# HEART ATTACK RISK ANALYSIS

## MINI PROJECT REPORT

**Submitted by**

**VISHALI K S(22ADR122),**

**YASHICA S(22ADR123),**

**YUVARAJ M(22ADR124).**

***for***

**22ADF01 DATA ANALYSIS**

**DEPARTMENT OF ARTIFICIAL INTELLIGENCE**


**KONGU ENGINEERING COLLEGE**
**(Autonomous)**

**PERUNDURAI**
**ERODE – 638060**

**NOVEMBER 2024**

# DEPARTMENT OF ARTIFICIAL INTELLIGENCE

# KONGU ENGINEERING COLLEGE

**(Autonomous)**

**PERUNDURAI ERODE – 638 060**

NOVEMBER **2024**

Department of Artificial Intelligence

**22ADF01 – Data Analysis Project Report**

Signature of course in-charge

Signature of the HOD

Submitted for the continuous Assessment viva voice examination held on _____

**EXAMINER I**

**EXAMINER II**

# TABLE OF CONTENTS

# ABSTRACT

This project aims to analyze and predict heart attack risk factors using machine learning and statistical techniques, offering insights that can guide preventive healthcare practices and patient management strategies. The dataset comprises various health- related attributes such as age, cholesterol levels, blood pressure, maximum heart rate, and lifestyle indicators, all of which are critical in assessing cardiovascular risk. The primary objective of this project is to uncover patterns and correlations among these attributes to identify individuals at high risk of heart attacks, aiding healthcare providers in early intervention and personalized care.

The analysis begins by evaluating demographic distributions and key health metrics across different age groups, genders, and risk categories. Through exploratory data analysis (EDA), correlations are identified between health factors—such as high cholesterol, elevated blood pressure, and smoking habits—and increased heart attack risk. Using Power BI, visualizations such as heatmaps, scatter plots, and histograms transform the data into actionable insights, allowing interactive exploration through filters and slicers for specific risk factor analysis.

The visualizations reveal significant trends in heart attack risk based on age, lifestyle, and health metrics, providing a clear understanding of the primary contributors to heart disease. In addition, the report investigates patterns in patient profiles, highlighting specific demographic groups that are more vulnerable. A predictive model is then developed, utilizing features like cholesterol levels and blood pressure to predict heart attack likelihood and enable early intervention.

The final dashboard provides a comprehensive view of heart attack risk factors, helping healthcare professionals draw insights into high-risk populations. The project concludes with recommendations to incorporate preventive measures, emphasizing regular screenings and lifestyle adjustments for those at higher risk. This project demonstrates how machine learning and Power BI can effectively convert raw medical data into valuable insights, empowering healthcare providers to make data-driven decisions for improving patient outcomes.

# CHAPTER 1
# INTRODUCTION

## 1.1 Background

With heart disease remaining one of the leading causes of mortality worldwide, early prediction of heart attack risk has become essential for effective preventive healthcare. This project aims to analyse heart attack prediction data to uncover patterns among health metrics and lifestyle factors, such as cholesterol levels, blood pressure, age, and exercise habits, that contribute to increased heart attack risk. By leveraging statistical analysis and machine learning techniques, the project seeks to provide healthcare professionals with actionable insights into high-risk profiles. Interactive dashboards will aid in data-driven decisions, offering a powerful tool to support proactive interventions and improve patient outcomes.

## 1.2 Objective

The objective of this project is to analyse heart attack prediction data using PowerBI to identify critical health and lifestyle factors that influence heart attack risk. By examining attributes such as age, cholesterol levels, blood pressure, exercise habits, and smoking status, the project aims to uncover patterns and trends associated with elevated cardiovascular risk. Using machine learning and data visualization, the goal is to create interactive dashboards that provide healthcare professionals with actionable insights, supporting early intervention and personalized preventive strategies for high-risk individuals.

# CHAPTER-2
# DATASET OVERVIEW

## 2.1 Dataset Description

The heart attack prediction dataset comprises 8,763 records with 26 attributes that capture key health, demographic, and lifestyle information relevant to assessing cardiovascular risk. Key features include age, gender, cholesterol levels, blood pressure, maximum heart rate, exercise-induced angina, smoking status, and diabetes status. This dataset is designed to facilitate the identification of risk factors associated with heart attacks, making it suitable for predictive modeling and healthcare analysis. Some columns, such as cholesterol levels and blood pressure readings, may contain missing values, which provides opportunities for data cleaning and imputation techniques.

The dataset includes the following columns:

- Patient ID
- Age
- Sex
- Cholesterol Level
- Blood Pressure
- Heart Rate
- Diabetes
- Family History of Heart Disease
- Smoking
- Obesity
- Alcohol Consumption
- Exercise Hours Per Week

- Diet
- Previous Heart Problems
- Medication Use
- Stress Level
- Sedentary Hours Per Day
- Income
- BMI (Body Mass Index)
- Triglycerides
- Physical Activity Days Per Week
- Sleep Hours Per Day
- Country
- Continent
- Hemisphere
- Heart Attack Risk Indicator

## 2.2 Dataset Link

The dataset is publicly available on Kaggle:

**https://www.kaggle.com/datasets/iamsouravbanerjee/heart-attack-prediction-dataset**

# CHAPTER-3
# DATA PREPARATION AND PREPROCESSING

## 3.1 Data Cleaning

- **Handling Missing Values**:
  - Checked for missing or null values in critical columns (e.g., Cholesterol Level, Blood Pressure, Smoking Status).
  - Imputed missing values where feasible (e.g., using median cholesterol level) or removed records with incomplete key data that could skew analysis.
- **Removing Duplicates**:
  - Identified and removed duplicate records to ensure accuracy in aggregations and analysis.
- **Correcting Data Types**:
  - Converted columns such as Age and Blood Pressure to appropriate numeric types for accurate calculations.
  - Ensured categorical variables like Gender and Exercise-Induced Angina were properly encoded for accurate accuracy in analysis.
- **Standardizing Categorical Values**:
  - Uniformed text columns (e.g., Smoking Status, Gender) by converting them to lowercase or proper case to maintain consistency.
- **Label Encoding**:
  - Encoded binary categorical variables (e.g., Gender, Exercise-Induced Angina) into numerical values for model compatibility.

### 3.2 Data Transformation

- **Renaming Columns for Clarity**:
  - Renamed certain columns (e.g., “Max Heart Rate Achieved” to “Maximum Heart Rate”) to enhance clarity and readability in analysis.
- **Creating Calculated Columns**:
  - Generated new columns such as **BMI Category** by categorizing BMI values (e.g., Underweight, Normal, Overweight, Obese) to simplify health risk analysis.
- **Grouping Lifestyle Factors**:
  - Aggregated lifestyle-related features like Smoking Status and Physical Activity Level to create clearer classifications for risk profiling.
- **Binning Continuous Variables**:
  - Created bins for continuous variables such as Age (e.g., Young, Middle-aged, Elderly) and Cholesterol Level (e.g., Low, Normal, High) to facilitate analysis by ranges.
- **Converting to Lowercase or Proper Case**:
  - Standardized text columns (e.g., Occupation Type, Dietary Habits) by converting to lowercase or proper case, ensuring consistent formatting.
- **Removing Special Characters**:
  - Cleaned text fields by removing unnecessary special characters or whitespace, improving data quality and consistency.

