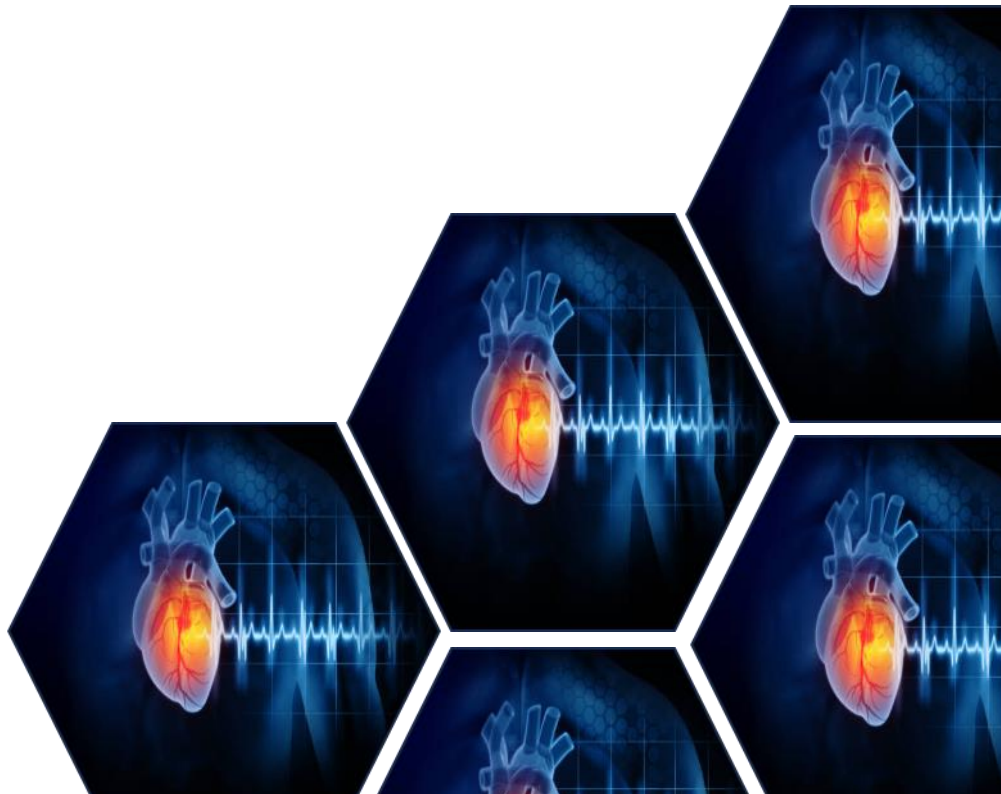
This study involved the analysis of Heart disease patients' data set with proper data processing. Here, I analyzed various factors and their effects which are suspectable in the case of heart diseases and heart attacks.

- **The age group 50-60 has the highest count of people having a high chance of heart attack.**
- **More than 83% of individuals with heart disease display cholesterol levels exceeding the desirable limit.**
- **More than 85% of people inspected for heart disease have Fasting blood sugar.**
- **About 50% of overall patients investigated having Fasting blood sugar and high cholesterol have a higher chance of Heart attack.**
- **Typical angina is the most common chest pain among heart disease patients even though Non Anginal pain has a higher probability of resulting in a Heart attack.**
- **More than 60% of patients having S-T abnormality have a high chance of Heart attack.**
- **Patients having Exercise-induced angina are having Typical angina in more than 80% of cases and has a lower chance of resulting in a Heart attack.**

To reduce the risk of heart disease, focus on regular exercise, healthy eating, and stress management. Quitting smoking, limiting alcohol, and maintaining a healthy weight are also vital. Regular medical check-ups, including blood pressure and cholesterol screenings, along with adherence to prescribed medications, are crucial for maintaining heart health. These lifestyle changes and proactive medical care can significantly reduce the risk of heart disease.



HEART DISEASE MORTALITY RATE DATA ANALYSIS



INTRODUCTION

In this data analysis, we aim to comprehensively examine the heart disease mortality rate in the United States in the year 2015. The dataset contains detailed information at the county, state, and national levels, offering a rich source of insights into the prevalence and impact of heart disease across different geographic regions within the country. By delving into this dataset, we can explore the geographical distribution, trends, and potential influencing factors associated with heart disease mortality, thereby paving the way for informed strategies and interventions to address this significant public health concern. We explore the data to gain insights about the relationships between Gender and Race/Ethnicity in the mortality rate.

Data Analysis is carried out using Python and its data analysis libraries such as Pandas, NumPy, Matplotlib, Seaborn, osgeo, etc.

Contents on Data

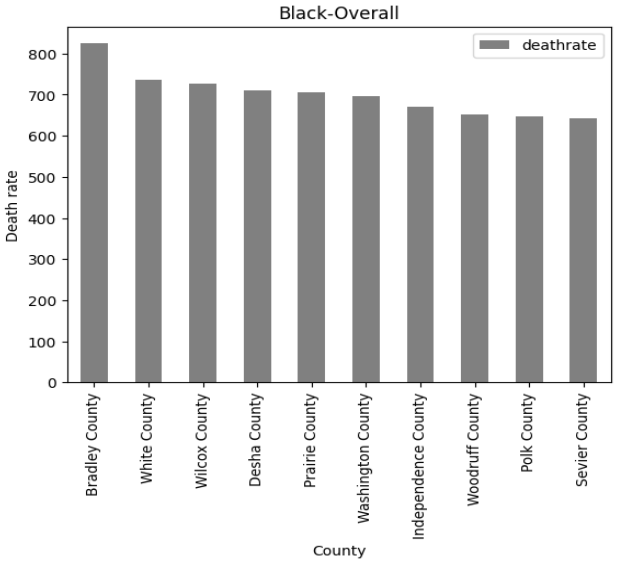
- LocationDesc (County, State, Nation)
- Geographical Level
- Data Value
- Data_Value_Unit
- Stratification Category 1(Gender)
- Stratification 1(Gender type)
- Stratification Category 2(Race/Ethnicity)
- Stratification 2(Race/Ethnicity type)
- Location(latitude and longitude)

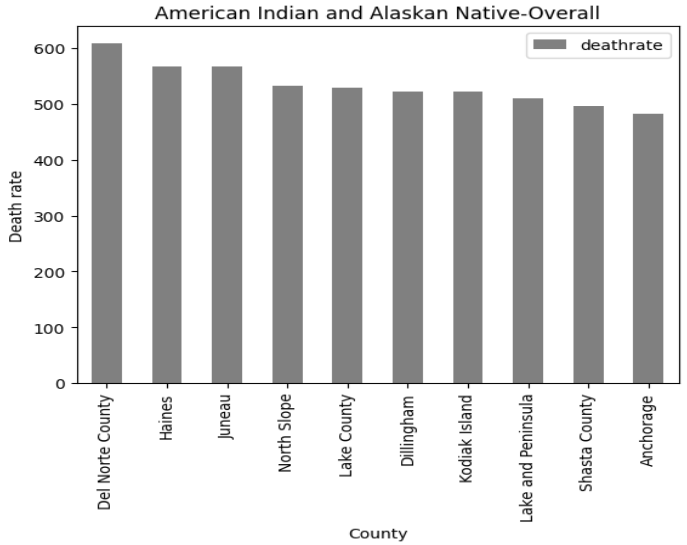
Analyze the data and give input about the county or state that has more heart disease and needs further inspection reports.

