Final Report - Healthcare Analytics

By

Shahana T

12219504

Analyse Health and Demographic Data to identify common traits leading to heart disease by Practo

School of Computer Science and Engineering

Lovely Professional University

Phagwara, Punjab (India)

October, 2023

Introduction

This report presents a comprehensive analysis of heart disease and its associated health metrics, based on a detailed exploratory data analysis (EDA) and Data visualization techniques conducted on a dataset containing various health indicators. The primary objective of this analysis is to identify key factors associated with heart disease, provide actionable insights, and recommend further research or health strategies.

We explore the distribution of individual variables using histograms, bar charts, pie charts and the relationships between various health indicators and heart disease. This includes comparing the distribution of key variables between individuals with and without heart disease and visualizing correlations between continuous variables using scatter plots and correlation matrices.

Furthermore, we employ various visualization techniques to provide a deeper understanding of the data which includes categorical variables, continuous variables. We plot scatter plots to investigate relationships between pairs of variables, focusing on those potentially associated with heart disease.

Advanced visualization techniques were used to display trends over time or across ordered categories, to show the distribution of data, and to plot pairwise relationships across multiple variables.

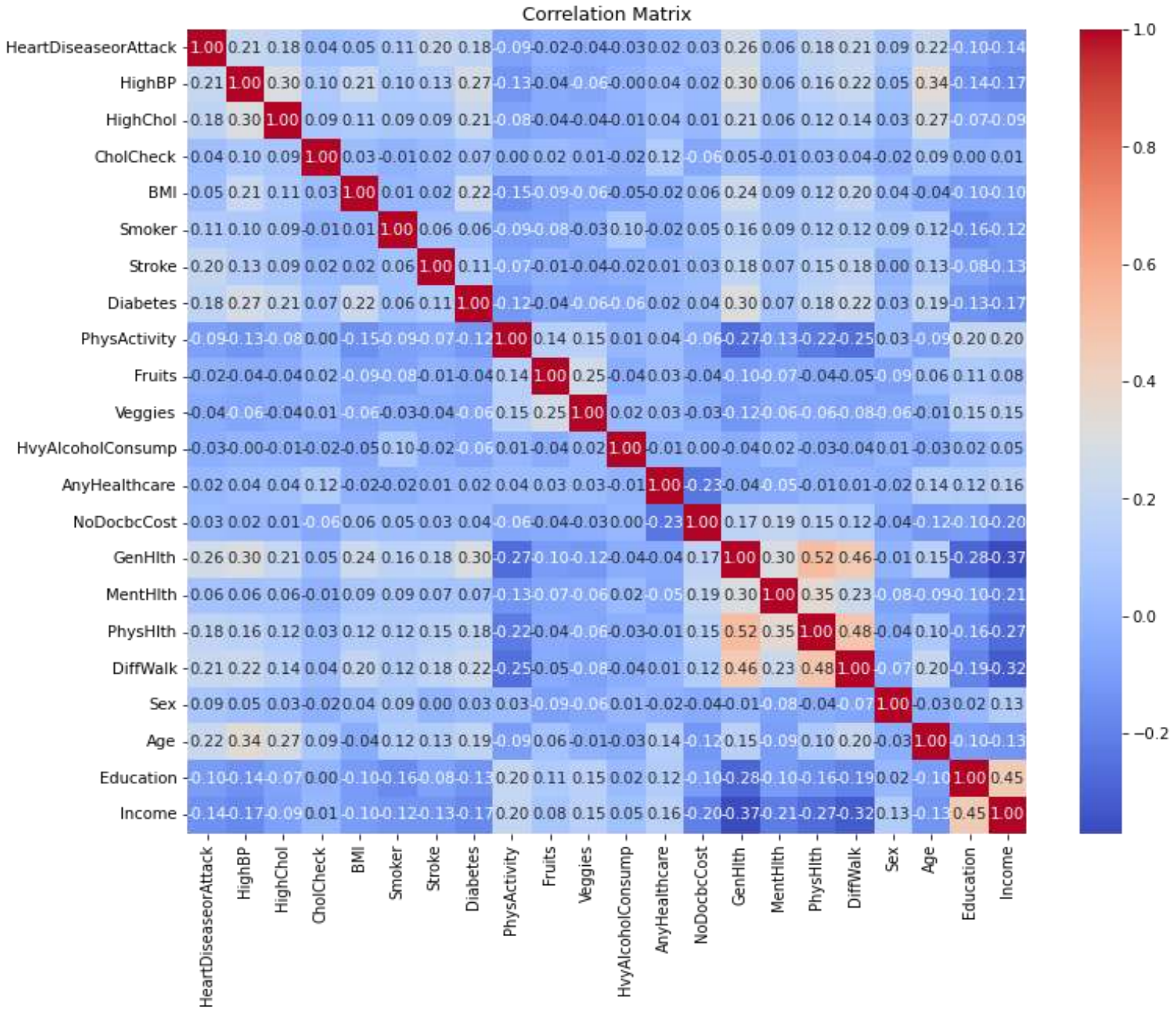
Through this multifaceted analysis, we aim to uncover significant patterns and associations that can inform targeted health interventions and strategies to mitigate the risk of heart disease.

Factors Associated with Heart Disease

To identify and discuss the factors most strongly associated with heart disease, we can perform a correlation analysis and visualize the results using a heatmap.