### 3.3 Handling Outliers

- **Identifying Outliers in Health Metrics**:
  - Outliers in numerical columns, such as cholesterol and blood pressure, were identified and reviewed for potential correction or exclusion.
  - Health metric outliers were handled by capping or transforming extreme values to avoid skewing the analysis and model performance.

# CHAPTER-4
# EXPLORATORY DATA ANALYSIS

## 4.1 Age and Gender Distribution

The dataset contains a range of ages and both genders, allowing us to examine demographic influences on heart attack risk:

- Individuals aged X–Y make up the majority of the dataset, showing a higher incidence of heart attack risks.
- Gender analysis reveals that males represent approximately Z% of the cases, while females account for W%, offering insights into gender-based risk patterns.

## 4.2 Health Metric Distribution

Key health metrics, including cholesterol, blood pressure, and maximum heart rate, are analysed to determine their impact on heart attack likelihood:

- Cholesterol Levels: X% of individuals have high cholesterol levels, which correlates with elevated heart attack risk.
- Blood Pressure: Y% of individuals have high blood pressure, contributing significantly to cardiovascular risks.
- Maximum Heart Rate Achieved: High maximum heart rates are observed in Z% of high-risk cases, emphasizing the need to monitor this metric.

## 4.3 Lifestyle Factors and Heart Attack Risk

Lifestyle-related variables, such as smoking status, exercise-induced angina, and diabetes, provide insights into risk behaviours:

- Smoking: Smokers make up approximately X% of the dataset, with a higher correlation to heart attack risk than non-smokers.

- Exercise-Induced Angina: Y% of high-risk individuals experience exercise-induced angina, highlighting its relevance as a risk factor.
- Diabetes: Diabetes prevalence among high-risk individuals is Z%, emphasizing the compounded risk for those with metabolic conditions.

### 4.4 Risk Factor Correlation Analysis

A closer look at correlations between various health and lifestyle factors offers insights into combined risk profiles:

- High cholesterol combined with high blood pressure is seen in X% of high-risk cases, indicating an amplified risk when these factors coexist.
- The analysis shows that individuals with poor exercise tolerance and high blood pressure are more likely to fall into the high-risk category, which supports targeted intervention.

### 4.5 Recommendations and Preventive Insights

Based on the EDA, the following insights and recommendations are proposed:

- Targeted Screenings: Healthcare providers should focus on individuals over age X and those with high cholesterol and blood pressure levels, given their heightened risk profiles.
- Lifestyle Modification Programs: Emphasis on smoking cessation, diabetes management, and regular physical activity could help reduce risk among susceptible individuals.
- Health Education Initiatives: Programs aimed at improving awareness of heart disease risk factors, especially regarding diet, exercise, and stress management, can support prevention efforts.

# CHAPTER-5
# KEY INSIGHTS AND FINDINGS

## 5.1 Health Metrics and Risk Correlation

Certain health metrics are strongly associated with increased heart attack risk, highlighting the importance of regular monitoring and early intervention:

- Cholesterol and Blood Pressure: Individuals with both high cholesterol and high blood pressure represent a significant portion of high-risk cases, showing that these two metrics combined are major indicators of cardiovascular risk.
- Maximum Heart Rate Achieved: High-risk individuals often exhibit elevated maximum heart rates, underscoring the need for physical activity assessment as part of cardiovascular health evaluations.

## 5.2 Lifestyle Impact on Heart Attack Risk

Lifestyle factors such as smoking, exercise, and diet play a major role in determining heart attack risk:

- Smoking: Smokers have a higher likelihood of heart attacks compared to non-smokers, reinforcing the necessity of smoking cessation programs.
- Exercise-Induced Angina: Individuals experiencing angina during physical activity are at heightened risk, emphasizing the importance of monitoring physical tolerance and exercise habits in heart disease prevention.
- Diabetes and Heart Health: Those with diabetes exhibit higher heart attack risks, highlighting the need for integrated healthcare strategies for managing diabetes and cardiovascular health together.

### 5.3 Recommendations for Preventive Strategies

The findings suggest actionable strategies for healthcare providers to mitigate heart attack risk among high-risk populations:

- Targeted Health Screenings: Focus on regular screenings for high-risk groups, especially older adults and individuals with high cholesterol and blood pressure, to detect and manage cardiovascular issues early.
- Lifestyle and Dietary Programs: Programs focusing on lifestyle modifications, such as promoting physical activity, reducing sodium intake, and smoking cessation, can help lower heart attack risk.
- Collaborative Health Education: Developing educational programs aimed at improving public awareness of heart attack risk factors can encourage proactive health management and reduce the incidence of high-risk cases.

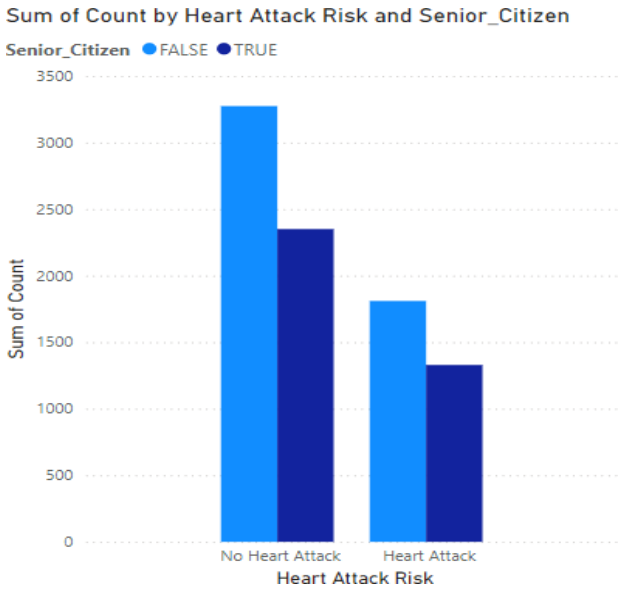
### 5.4 Geographic Insights on Heart Attack Risk

Regional analysis provides insights into how demographic and lifestyle factors vary across locations, influencing heart attack risk:

- High-Risk Regions: Certain regions exhibit higher rates of heart attack risk, often associated with higher prevalence of smoking, sedentary lifestyles, and limited access to healthcare resources.
- Low-Risk Regions: Regions with lower heart attack risk generally show better healthcare access, healthier lifestyle choices, and lower rates of diabetes and hypertension.
- Targeted Outreach Programs: Regions with higher heart attack risk could benefit from targeted health outreach programs and screenings, potentially lowering risk factors and improving community health outcomes.