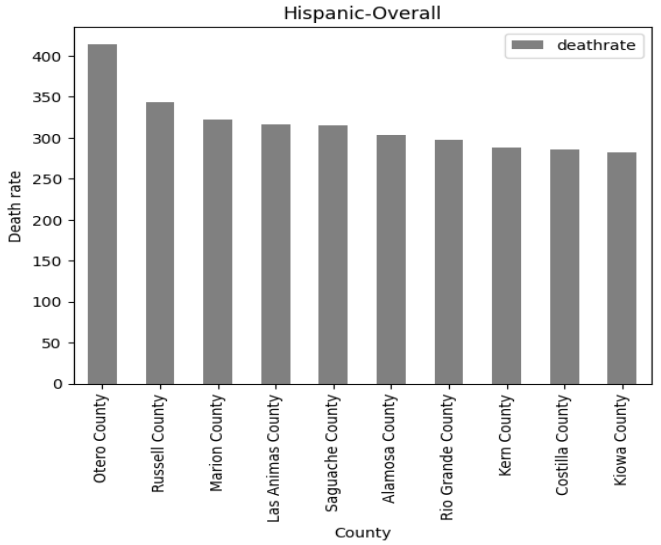
- Analysis of Counties with the most heart disease mortality rate. Analysis was done by inspecting the overall death mortality rate of counties for overall gender and race/ethnicity.

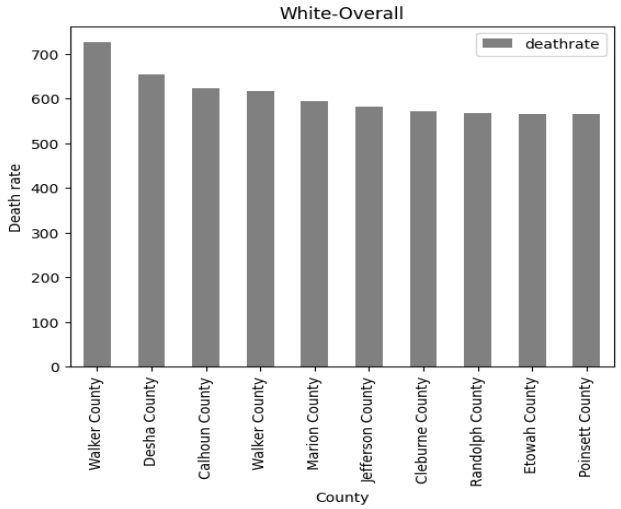
  - The county named Franklin Parish has the highest death mortality rate for overall gender and race among all counties with a death rate of 1170.5 per 1 Lakh per population.
  - Caldwell Parish, Washington County, Breathitt County, Desha County, Catahoula Parish, Wolfe County, Walker County, Marlboro County, and Tensas Parish are the top nine most death-rated counties after Franklin Parish.
  - More inspections to be done in these counties regarding lifestyles, living conditions, etc. to study the reasons for this higher mortality rate and to find effective methods to reduce those.

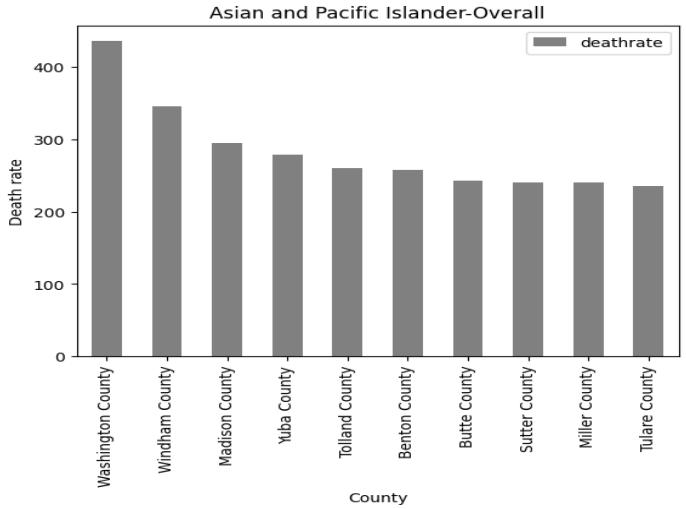
Most death-rated counties for overall gender for each race. Instructions are to be given particularly among the people.