Correlation matrix represented using a Heatmap

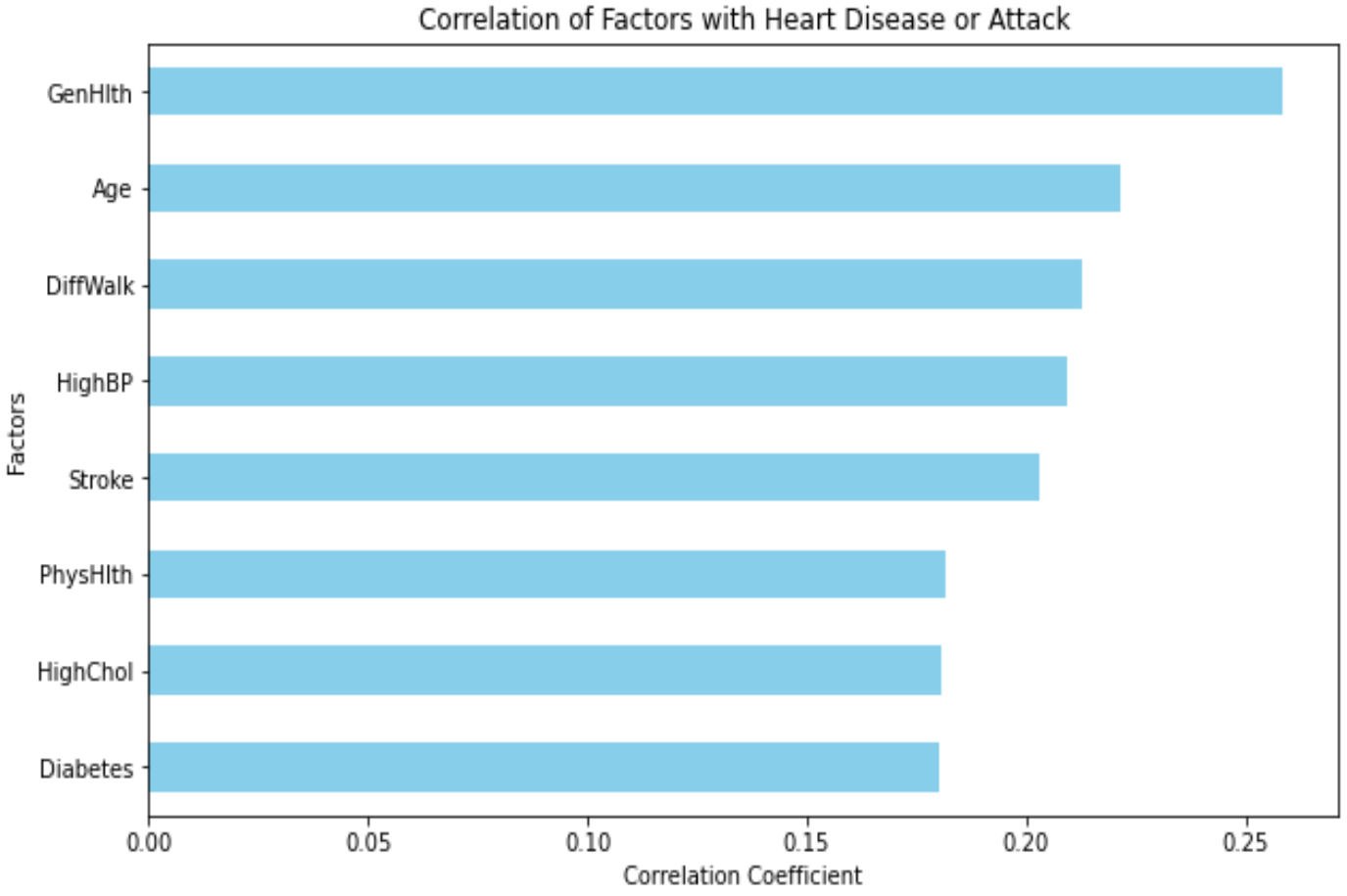
The heatmap visually represents the correlation coefficients between "HeartDiseaseorAttack" and other variables. Variables with warmer colours (closer to red) indicate stronger positive correlations, while cooler colours (closer to blue) indicate negative correlations or weaker associations



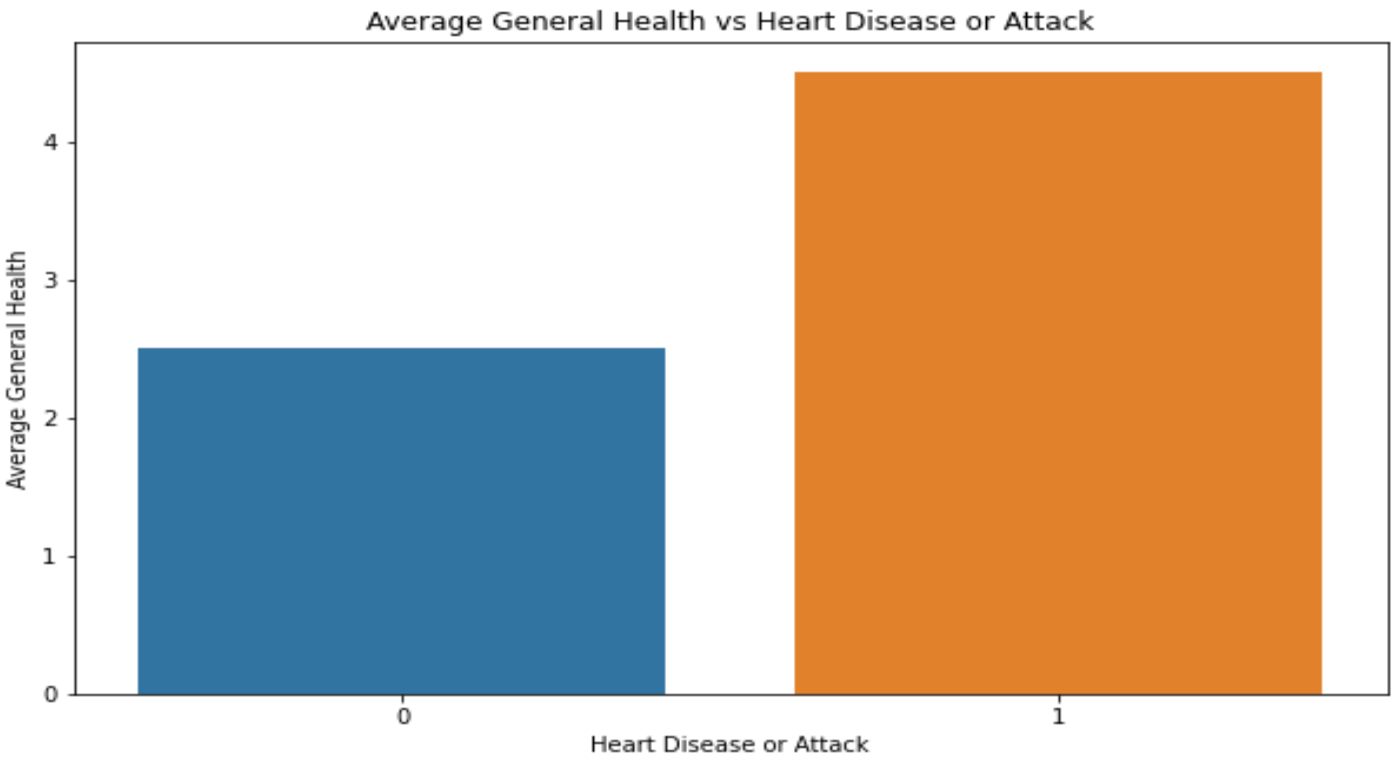
```
HeartDiseaseorAttack    1.000000
GenHlth                 0.258383
Age                     0.221618
DiffWalk                0.212709
HighBP                  0.209361
Stroke                  0.203002
PhysHlth                0.181698
HighChol                0.180765
Diabetes                0.180272
Smoker                  0.114441
Sex                     0.086096
MentHlth                0.064621
BMI                     0.052904
CholCheck               0.044206
NoDocbcCost             0.031000
AnyHealthcare           0.018734
Fruits                 -0.019790
HvyAlcoholConsump      -0.028991
Veggies                -0.039167
PhysActivity           -0.087299
Education              -0.099600
Income                 -0.141011
Name: HeartDiseaseorAttack, dtype: float64
```