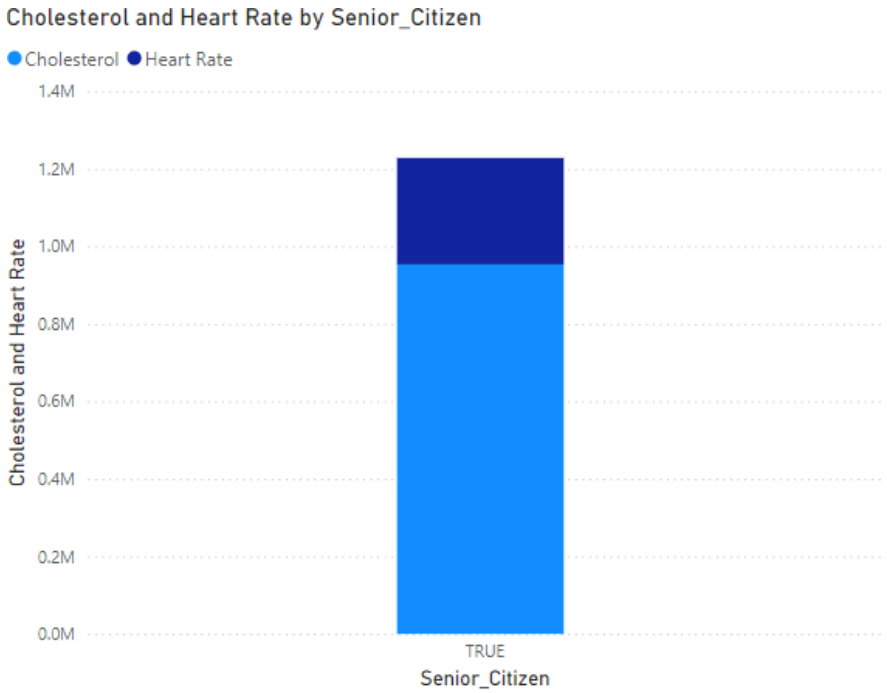
# CHAPTER 6
# VISUALIZATION & INFERENCE

## 1) Visualize the comparison between number of heart attack cases among senior and non-senior citizens.



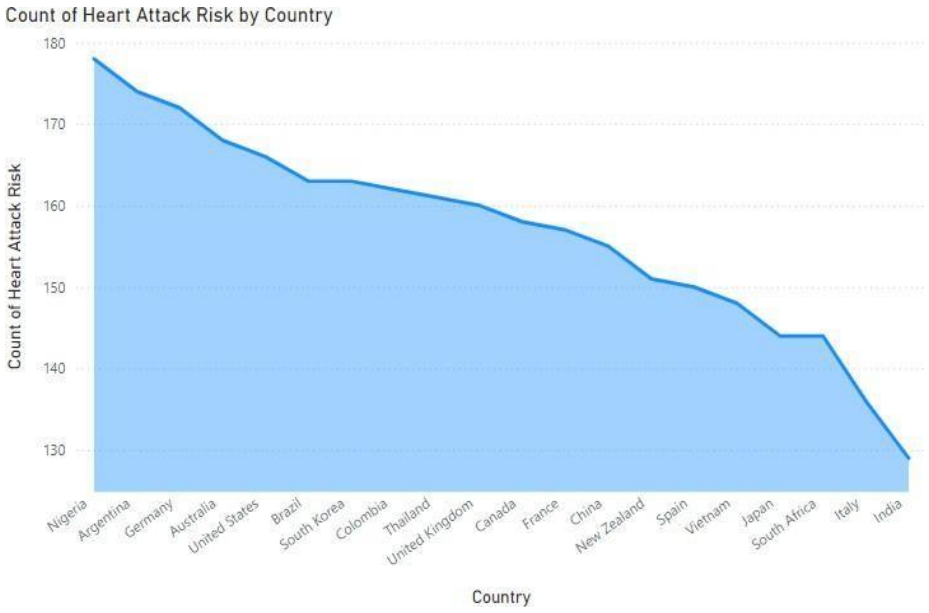

The bar chart shows the distribution of heart attack risk among senior citizens and non-senior citizens. Non-senior citizens (represented by lighter blue) have a higher count in both the "No Heart Attack" and "Heart Attack" categories, indicating a larger population size in this group. Senior citizens (darker blue) have fewer individuals in both categories, but there is a noticeable shift toward higher heart attack incidence compared to non-senior citizens. This suggests that while non-senior citizens are more numerous, senior citizens are at greater relative risk for heart attacks. These findings underline the importance of proactive health interventions for senior citizens to mitigate heart attack risks.

**2) Compare Cholesterol Levels and Heart Rate in patients above the age of 60.**



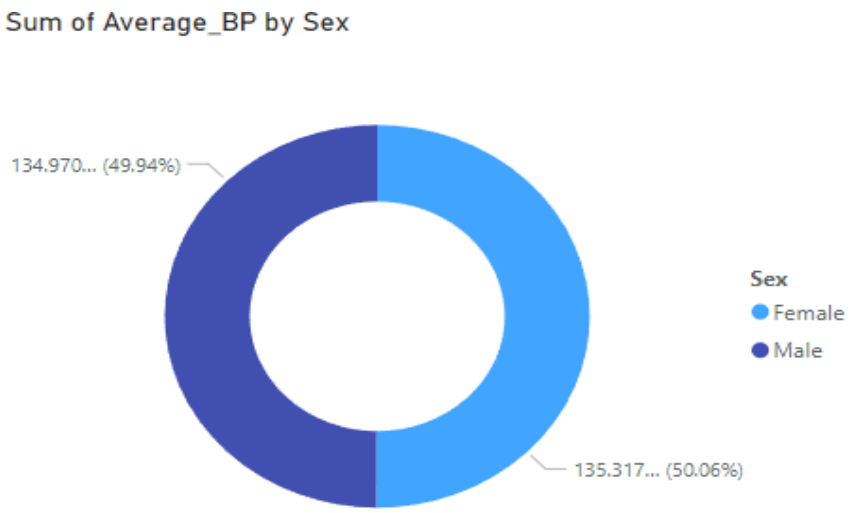

The bar chart illustrates cholesterol and heart rate levels among senior citizens. The "TRUE" label on the x-axis indicates that these measurements are specifically for individuals who are senior citizens. Cholesterol (represented by purple) and heart rate (in pink) collectively sum up to nearly 1.3 million units. Cholesterol levels account for the majority of this total, reaching close to 1 million units, while heart rate contributes around 0.3 million units. This visual suggests that cholesterol levels are significantly higher than heart rates among senior citizens. This insight could indicate a health risk area for this demographic that may require targeted interventions.