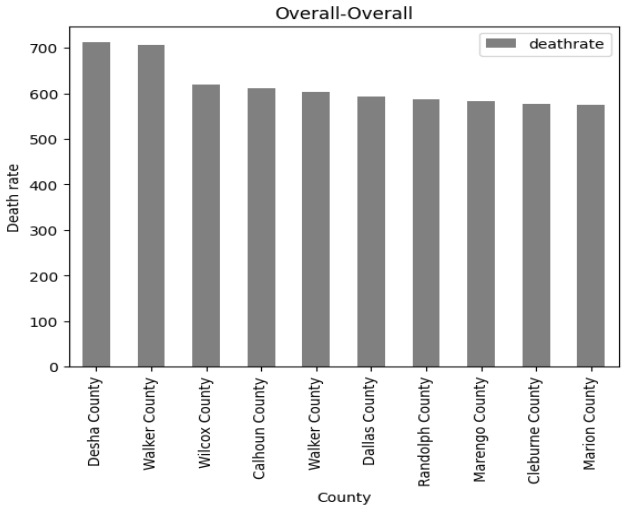




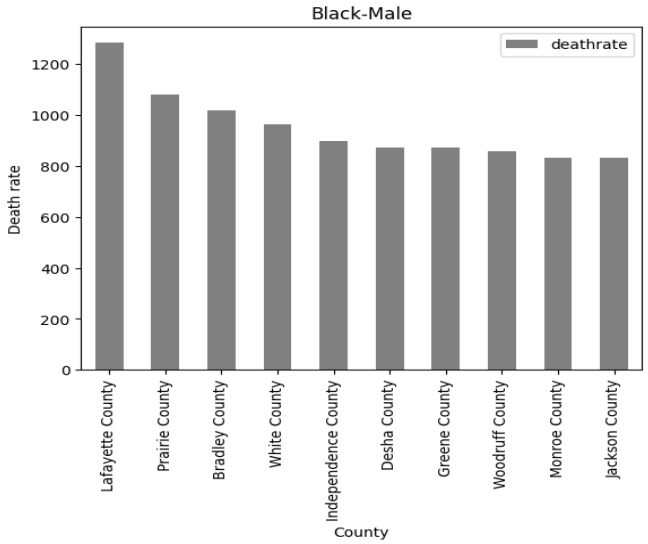


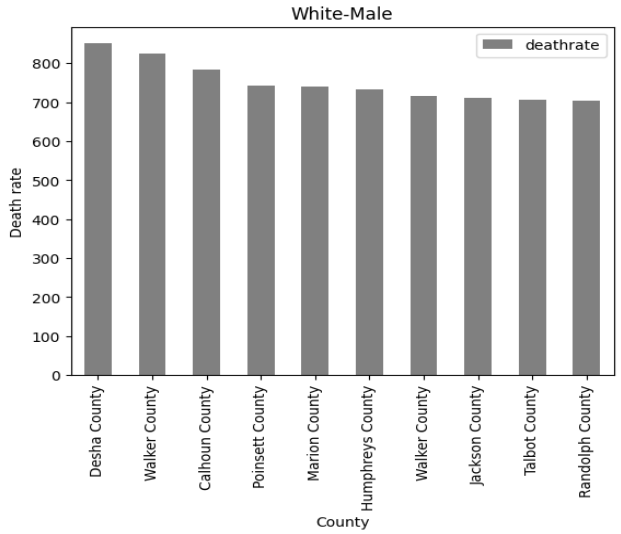


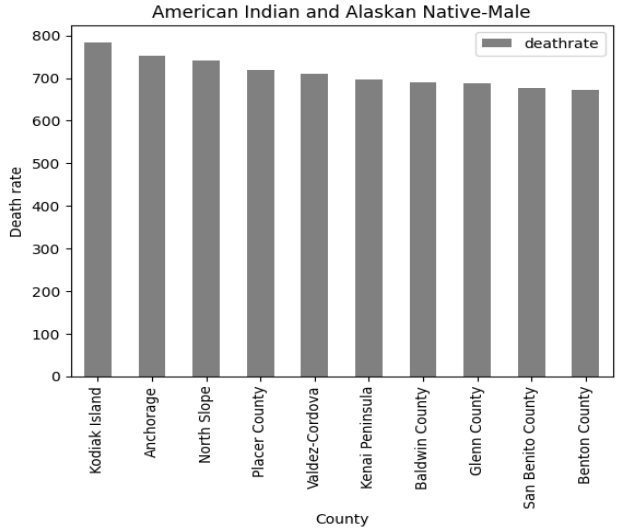


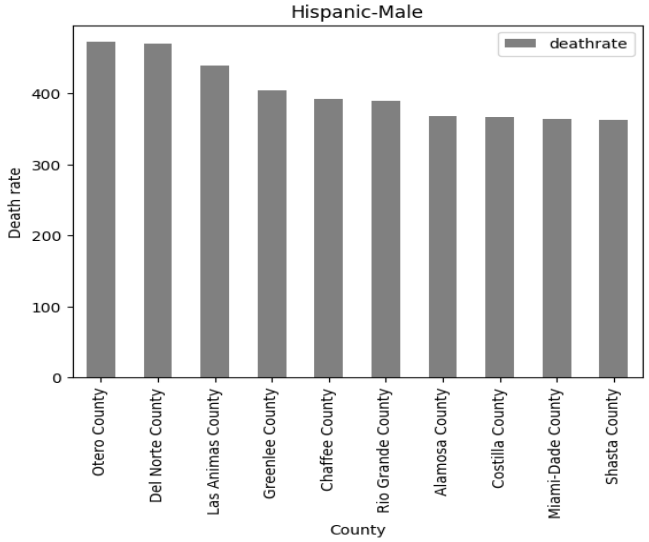


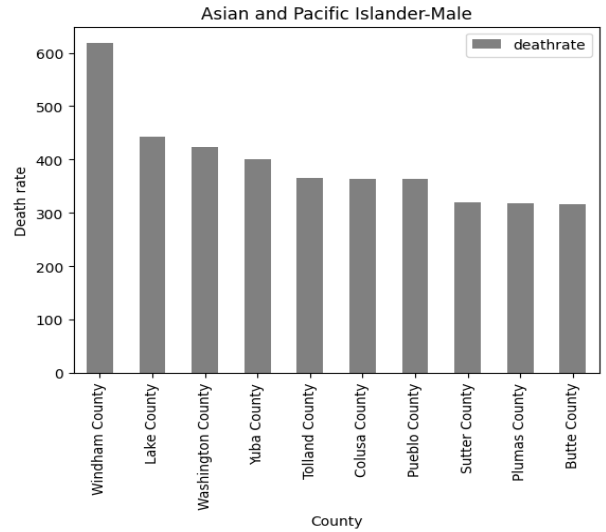
Most death-rated counties for Males for each race. Instructions are to be given particularly among the people.