When variables with positive correlation increase or change, there tends to be a higher likelihood or risk of heart disease or heart attacks and factors with negative correlation, where higher values in these variables tend to correlate with a lower risk or likelihood of heart disease or heart attacks.

Based on the correlation analysis, several factors were found to significantly influence the likelihood of heart disease or heart attacks are mentioned below:



1.1 General Health (GenHlth) - 0.258



Observation: This factor showed the strongest positive correlation (approximately 0.258) with "HeartDiseaseorAttack".

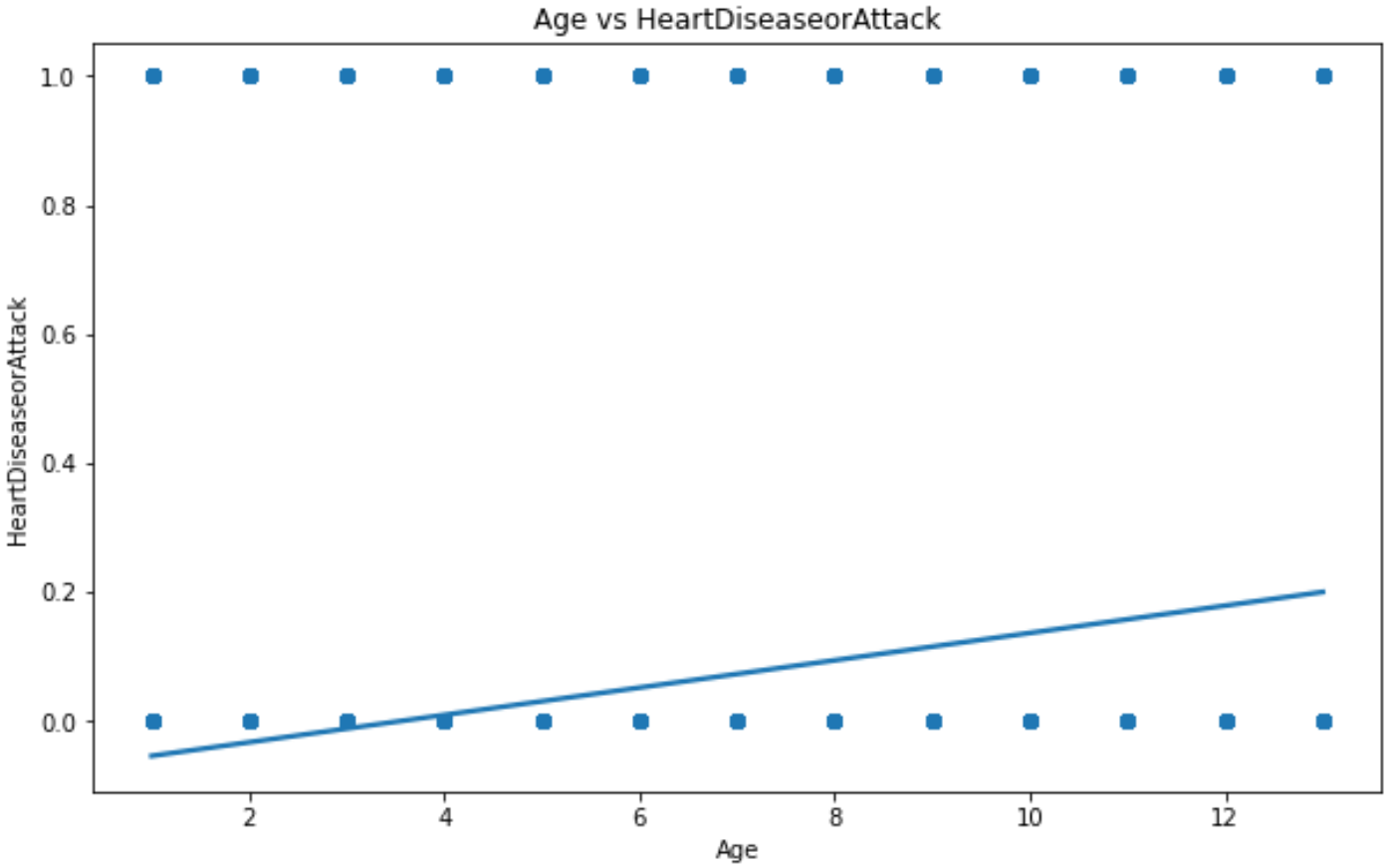
Data Representation:

- The x-axis represents “Heart Disease or Attack,” with two categories: 0 (no heart disease or attack) and 1 (heart disease or attack).
- The y-axis represents “Average General Health” and ranges from 1 to 4.

Remarks:

- Those with heart conditions report higher average general health, which may seem counterintuitive.
- Individuals reporting poorer general health were more likely to have heart disease.