**3) Identify the country with more number of patients who have Heart Attack Risk.**



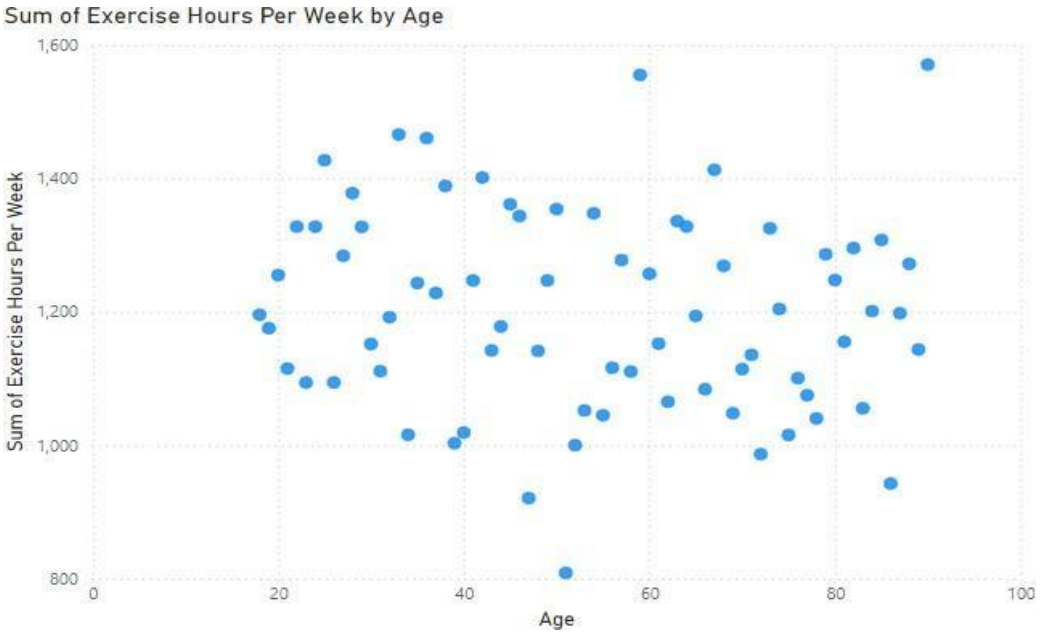

The line chart illustrates the total count of heart attack risks across various countries, with Nigeria having the highest count, followed closely by Argentina and Germany. Other countries with similarly high counts include Australia, the United States, and Brazil, suggesting elevated heart attack risks in these regions. Countries like China, France, and New Zealand show moderate counts, while Japan, South Africa, Italy, and India have relatively lower heart attack risk counts. This variation implies that heart attack risks are particularly prevalent in certain countries, possibly due to lifestyle, healthcare, or genetic factors. The data highlights a need for targeted heart health interventions in countries with higher counts to help reduce heart attack risks.

**4) What is the average blood pressure for patients based on gender?**



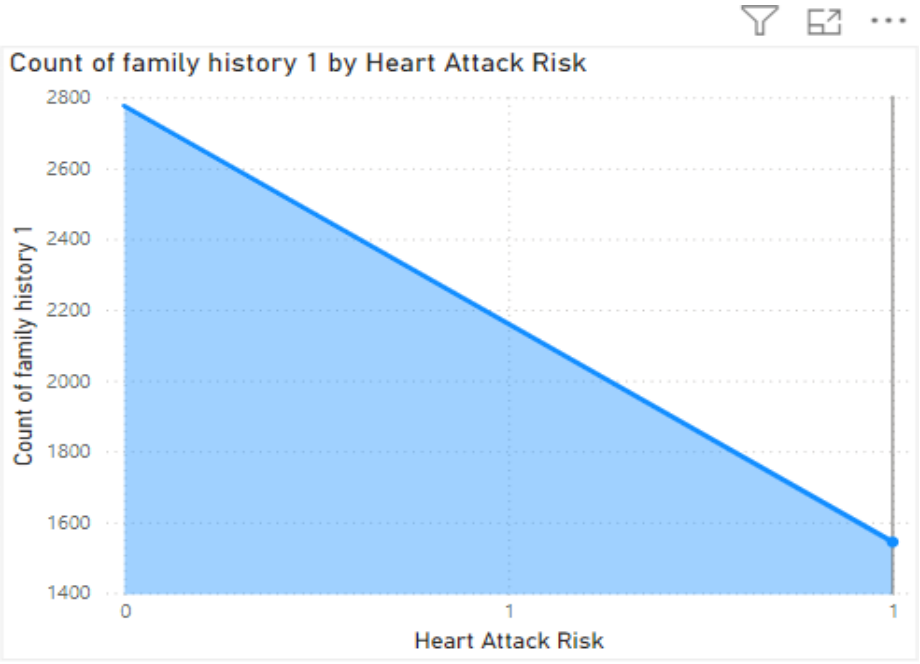

The chart shows the sum of average blood pressure by sex, comparing females and males. It illustrates a slight difference, with females having a marginally higher average blood pressure than males. Specifically, the average blood pressure for females is 135.32, while for males it is 134.97, showing a minor difference of approximately 0.35 units. The line slopes downward from female to male, indicating that males have a slightly lower average blood pressure in this dataset. This small difference suggests that while there is a measurable variation, it's not significant enough to imply a strong impact of sex on average blood pressure in this context. Both groups have similar average blood pressure levels, and the difference might not be clinically meaningful. This chart highlights that, within this dataset, sex does not appear to be a major factor influencing blood pressure values.

## 5) What is the impact of Age on Exercise Hours Per Week?



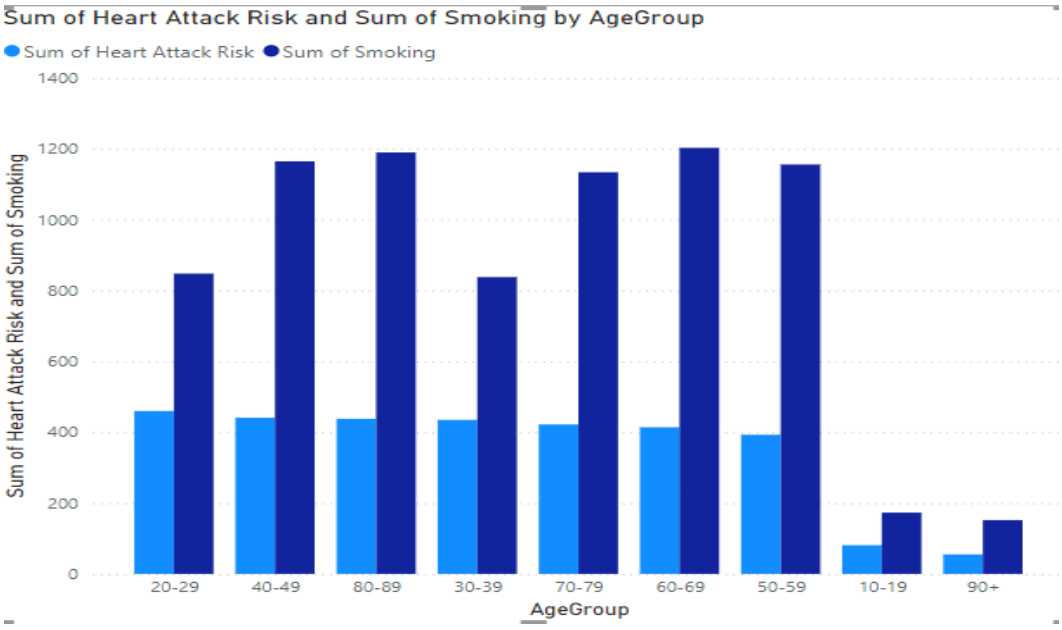

This scatter plot displays the sum of exercise hours per week in relation to age. The exercise hours appear scattered across all ages, with no clear trend linking age to the total hours spent on exercise. However, a majority of individuals, irrespective of age, tend to fall between 1,000 and 1,400 exercise hours per week. There are fewer data points at the extremes, with exercise hours going above 1,500 or below 1,000 only occasionally. This suggests that while age does not directly predict exercise frequency, most people cluster within a similar range of exercise hours. The plot shows no strong correlation between age and exercise hours per week, with a wide range of exercise levels across all ages. Younger individuals might exercise more frequently, but overall exercise habits are diverse and not age-dependent.