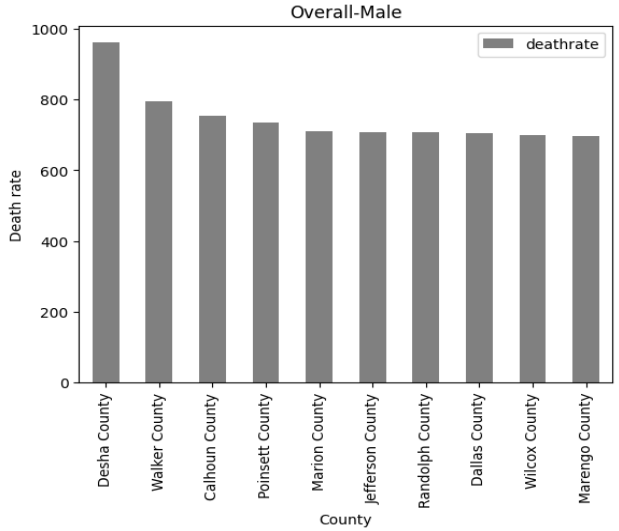




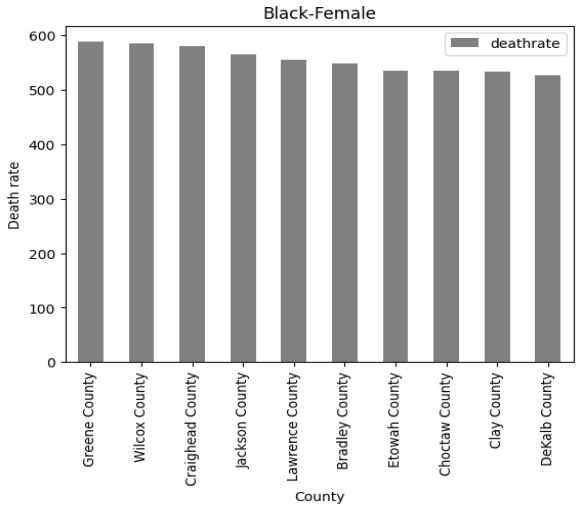


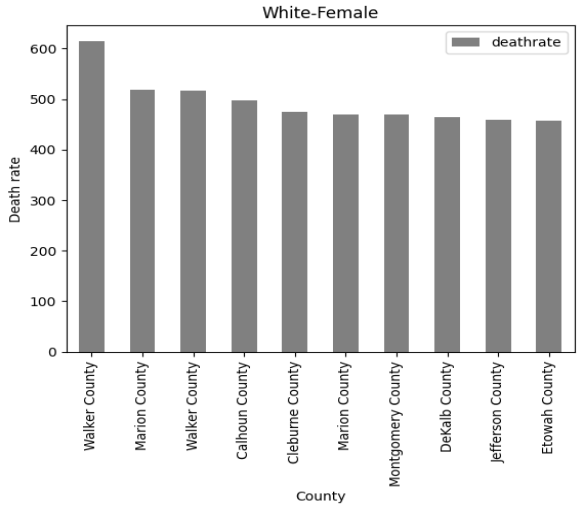


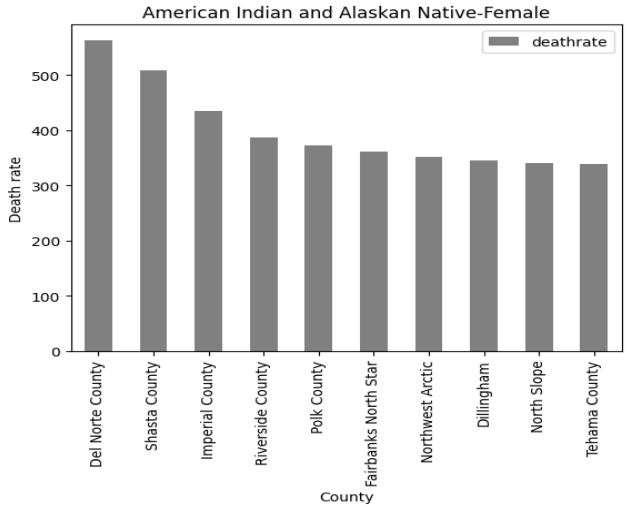


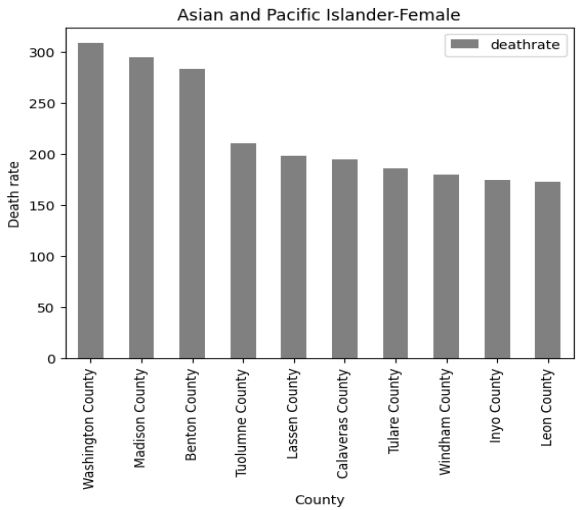


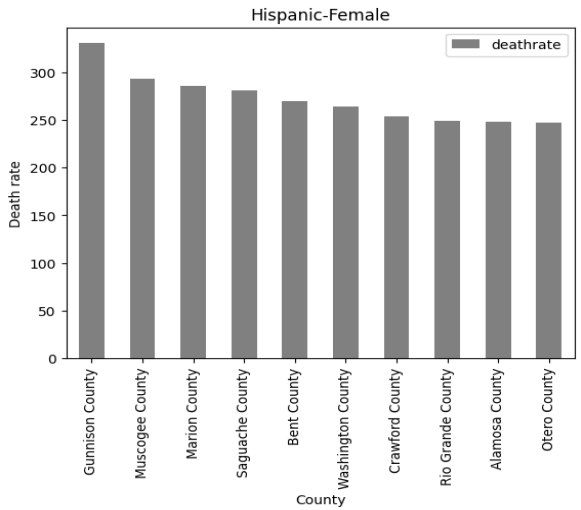
Most death-rated counties for Females for each race. Instructions are to be given particularly among the people.

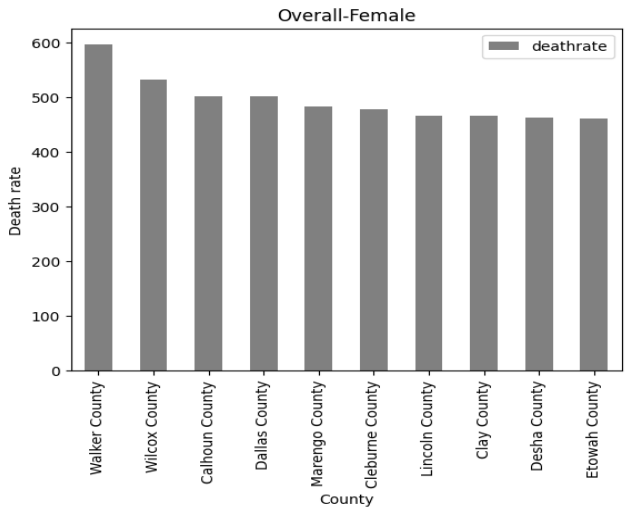












- **Analysis of States with the most heart disease mortality rate. Analysis was done by inspecting the overall death mortality rate of counties for overall gender and race/ethnicity.**

  - **The State named GUAM has the highest death mortality rate for overall gender and race among all counties with a death rate of 560.5 per 1 Lakh per population.**
  - **Alabama, Arkansas, District of Columbia, American Samoa, Indiana, and Georgia are the most Deathrate reported states after Guam.**
  - **More inspections to be done in these states regarding lifestyles, living conditions, etc. to study the reasons for this higher mortality rate and to find effective methods to reduce those.**

Most death-rated States for Males for each race.

Male-White	
Alabama	543.4
Arkansas	534.1
Louisiana	508.4
Indiana	452.9
Georgia	440.6
Illinois	421.2
Delaware	415.6
Iowa	402.6
California	400.5
Kansas	397.1

Male-Black	
District of Columbia	713.0
Arkansas	660.5
Louisiana	604.7
Alabama	585.2
Illinois	566.9
California	516.2
Indiana	511.8
Iowa	492.4
Georgia	491.4
Maryland	490.9

Male-Asian and Pacific Islander	
Hawaii	335.5
Arkansas	270.0
Idaho	238.7
Indiana	220.2
California	213.1
Iowa	209.6
Alaska	209.2
Illinois	202.0
Delaware	188.2
Colorado	183.1

Male-American Indian and Alaskan Native	
Kansas	706.0
Alaska	606.8
Idaho	422.7
California	411.0
Arizona	363.0
Colorado	290.4
Alabama	220.5
Florida	199.3
Illinois	196.7
Arkansas	150.9

Male-Hispanic	
Hawaii	523.6
Florida	312.5
Arizona	275.2
Colorado	274.9
California	271.0
Connecticut	241.5
Idaho	233.8
Kansas	222.7
Illinois	216.7
District of Columbia	207.0

Most death-rated States for Females for each race.

Female-White	
Alabama	350.4
Arkansas	337.5
Louisiana	312.4
Indiana	276.8
Georgia	276.3
Illinois	257.6
Delaware	248.6
Idaho	245.8
Iowa	243.8
California	242.7

Female-Black	
District of Columbia	408.8
Arkansas	393.3
Louisiana	390.9
Alabama	373.5
Illinois	350.3
California	335.0
Indiana	335.0
Iowa	321.7
Georgia	313.0
Kansas	294.5

Female-Asian and Pacific Islander	
Hawaii	172.4
Arkansas	167.6
Louisiana	143.9
Alaska	134.2
California	133.8
Georgia	130.7
Illinois	125.3
Arizona	117.4
Connecticut	116.8
Colorado	115.8

Female-American Indian and Alaskan Native	
Kansas	384.8
Alaska	320.1
California	256.0
Colorado	247.5
Arizona	181.6
Alabama	130.4
Florida	125.7