1.2 Age - 0.222

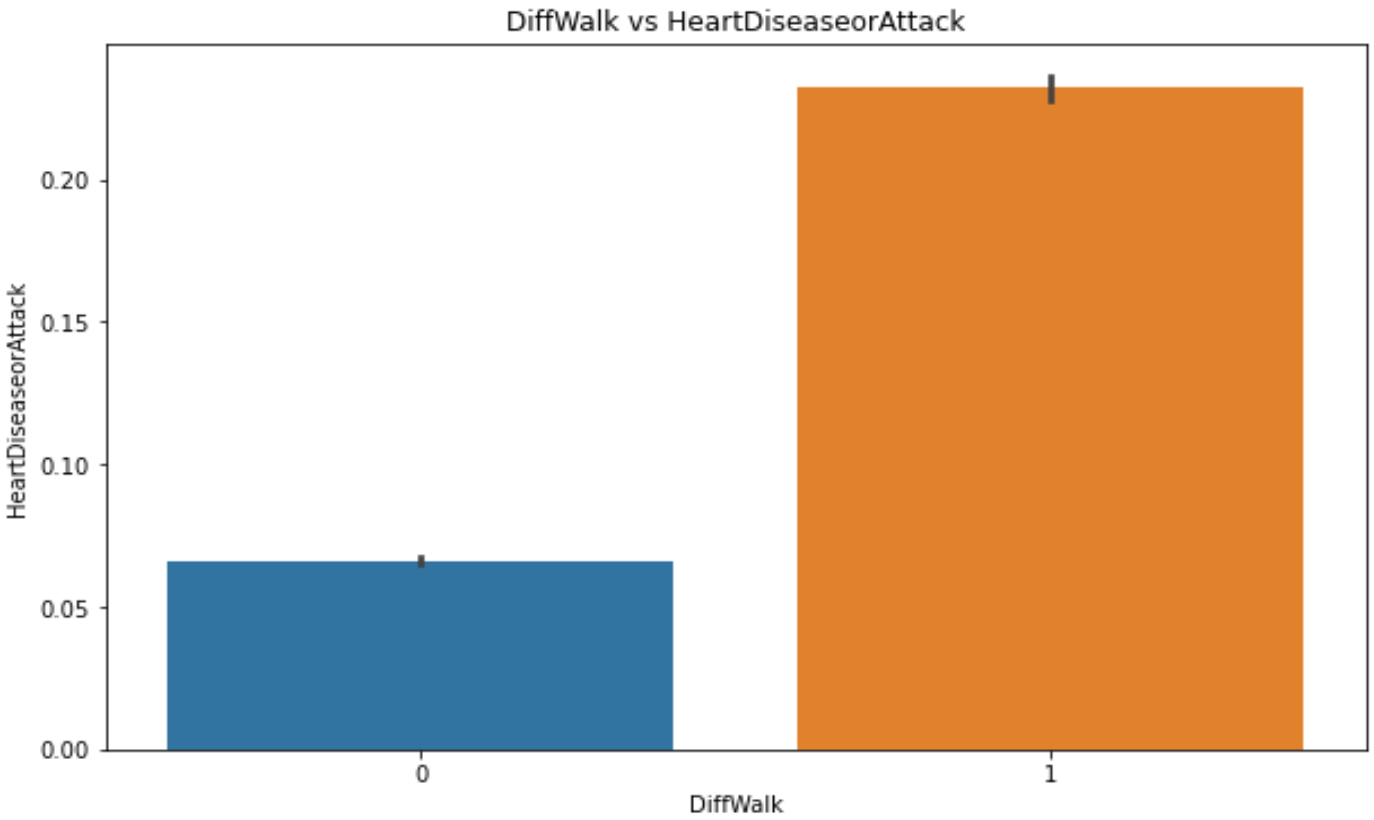


Observation: As age increases, there appears to be an increase in heart disease or attack incidents.

Trend Line: A trend line has been added which indicates a general upward trend, suggesting that as one ages, the likelihood of having heart disease or an attack increases.

Data Distribution: Most data points are clustered at the lower end of both axes, indicating that younger ages have lower incidences of heart diseases or attacks.

1.3 Difficulty Walking (DiffWalk) - 0.213



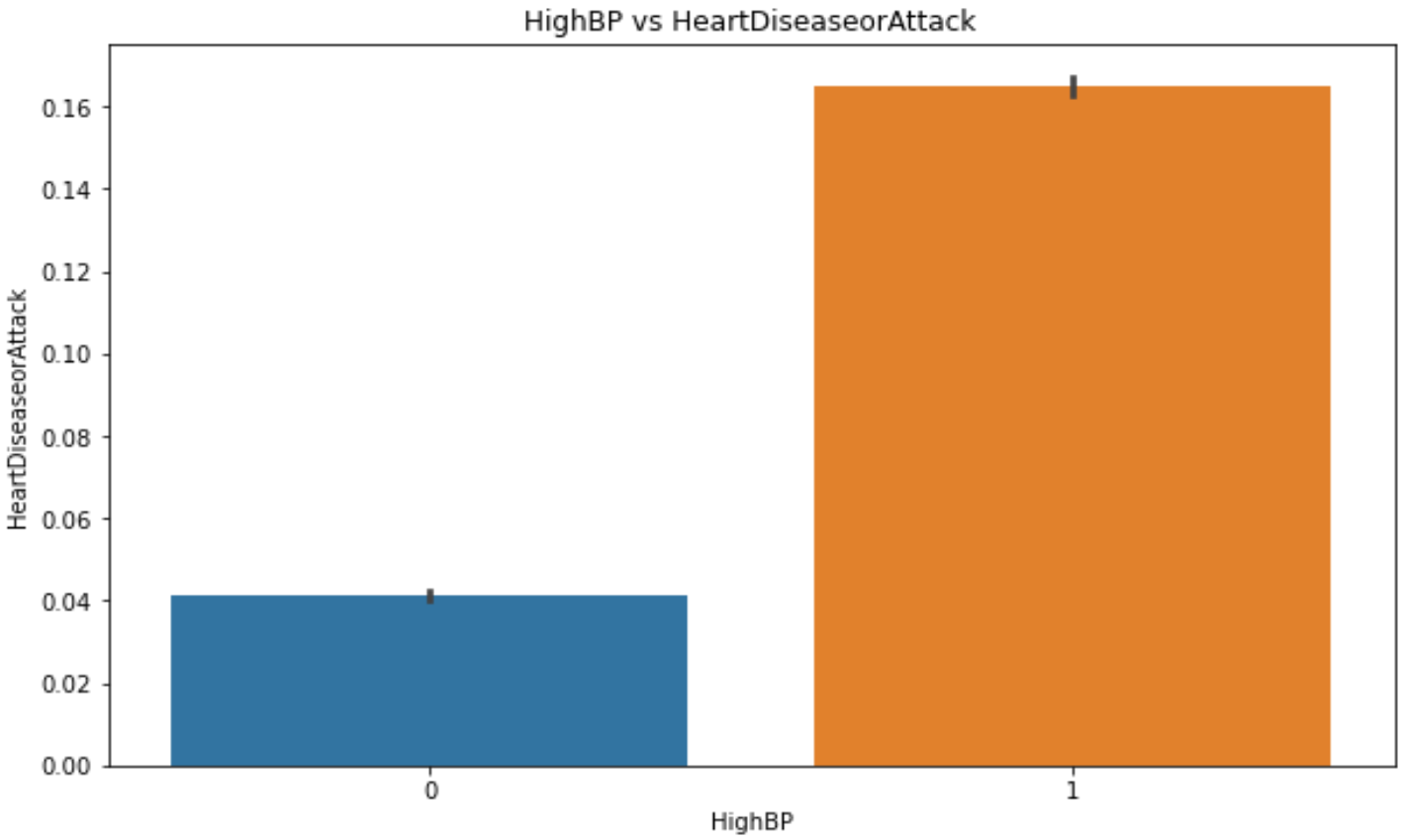
Observation:

- Group ‘0’ represents individuals with less difficulty in walking, while group ‘1’ likely includes those with more difficulty.
- The y-axis (labelled “HeartDiseaseorAttack”) shows values ranging from 0.00 to 0.20.

Remarks:

- The blue bar (group ‘0’) has a significantly lower value on the y-axis, suggesting a lower association with heart disease or attacks.
- The orange bar (group ‘1’) indicates a higher association with heart disease or attacks.
- Individuals in group ‘1’ (more difficulty walking) are at a higher risk for heart-related health issues.

1.4 High Blood Pressure (HighBP) - 0.209



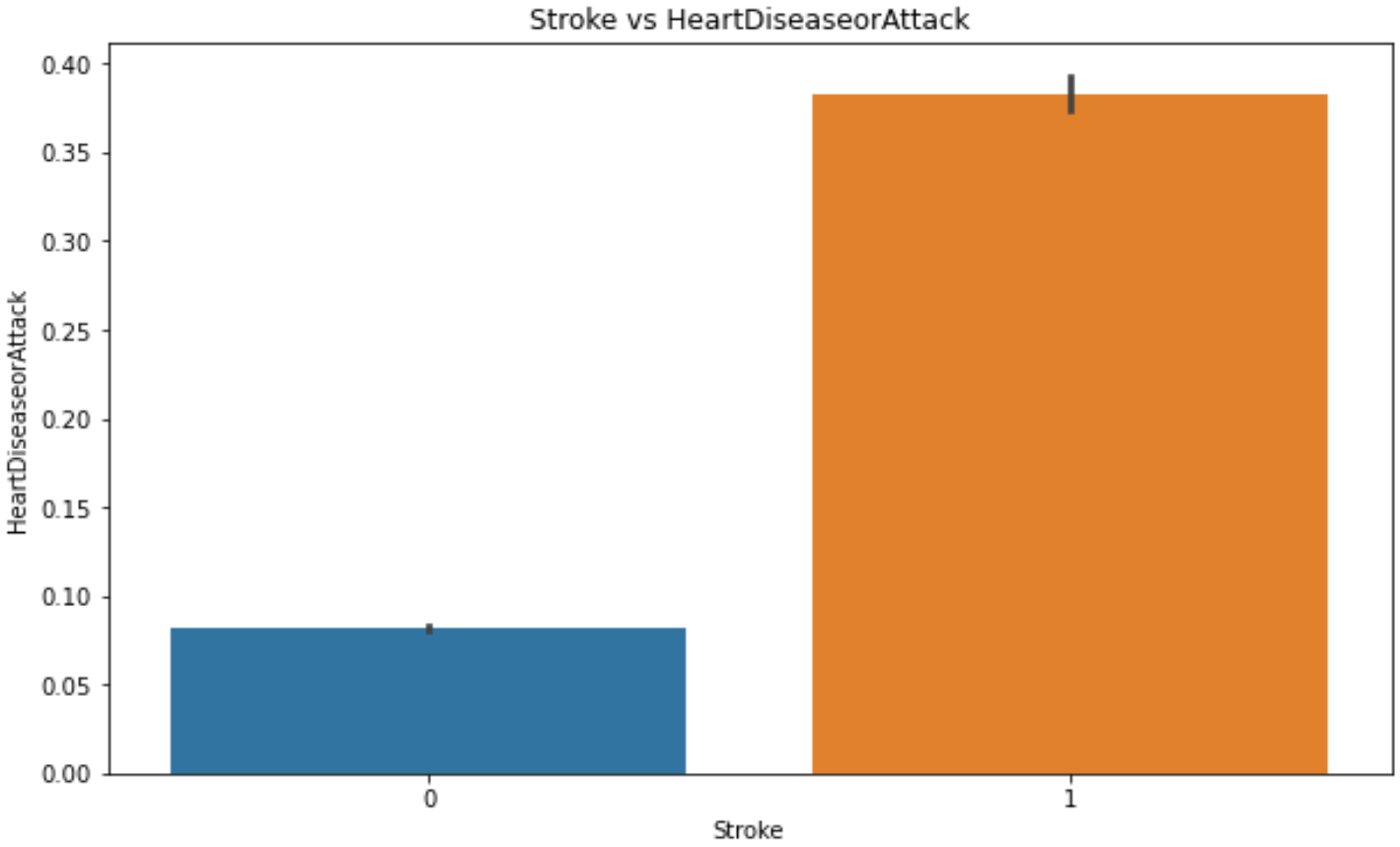
Observation:

- Individuals without high blood pressure (HighBP = 0) have a lower proportion of heart disease or attacks.
- Individuals with high blood pressure (HighBP = 1) have a significantly higher proportion of heart disease or attacks.

Remarks:

- The visualization suggests a strong correlation between high blood pressure and the occurrence of heart disease or attacks.
- High blood pressure can cause damage to arteries, leading to coronary artery disease, a risk factor for heart attacks.
- Lifestyle factors (diet, exercise, stress) and other health conditions (diabetes, obesity) contribute to both high blood pressure and heart diseases.