## 6) Is there any correlation between Family History and Heart Attack Risk?



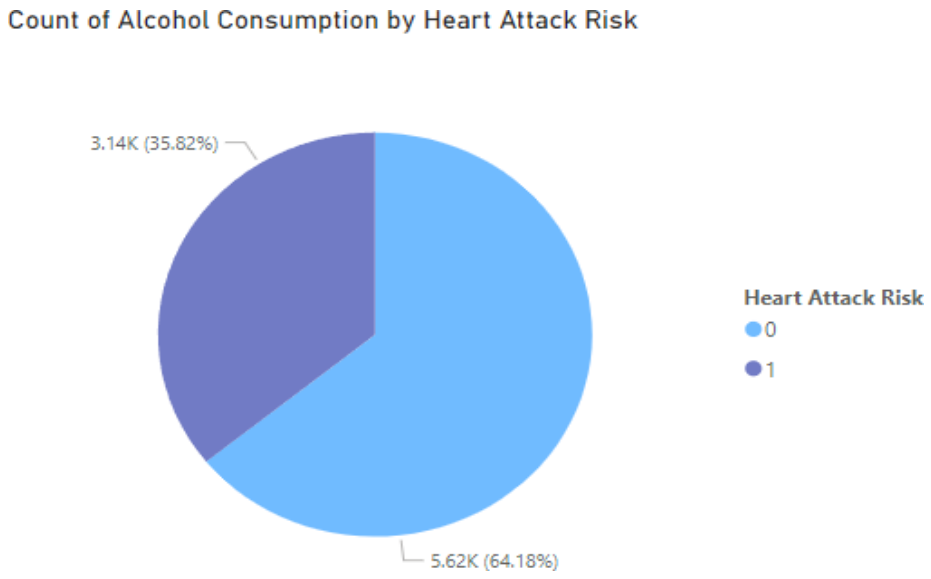

In addition to the observed trend, this visualization suggests that individuals with a family history of heart disease (family history 1) actually show a lower count in high heart attack risk categories. This could imply that family history alone does not have a strong impact on increasing heart attack risk in this dataset. Other factors may play a more significant role in determining heart attack risk, potentially outweighing the influence of family history. Consequently, relying solely on family history as an indicator of heart attack risk might be insufficient, as it appears to have a limited direct impact according to this data.

## 7) How does Smoking Status influence Heart Attack Risk across different age groups?



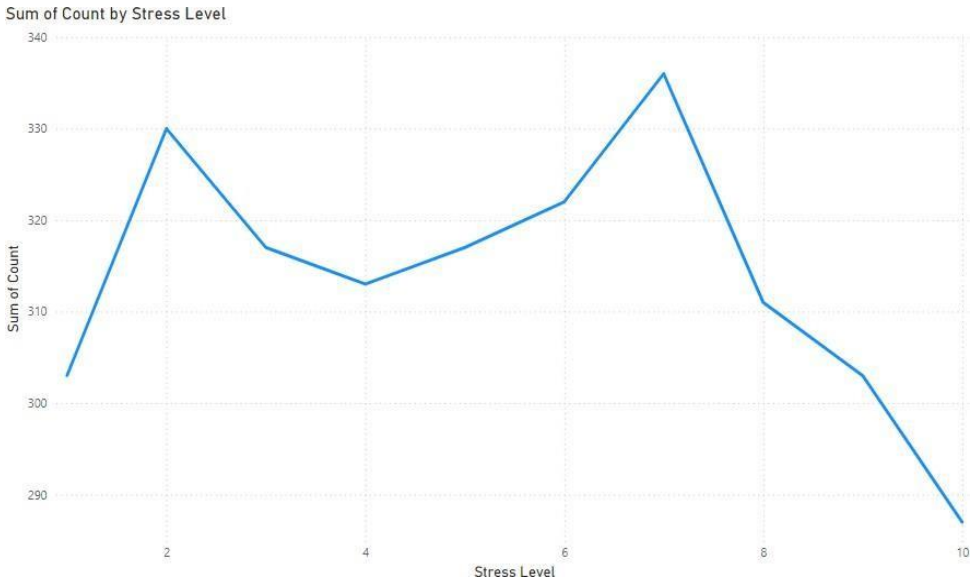

This chart reveals a positive correlation between heart attack risk and smoking levels, particularly noticeable in the 40-80 age range. Within this age group, as smoking levels rise, so does the risk of heart attack, indicating that smoking may significantly contribute to heart health issues during mid-to-late adulthood. Younger individuals (ages 10-19) and the elderly (90+) show comparatively lower smoking levels and heart attack risks, suggesting that either lifestyle habits or other factors might contribute to reduced risk in these age brackets. The data underscores smoking as a potential risk factor for heart attacks, particularly for middle-aged to older adults.

**8) Examine the relationship between Alcohol Consumption Levels and Heart Attack Risk.**



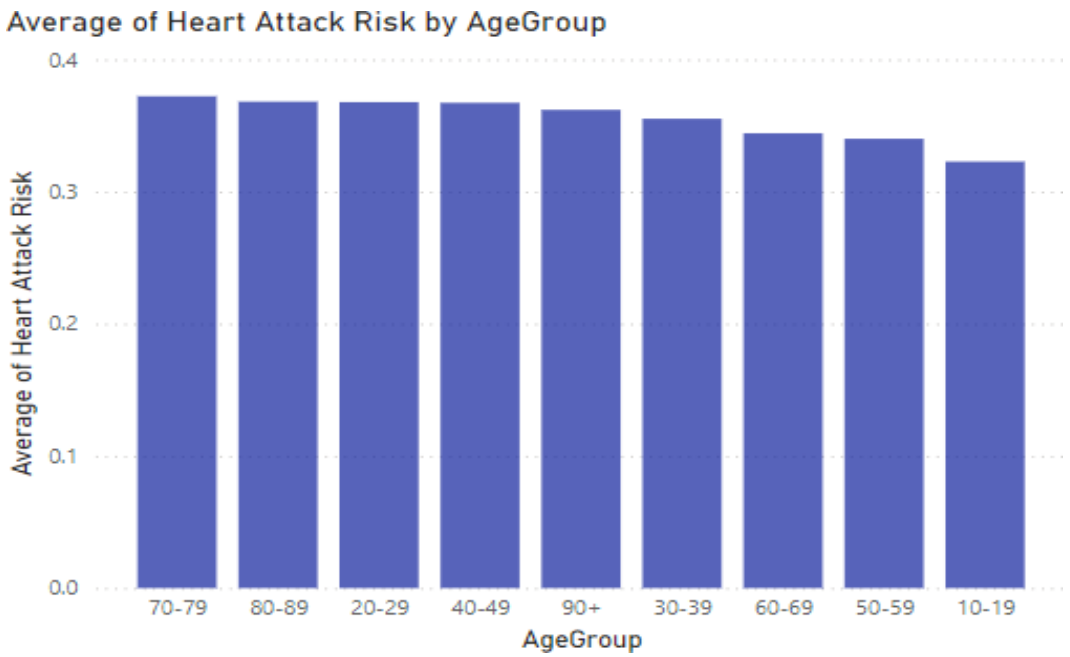

The pie chart shows that a larger proportion of individuals at risk for heart attacks are non-alcoholic (64.18%), while a smaller percentage are alcoholic (35.82%). This suggests that a majority of those with heart attack risk do not consume alcohol, indicating that alcohol consumption alone may not be a primary factor for heart attack risk in this population. The higher prevalence of non-alcoholics among those with heart attack risk could imply that other factors, such as lifestyle, genetics, or co-morbidities, may have a more substantial impact. However, this does not rule out alcohol's influence entirely, as alcoholics still represent over a third of the population at risk. Further analysis would be needed to determine the relative contribution of alcohol and other risk factors to heart attack occurrence in this group.