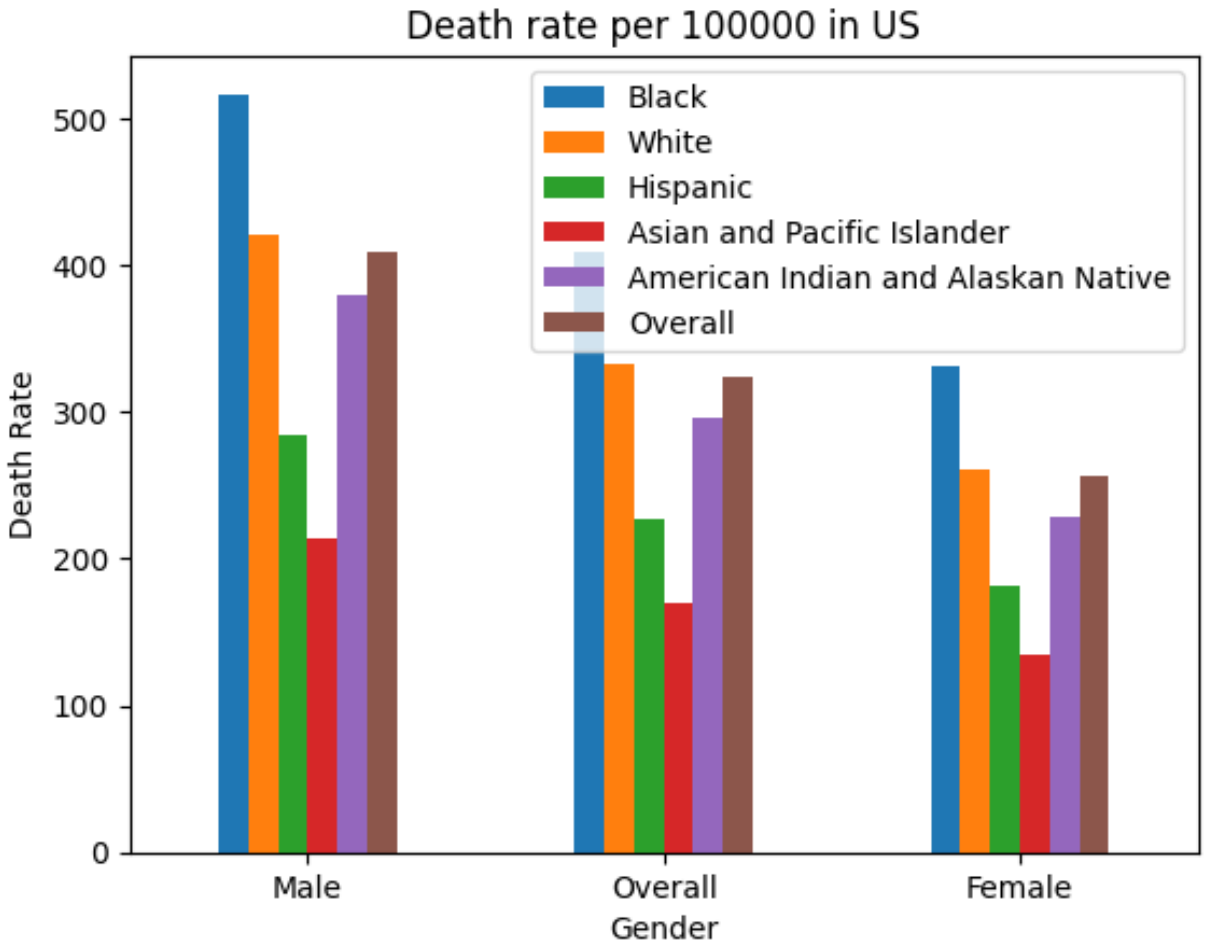
Female-Hispanic Hawaii	370.7
Florida	191.6
Arizona	174.2
California	173.2
Colorado	170.6
Connecticut	153.9
Kansas	147.5
Illinois	143.6
Idaho	139.5
Indiana	135.9

Analysis aspect of the medical advisor of the country.

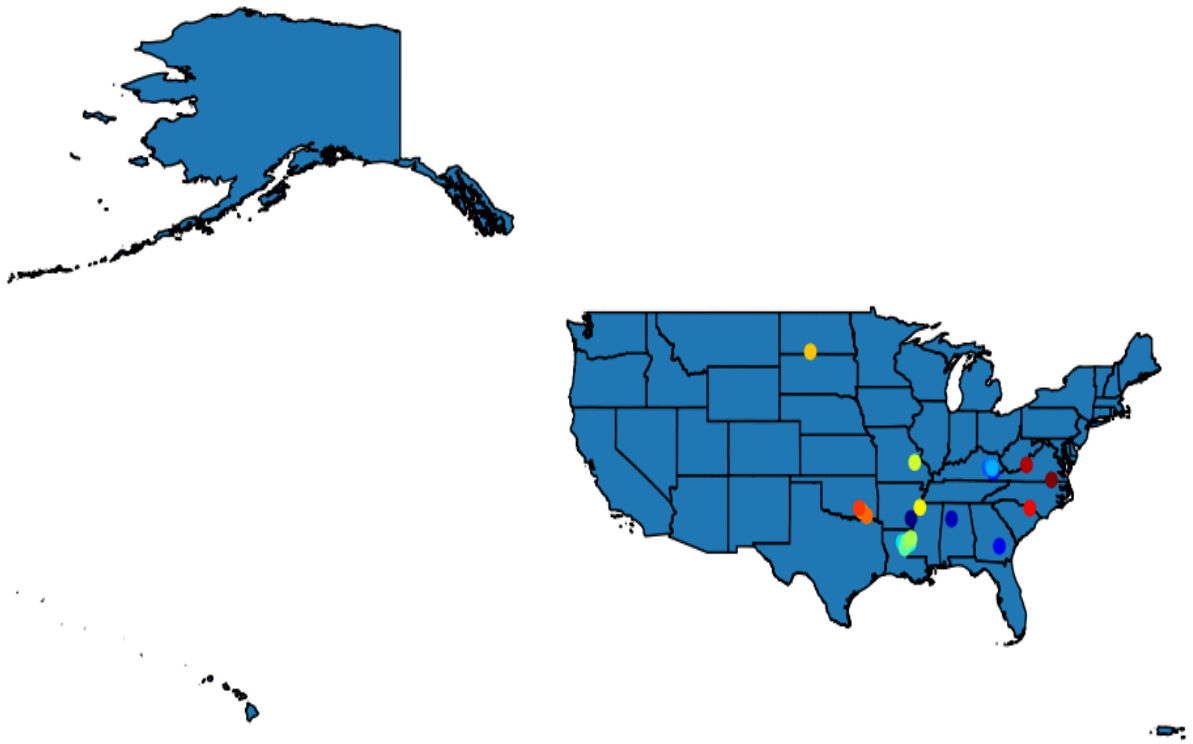
- Analysis of the mortality rate in the Counties or States we sort out the Counties and States which has the highest Death mortality rate. More inspections are to be done in these counties regarding lifestyles, living conditions, etc. to study the reasons for this higher mortality rate and to find effective methods to reduce those.
- Initiate a thorough analysis of health risk factors within the identified counties/states, including the prevalence of conditions such as obesity, diabetes, hypertension, and smoking.
- Evaluate the accessibility and quality of healthcare services, including primary care, preventive screenings, and cardiac care facilities, to identify potential gaps in healthcare provision.
- We did a detailed study of the Counties and States for each Gender-Ethnicity pair and found the Counties and States with a higher mortality rate. This will be helpful in the study of a particular part of society and in finding effective methods for making decisions accordingly. A separate study can be done within each category and can take further necessary actions.
- Study the socioeconomic and environmental factors influencing health, encompassing education levels, income disparities, food security, and environmental hazards.

US nation death rate of different races compared with gender.

- Overall Heart disease Death rate mortality rate in the United Nations is compared among genders.
- In every category of Race /Ethnicity, the Male category has a significantly higher mortality rate than Females.
- Black male has an enormous mortality rate which needs further study on the category.