1.5 Stroke - 0.203



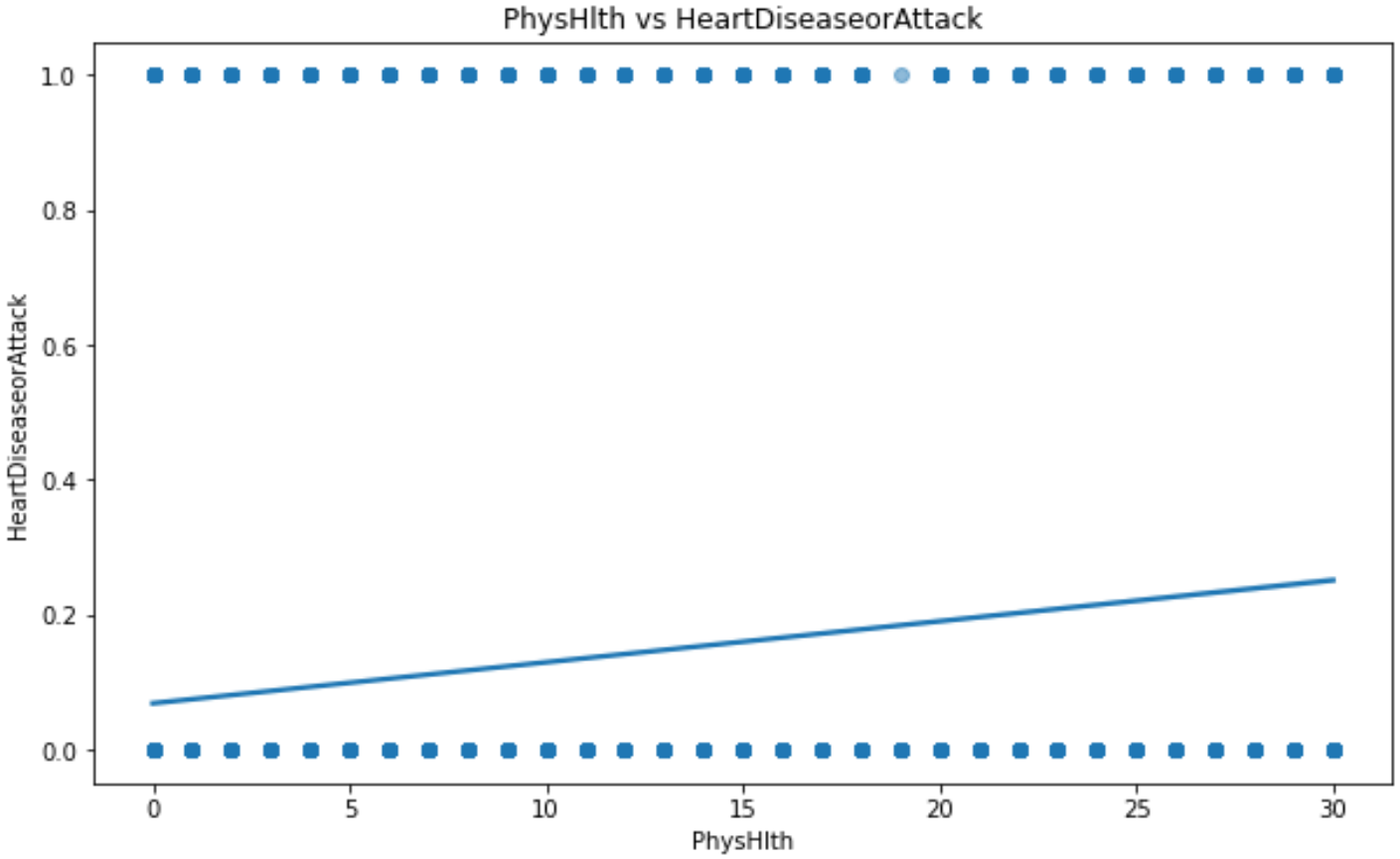
Observation:

- The visualization suggests that experiencing a stroke may be an important factor or indicator for heart disease or attacks.
- Individuals in category '1' (stroke) have a significantly higher rate of heart disease or attacks compared to those in category '0'.
- The height difference between the bars indicates a strong association between strokes and heart conditions.

Reasons for Heart Disease due to Strokes:

- Strokes can damage the cardiovascular system, increasing vulnerability to heart diseases.
- A history of strokes might reflect underlying health issues that contribute to heart conditions.

1.6 Physical Health (PhysHlth) - 0.182



Observation:

- As 'PhysHlth' increases, there appears to be an increase in heart disease or attack incidents.
- Most data points cluster near zero on the 'HeartDiseaseorAttack' scale, indicating that many individuals have low risk relative to their physical health status.

Remarks:

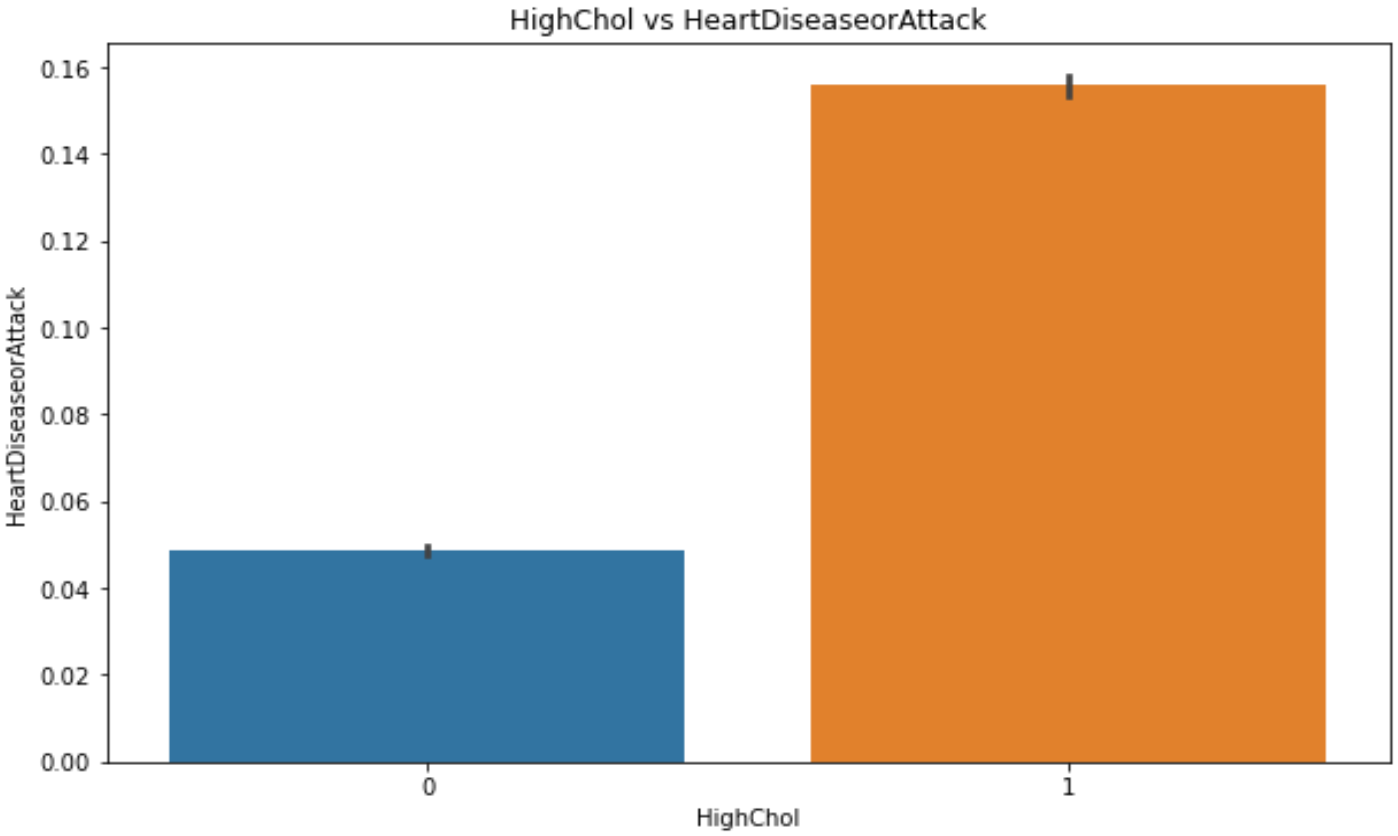
- The trend suggests that individuals with higher 'PhysHlth' values tend to have a higher incidence or risk of heart disease or attack.
- Poor physical health may lead to increased strain on the cardiovascular system, contributing to heart diseases.