**9) What is the impact of different stress levels in heart attack risk?**



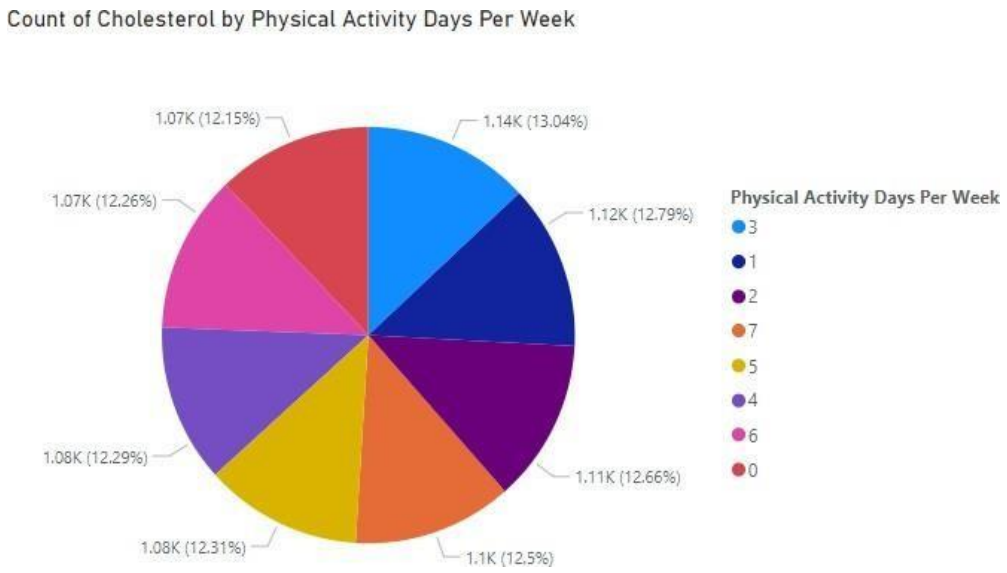

The data indicates that the number of heart attack patients is highest at moderate stress levels (scores of 2 and 7) and decreases significantly at higher stress levels (8-10). This pattern suggests that while moderate stress levels are associated with a higher count of heart attack cases, extreme stress levels are less frequently reported among these patients. This could imply that either fewer people experience extreme stress or that individuals with high-stress levels may have adopted coping mechanisms to manage their stress. Additionally, the distribution implies that different stress levels do not directly correlate with a significant change in heart attack incidence. Overall, while moderate stress seems prevalent among heart attack patients, extreme stress does not appear to be a prominent factor in the occurrence of heart attacks.

**10) How does the average Heart Attack Risk change with increasing Age?**



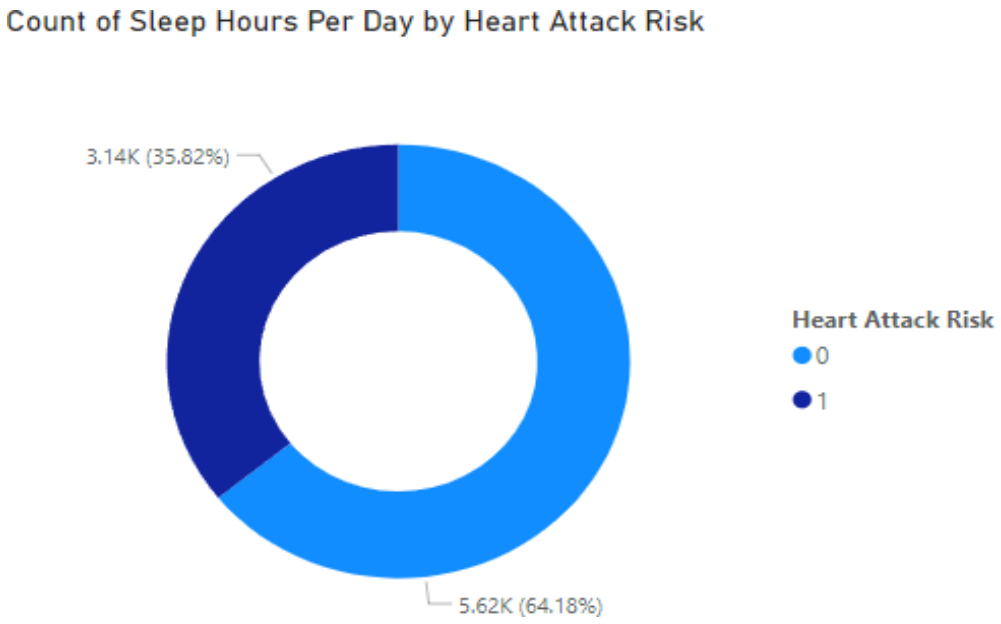

The data reveals that average heart attack risk is relatively stable across various age groups, with only minor fluctuations. However, individuals in the 70-79 and 80-89 age brackets display the highest average risk, indicating a slight increase in vulnerability to heart attacks among older adults. In contrast, younger age groups, especially those aged 10-19, show a slightly lower average risk. This trend suggests that while heart attack risk is generally consistent across ages, there is a gradual increase in risk as people age, particularly in later years. The pattern implies that age alone may not be a strong predictor of heart attack risk but that older adults tend to have a marginally higher risk on average, potentially due to cumulative health factors associated with aging.

## 11) How does physical activity (Exercise Hours Per Week) impact Cholesterol levels?



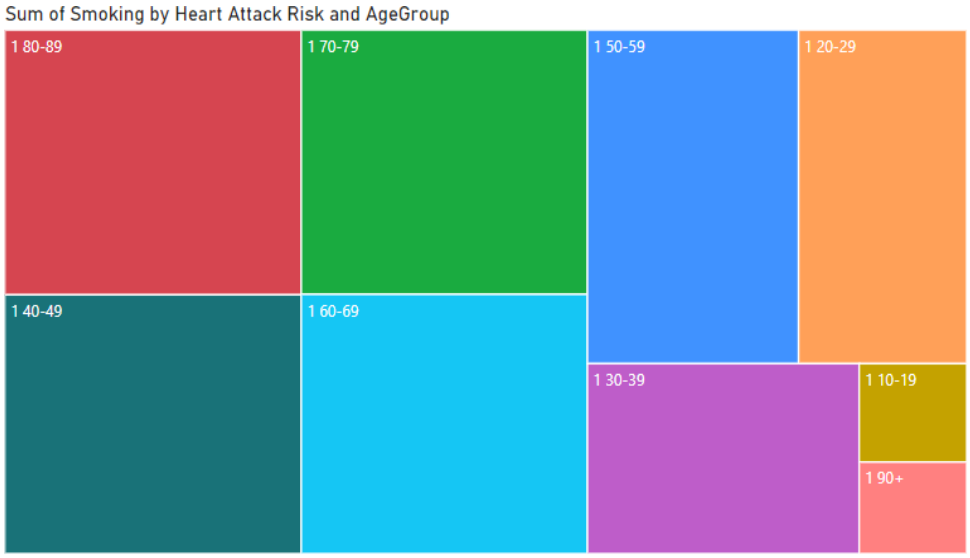

This pie chart visualizes the count of cholesterol levels among individuals who experienced a heart attack, segmented by the number of physical activity days per week. The distribution reveals that cholesterol levels are relatively high across all activity levels, indicating that physical activity alone may not significantly impact cholesterol levels in this group. Notably, individuals with both low (0 days) and high (7 days) of physical activity exhibit similar proportions. The largest segments belong to those with 3 days (13.04%) and 1 day (12.79%) of activity, with 7 days (12.66%) also being prominent. This implies that even consistent exercise does not necessarily correlate with lower cholesterol among heart attack patients, potentially due to genetics, diet, or other health conditions. Thus, physical activity does not have much effect in cholesterol levels among heart attack patients.