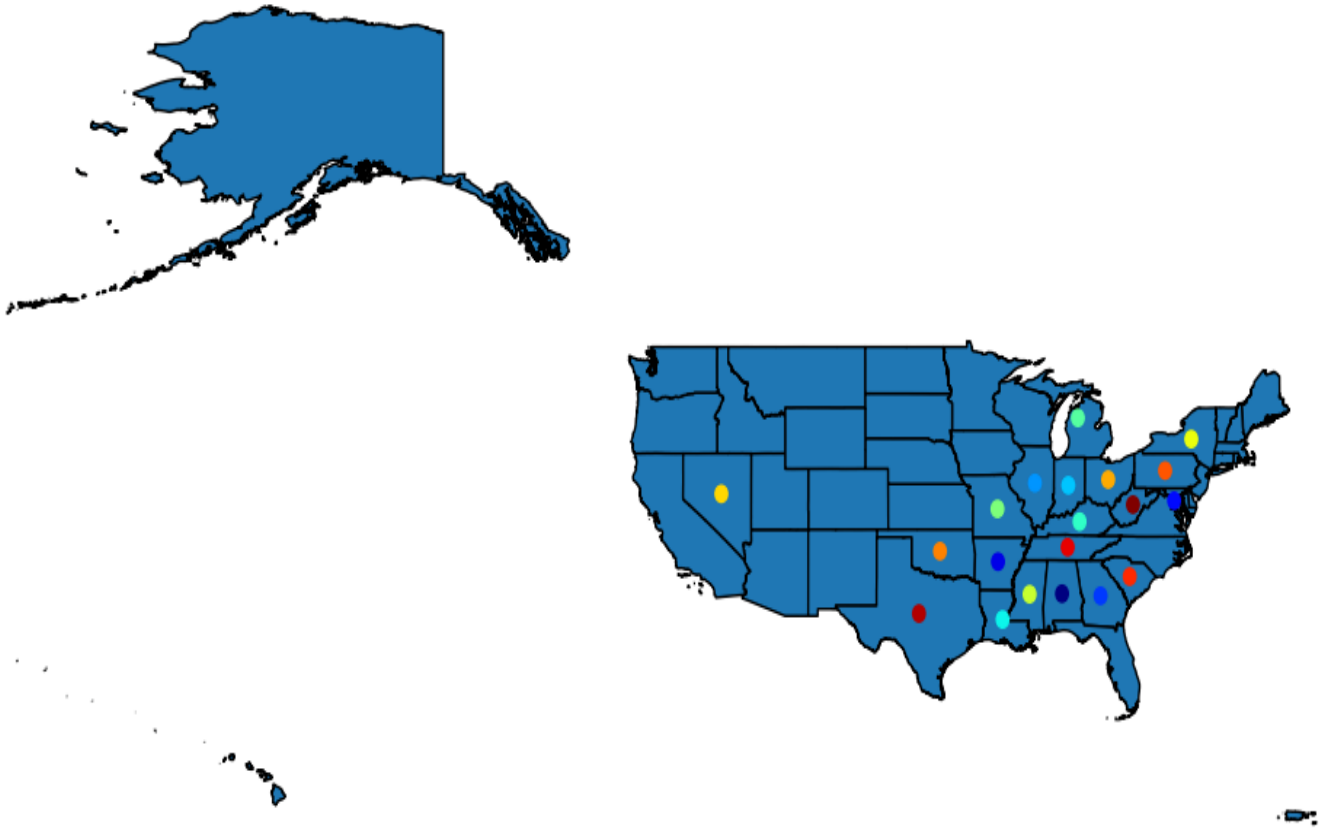
Analyzing the geographical location of the counties having a high mortality rate



- Including the shapefile of the United States we plot the counties having the highest Heart Disease mortality rate and plot using dots (as represented in figure).
- Most of the counties which have the highest mortality rate are found in the Southeast region of the United States.
- Location needs to be inspected further to find more relationships regarding the analysis.

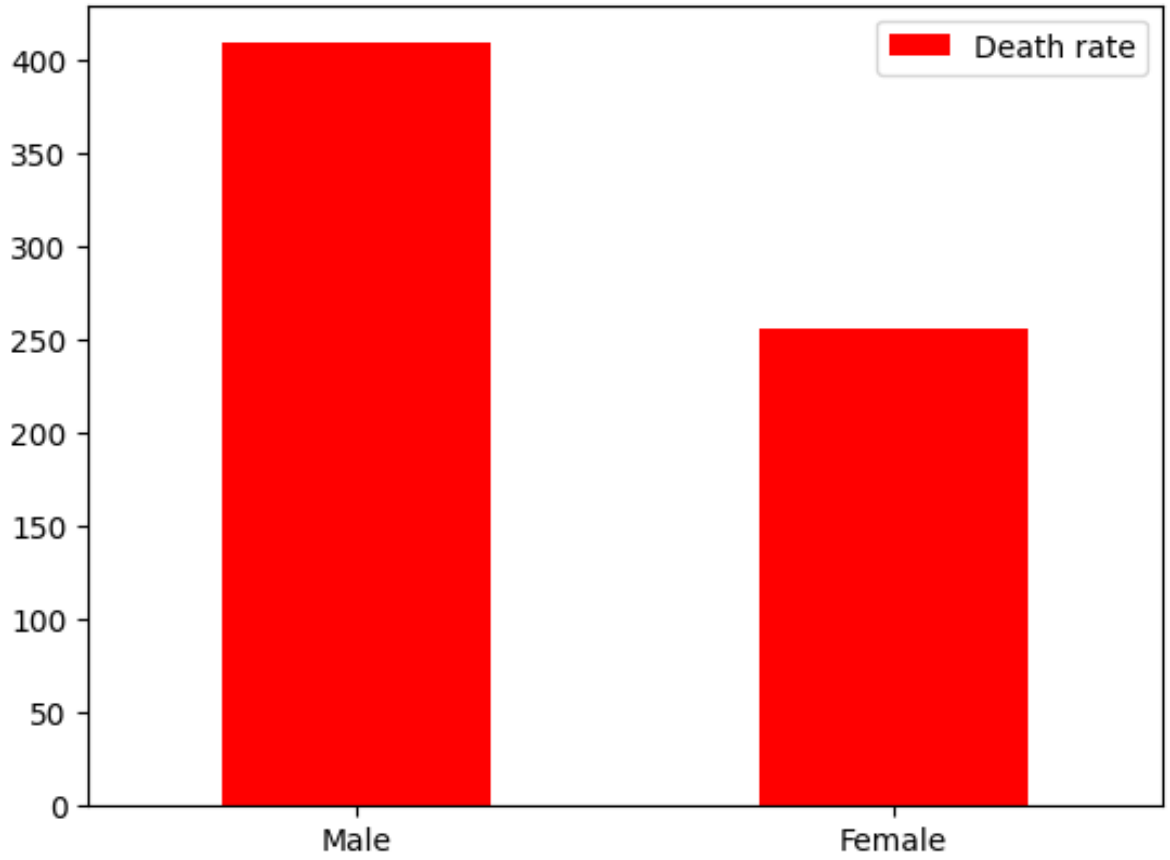
Analyzing the geographical location of the states having mortality rates higher than the national average

- Including the shapefile of the United States we plot the States having the Heart Disease mortality rate higher than the National average and plot using colored dots (as represented in the figure).
- Almost all the States with mortality rates higher than the National average are located in the East and Southeast parts of the United States.
- From the data we can infer that more studies should be done in these states about the living conditions of people, lifestyles, Medical care, etc., and about the Geographical conditions to recognize the reasons for the higher mortality rate .