Reasons for Heart Disease Due to Above Factor:

- Lifestyle factors: Smoking, unhealthy diet, and lack of exercise.

- Medical conditions: Hypertension, high cholesterol, and diabetes (reflected in poor physical health scores).

1.7 High Cholesterol (HighChol) - 0.181



Observation:

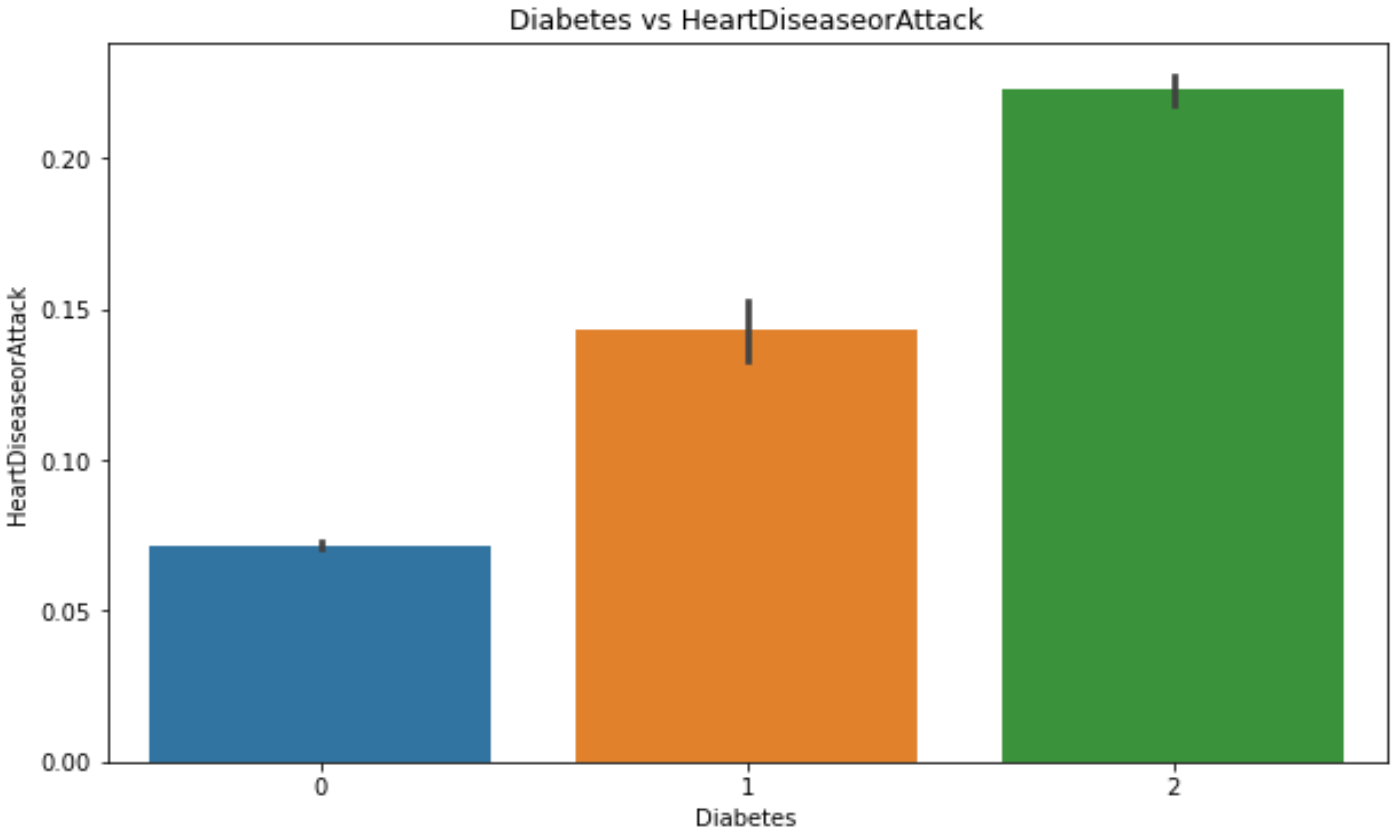
- This suggests that individuals with high cholesterol (‘1’) have a higher incidence of heart disease or attacks compared to those without high cholesterol (‘0’).
- The visualization highlights a clear difference in heart disease rates based on cholesterol levels.
- High cholesterol appears to be a significant factor in heart disease or attacks

Factors Responsible for High Cholesterol and Heart Disease:

- Diet: Consuming saturated fats and trans fats raises blood cholesterol levels.

- Physical Inactivity: Lack of exercise contributes to weight gain and higher cholesterol.

1.8 Diabetes - 0.180



Observation:

- The graph shows three vertical bars representing different categories of diabetes (0, 1, and 2).
- The height of each bar corresponds to the proportion of heart disease or attacks within that diabetes category.