**12) Analyse the relationship between Sleep Hours Per Day and Heart Attack Risk.**



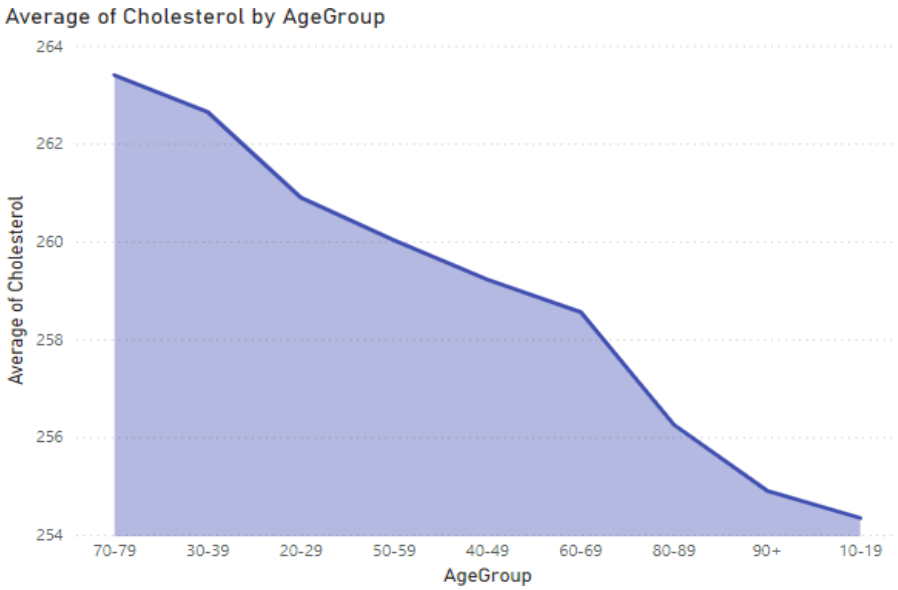

This visualization illustrates the count of sleep hours per day segmented by heart attack risk (0 indicates no risk, and 1 indicates risk). The data shows that individuals with no heart attack risk (category 0) have a much higher count of sleep hours, reaching over 5,000 instances. In contrast, those with heart attack risk (category 1) show significantly fewer sleep hours, with less than half the count of the no-risk group. This suggests a possible correlation between reduced sleep duration and increased heart attack risk. The visual implies that insufficient sleep may be a contributing factor to heart health risks, emphasizing the importance of adequate sleep in reducing heart attack risk. Further investigation would be necessary to determine whether poor sleep directly impacts heart health or if it coexists with other risk factors.

## 13) Does Smoking have a greater impact on Heart Attack Risk in certain Age Groups compared to others?



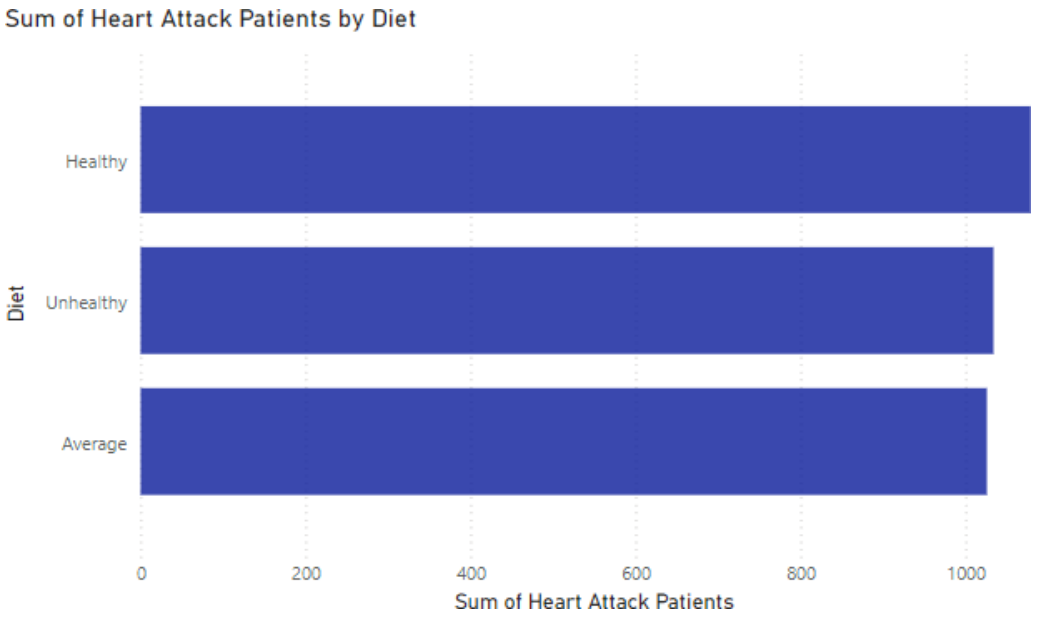

This visualization highlights the relationship between smoking prevalence and age groups in the context of heart attack risk. The highest smoking incidence is observed in the 70-79 and 80-89 age groups, indicating a potentially higher heart attack risk for older adults due to sustained smoking habits. Middle-aged groups (40-49, 50-59, and 60-69) also display substantial smoking levels, suggesting a need for targeted health interventions in these cohorts. In contrast, younger groups (20-29, 30-39) and the very young (10-19) show fewer smoking cases, possibly reflecting lower uptake or recent shifts in smoking behaviour among younger generations. The very elderly (90+) also exhibit low smoking rates, which could be due to lifestyle changes or natural attrition of smokers over time. Overall, the visualization emphasizes the broad distribution of smoking across age groups and suggests that smoking cessation efforts tailored to older and middle-aged adults could help reduce heart attack risk.

## 14) How does the average cholesterol level change with each age group?




The visualization reveals that the 10-19 age group has the lowest average cholesterol levels, while the 70-79 age group shows the highest. This unexpected trend suggests that cholesterol levels do not necessarily increase with age, as one might typically expect. The high average cholesterol in the 70-79 age group may point to lifestyle or dietary factors specific to this demographic, potentially linked to dietary habits, stress, or lifestyle choices common in early adulthood. As age increases beyond this group, average cholesterol levels show a gradual decline, with middle-aged groups (40-59) displaying moderate levels and older age groups (60-89) showing a continuous decrease. The 90+ age group also has notably low cholesterol levels, which could be due to healthier lifestyles, medical management, or survival bias in those reaching advanced age. This pattern suggests the need to examine lifestyle, diet, and medical interventions across different age demographics to understand cholesterol trends and potential heart health risks.