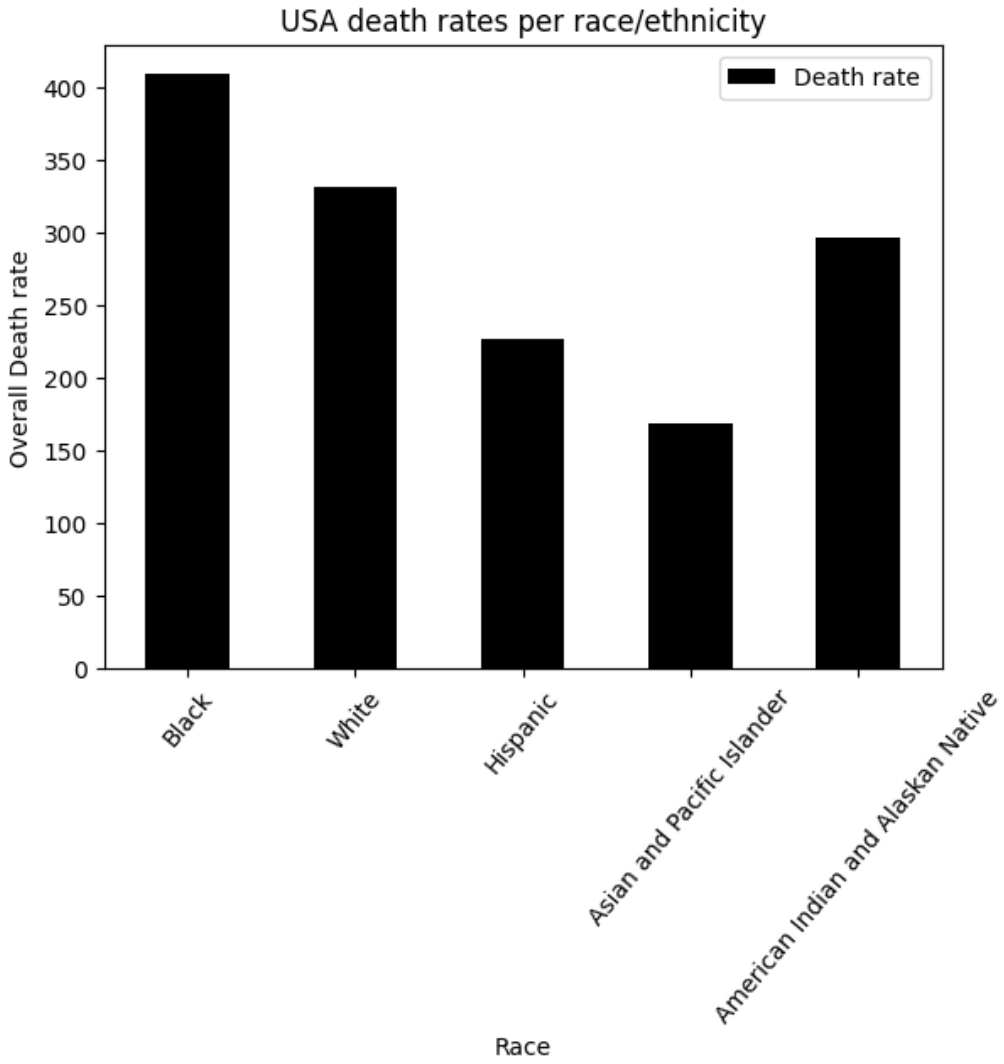


Heart disease mortality rate compared among genders

- Heart disease mortality rate of National level data compared among Gender.
- The mortality rate of Males is significantly higher than Females.
- Heart disease and death related to heart diseases are more common in males than females. Various biological and lifestyle factors contribute to this gender disparity, such as differences in hormone levels, genetic predispositions, and variations in risk behaviors including smoking and alcohol consumption.



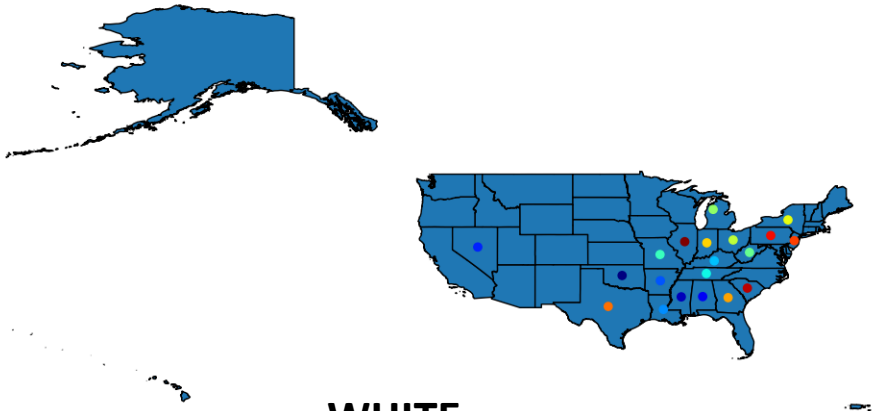
Heart disease mortality rate compared among race/ethnicity



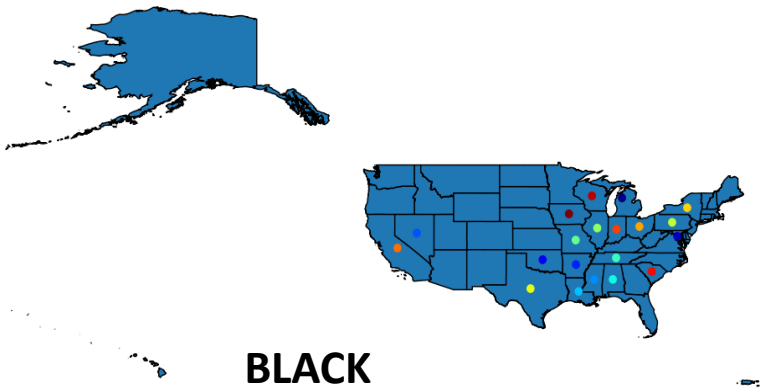
- **Heart disease mortality rate of National level data compared among Race/Ethnicity.**
- **Black people have the highest mortality rate among all Races for overall gender. White people are second highest and American Indian and Alaskan Natives are third.**
- **Black individuals often face disparities in socioeconomic status, education, and access to healthcare, which can lead to a higher prevalence of risk factors for heart disease such as hypertension, obesity, and diabetes.**
- **Genetic predispositions and variations in the expression of genes related to heart disease may also play a role in the increased mortality rate among Black populations.**

Analysis of States with the highest mortality rate for each race/ethnicity by geographic location

- States with high mortality rates for White people found in the Eastern region of the United States.
- Geographical features, lifestyles, etc. may be affecting factors. Also, the White people population may be much higher in these States.