Remarks:

- Category ‘0’ (blue bar) has the lowest proportion of heart disease or attacks.
- Category ‘1’ (orange bar) has approximately twice the proportion compared to category ‘0’.

- Category ‘2’ (green bar) has the highest proportion of heart disease or attacks.

Reasons for Heart Disease due to Diabetes:

- High blood sugar levels associated with diabetes can damage blood vessels and nerves, affecting the heart.
- Diabetes often coexists with other risk factors like high blood pressure and cholesterol levels.

Recommendations or Insights

Based on the findings from the analysis of heart disease and its associated health metrics, here are several recommendations and insights for further research, interventions, and health strategies:

General Health (GenHlth)

Insights:

- Increase public health initiatives focused on improving general health. This could include community health programs, regular health screenings, and promoting a balanced lifestyle.
- Those with poorer self-reported general health are more likely to have heart disease, indicating a need for targeted interventions for those who perceive their health as poor.

Age

Insights:

- Implement age-specific heart disease prevention strategies, particularly for older adults. Regular cardiovascular health check-ups should be promoted among older populations.

- The risk of heart disease increases with age, emphasizing the importance of early intervention and continuous monitoring of heart health as people age.

Difficulty Walking (DiffWalk)

Insights:

- Develop programs to improve mobility and physical activity among those with difficulty walking. Physiotherapy and structured physical activity programs can be beneficial.
- Difficulty walking is associated with a higher risk of heart disease, suggesting that interventions to improve mobility can also reduce heart disease risk.

High Blood Pressure (HighBP)

Insights:

- Enhance efforts to control blood pressure through lifestyle modifications and medication adherence programs. Public education on the importance of maintaining optimal blood pressure levels is crucial.
- High blood pressure significantly correlates with heart disease, highlighting the need for effective hypertension management strategies.

Stroke

Insights:

- Strengthen post-stroke care and rehabilitation programs to address cardiovascular health and prevent further heart disease. Regular cardiovascular monitoring for stroke survivors is essential.
- Stroke history is strongly associated with heart disease, indicating that comprehensive care for stroke patients should include heart disease prevention.

Physical Health (PhysHlth)

Insights:

- Promote physical activity and healthy living to improve overall physical health. Community-based fitness programs and accessible recreational facilities can encourage a more active lifestyle.
- Poor physical health is linked to higher heart disease risk, underscoring the need for public health campaigns promoting physical well-being.

High Cholesterol (HighChol)

Insights:

- Implement dietary counselling and cholesterol-lowering interventions, such as promoting heart-healthy diets and increasing access to cholesterol screening.
- High cholesterol levels are a significant factor in heart disease, suggesting the need for dietary interventions and regular monitoring.

Diabetes

Insights:

- Integrate heart disease prevention into diabetes management programs. Educate diabetic patients on the importance of cardiovascular health and provide comprehensive care plans that address both conditions.
- Diabetes is associated with an increased risk of heart disease, indicating the importance of addressing cardiovascular risk factors in diabetic patients.

Further Research

Longitudinal Studies

- Conduct long-term studies to understand the progression of heart disease and its risk factors over time.
- Investigate how changes in lifestyle or medical interventions impact the development of heart disease.

Genetic and Environmental Interactions

- Study the genetic predisposition to heart disease and how it interacts with environmental factors.
- Explore the role of epigenetics in heart disease development.

Behavioural and Psychological Factors

- Research the impact of stress, mental health, and behavioural factors on heart disease.
- Assess the effectiveness of interventions aimed at reducing stress and improving mental health in preventing heart disease.

Technology and Heart Disease

- Investigate the role of emerging technologies, such as wearable devices and telemedicine, in monitoring and managing heart health.
- Explore the potential of artificial intelligence in predicting heart disease risk and tailoring personalized interventions.

Interventions

Community Health Programs

- Develop community-based programs to promote heart-healthy lifestyles, including regular physical activity, healthy eating, and smoking cessation.

- Offer free or low-cost health screenings to identify individuals at risk and provide early interventions.

Patient Education and Support

- Implement educational programs to inform patients about the importance of managing risk factors such as high blood pressure, high cholesterol, and diabetes.
- Provide support groups and counselling services for individuals with heart disease and their families.

Policy and Environmental Changes

- Advocate for policies that support heart-healthy environments, such as creating more green spaces for physical activity and regulating food labelling and advertising.
- Promote initiatives to reduce air pollution and other environmental risk factors for heart disease.

Integrated Care Models

- Develop integrated care models that involve multidisciplinary teams, including cardiologists, primary care physicians, nutritionists, and physiotherapists, to provide comprehensive care for heart disease patients.
- Ensure continuity of care through regular follow-ups and coordinated management plans.

Health Strategies

Preventive Health Screenings

- Encourage regular preventive screenings for blood pressure, cholesterol levels, and blood glucose levels, especially for high-risk populations.

- Utilize risk assessment tools to identify individuals who would benefit from more intensive monitoring and early intervention.

Lifestyle Modification Programs

- Offer programs focused on weight management, smoking cessation, and increasing physical activity.
- Provide resources and support for individuals to make sustainable lifestyle changes.

Public Awareness Campaigns

- Launch public health campaigns to raise awareness about the risk factors and prevention of heart disease.
- Use social media and other digital platforms to disseminate information and engage with the community.

Healthcare Access and Equity

- Address disparities in healthcare access by providing services in underserved communities and ensuring that all individuals have access to preventive and therapeutic care.
- Implement culturally sensitive health interventions that consider the specific needs and preferences of diverse populations.

Conclusion

The analysis identifies several key factors associated with heart disease, including general health, age, difficulty walking, high blood pressure, stroke, physical health, high cholesterol, and diabetes. By targeting these factors through tailored interventions and health strategies, it is possible to mitigate the risk of heart disease and improve overall cardiovascular health. Further research should focus on exploring the underlying mechanisms linking these

factors to heart disease and developing effective prevention and management programs. Comprehensive public health initiatives, early intervention, and continuous monitoring are essential to address the multifaceted nature of heart disease and enhance the well-being of affected populations.

The findings from the analysis of heart disease highlight the importance of a multifaceted approach to prevention and management. By conducting further research, implementing targeted interventions, and developing comprehensive health strategies, it is possible to reduce the burden of heart disease and improve cardiovascular health outcomes. Collaboration among healthcare providers, policymakers, researchers, and the community is essential to address the complex factors contributing to heart disease and to create an environment that supports heart-healthy living.