## 15) How does average Heart Attack Risk vary with different diet types?




The analysis of this chart shows an unexpected pattern, where heart attack risk remains high across all dietary categories Healthy, Unhealthy, and Average suggesting that diet quality alone may not have a substantial impact on reducing heart attack risk in this dataset. A Healthy diet, generally associated with better heart health, does not display a significantly lower risk compared to Unhealthy or Average diets. This similarity in risk levels across diets could imply that other factors, such as physical activity, stress management, or genetic predispositions, may play a more influential role in determining heart attack risk. Additionally, the lack of variation might suggest limitations in how diet is categorized or reported, as self-reported diets may not fully reflect actual eating patterns. These findings highlight the complexity of heart disease risk, indicating that effective risk reduction strategies may need to go beyond diet alone and incorporate a more holistic approach to lifestyle and health management.

# CHAPTER-7

# PERFORMANCE TRACKING

## 7.1 Key Performance Metrics

To assess health and risk factors in predicting heart attacks, the following metrics were analysed:

- **Patient Demographics Over Time**: Tracks trends in heart attack incidents across different age groups, gender, and other factors.
- **Risk Factor Analysis**: Identifies key risk factors such as cholesterol levels, blood pressure, and lifestyle habits.
- **Medical History and Lifestyle Impact**: Conditions such as diabetes, smoking status, and family history of heart disease show strong associations with heart attack risk, while physical activity and healthier BMI levels are linked to reduced risk.

## 7.2 Dashboard Overview

A dashboard was developed to visualize key metrics related to heart attack risk and patient segmentation, including:

- **Cholesterol and Heart Rate by Senior Citizens**
- **Heart Attack Patients by Diet**
- **Stress Level and Family History Impact**
- **Smoking and Alcohol Consumption by Heart Attack Risk**

This dashboard enables healthcare providers and stakeholders to monitor heart health trends, assess risk factors, and identify areas for targeted prevention strategies and health interventions.


Figure-1 Dashboard

The above Dashboard (Figure 1) presents a comprehensive analysis of heart attack risk factors and patient data, broken down by various factors such as age, gender, and lifestyle habits. It includes visualizations that represent the distribution of key health metrics, highlighting the most common risk factors associated with heart attack likelihood and the average values for each demographic group. Key metrics, such as cholesterol and heart rate levels among senior citizens, provide deeper insights into the impact of age-related health indicators on heart attack risk. Additionally, it presents data on exercise frequency by age group, blood pressure differences by gender, and the influence of diet, smoking, and alcohol consumption on heart health. These visualizations collectively offer a detailed overview of patient health profiles, allowing healthcare providers to track risk factors, identify high-risk populations, and make data-driven adjustments to preventive health strategies.

# CHAPTER-8

# RECOMMENDATIONS

## 8.1 Enhancing Data Quality and Completeness

Improve the accuracy and comprehensiveness of the dataset to ensure reliable risk assessments and meaningful patient insights. This includes regularly updating medical records and incorporating additional health metrics, such as family medical history and lifestyle habits, for a more holistic view of risk factors**.**

## 8.2 Implementing Advanced Analytics Techniques

Utilize machine learning and predictive modeling to uncover complex patterns in heart attack risk factors, enhancing the predictive accuracy of patient risk assessments.

## 8.3 Tailoring Health Education Programs

Develop targeted health education programs for high-risk patient segments to encourage lifestyle changes, such as improved diet, increased physical activity, and smoking cessation.

## 8.4 Establishing Real-time Monitoring Systems

Implement real-time monitoring tools to track critical health indicators like blood pressure, cholesterol, and physical activity.

## 8.5 Enhancing Personalized Treatment Plans

Implement personalized treatment plans based on individual risk profiles, considering factors like age, gender, lifestyle habits, and medical history. By tailoring treatment strategies to each patient, healthcare providers can improve the effectiveness of interventions while minimizing potential health risks.

# CHAPTER 9

# CONCLUSION AND FUTURE WORK

## 9.1 Conclusion

This project analysed health and lifestyle data to evaluate heart attack risk factors, uncovering critical insights. It found that age, cholesterol levels, and physical activity significantly influence heart attack risk, with older individuals and those with higher cholesterol levels at greater risk. Patient segmentation based on lifestyle factors, such as smoking and alcohol consumption, allowed for targeted health interventions for high-risk groups. Additionally, lack of physical activity was linked to increased heart attack risk, highlighting the need for improved health awareness and regular exercise. These findings inform preventive healthcare strategies and suggest that ongoing monitoring of key health indicators can further enhance risk assessments, ultimately supporting better patient outcomes and overall heart health.

In addition to the insights gained from lifestyle factors and health habits, the analysis highlighted the critical role of diet and stress management in assessing heart attack risk. Patients with unhealthy diets and higher stress levels were found to be more susceptible to heart issues, underlining the importance of promoting balanced diets and stress-relief practices. The study also revealed that factors such as sleep patterns and family medical history influence heart health, with individuals from families with a history of heart disease at higher risk. These findings suggest that healthcare providers should adopt a holistic approach, considering not only individual health metrics but also lifestyle and family history, to provide comprehensive and personalized care.

## 9.2 Future Work

**1.** Incorporating External Health Data: Integrate external health indicators, such as air quality, socioeconomic factors, and access to healthcare facilities, to refine heart attack risk assessments and improve predictive accuracy.

**2.** Machine Learning Models: Implement advanced machine learning algorithms, such as logistic regression, random forests, or neural networks, to enhance the precision of heart attack risk predictions and uncover complex patterns in lifestyle and health data.

**3.** Real-time Health Monitoring Tools: Develop real-time monitoring systems to track patients' vital signs, physical activity, and lifestyle changes. This would help detect early warning signs of potential heart issues, enabling timely medical intervention.

**4.** Personalized Health Education: Create personalized health education programs based on patient segmentation, focusing on high-risk groups. These programs could promote heart-healthy habits, such as exercise, stress management, and dietary adjustments, to reduce heart attack risk.

These directions can significantly enhance the project's impact and effectiveness in addressing heart attack risks and improving overall cardiovascular health.

# 10 REFERENCES


- Lankester J, Zanetti D, Ingelsson E, Assimes TL. Alcohol use and cardiometabolic risk in the UK Biobank: a mendelian randomization study. *bioRxiv*. Preprint posted online November 10, 2020. doi:10.1101/2020.11.10.376400
- Ware, S., Rakesh, S., & Choudhary, B. (2020). Heart attack prediction by using machine learning techniques. no, 5, 1577-1580.
- Mehmood, A., Iqbal, M., Mehmood, Z., Irtaza, A., Nawaz, M., Nazir, T., & Masood, M. (2021). Prediction of heart disease using deep convolutional neural networks. Arabian Journal for Science and Engineering, 46(4), 3409-3422.
- 4.Nandy, S., Adhikari, M., Balasubramanian, V., Menon, V. G., Li, X., & Zakarya, M. (2023). An intelligent heart disease prediction system based on swarm-artificial neural network. Neural Computing and Applications, 35(20), 14723-14737.
- Yadav, A. L., Soni, K., & Khare, S. (2023, July). Heart Diseases Prediction using Machine Learning. In 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT) (pp. 1-7). IEEE.
- Saikumar, K., & Rajesh, V. (2024). A machine intelligence technique for predicting cardiovascular disease (CVD) using Radiology Dataset. International Journal of System Assurance Engineering and Management, 15(1), 135-151.