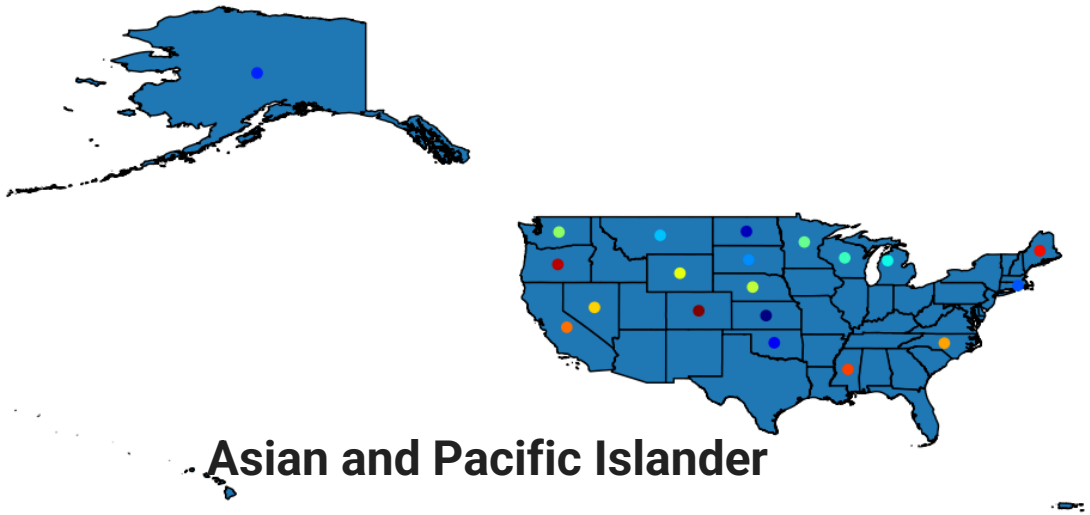


WHITE



BLACK

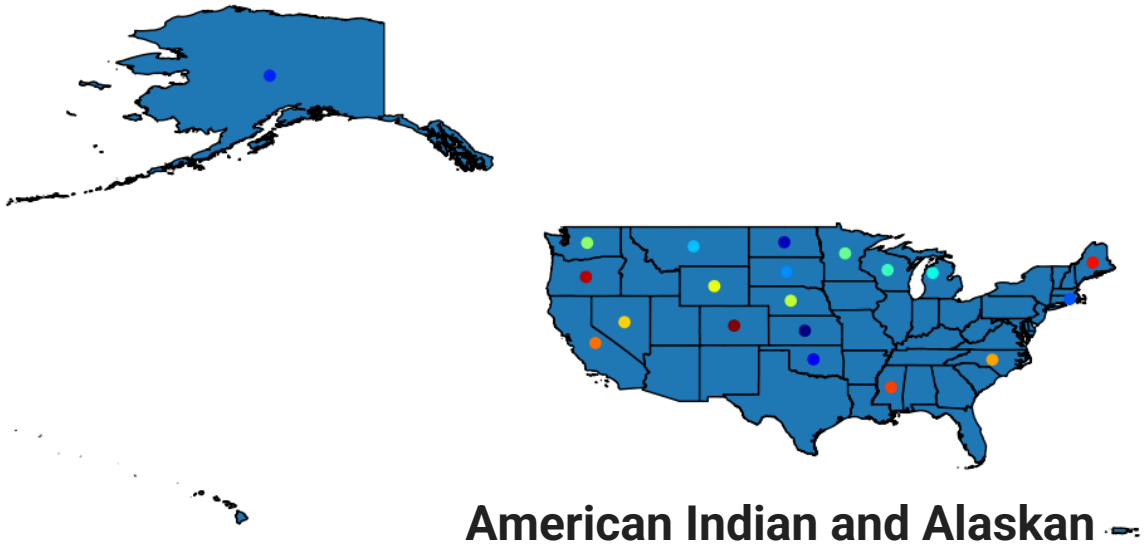
- States with high mortality rates for Black people also found in the Eastern region of the United States.
- Geographical features, lifestyles, etc. may be affecting factors. Also, the Black people population may be much higher in these States.



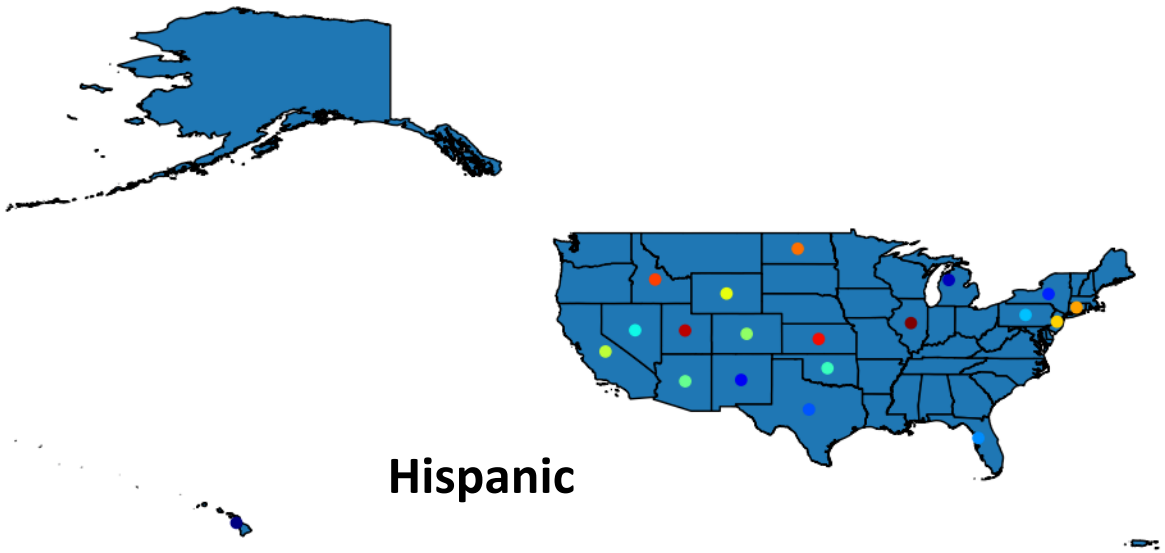
Asian and Pacific Islander

- States with high mortality rates for Asian and Pacific Islanders are depicted in the shapefile as colored dots.
- States are scattered in the Mid and Northern parts of the USA.

- States with high mortality rates for American Indian and Alaskan Native are depicted in the shapefile as colored dots.
- States are scattered in the Mid and Northern parts of the USA.



American Indian and Alaskan Native



Hispanic

- States with high mortality rates for Hispanic people found in the Western region of the United States.
- Geographical features, lifestyles, etc. may be affecting factors. Also, the Hispanic people population may be much higher in these States.

Conclusion

The analysis of heart disease mortality rate data in the USA in 2015 reveals a significant public health concern. The data underscores the urgent need for targeted interventions and public health policies to address the high incidence of heart disease-related deaths.

- **The Heart disease mortality rate of Males is significantly higher than Females in every race/ethnicity.** For **males at high risk of mortality, it's crucial to prioritize regular exercise, a healthy diet, and quitting smoking. Seek regular medical check-ups, monitor blood pressure and cholesterol levels, and adhere to prescribed medications to improve health outcomes and reduce the risk of premature death.**
- **Black people have the highest mortality rate among all Races for overall gender. White people are second highest and American Indian and Alaskan Natives are third.**
- **Further inspections to be done within the Black community to identify the valid reasons for their higher mortality rate**
- **States with high mortality rates for White people found in the Eastern region of the United States. Further studies should be done in these states about the living conditions of people, lifestyles, Medical care, etc., and about the Geographical conditions to recognize the reasons for the higher mortality rate.**

THANK YOU