

**PGDM- Big Data Analytics**

**BUSINESS INTELLIGENCE USING POWER BI**

**Mid-Term Project Report**

**Trimester 3**

**Interactive Dashboard on Power Bi**

**Topic: Heart Disease Analysis in India**

**Submitted to: Prof. Saurabh Mittal**

**Submitted By: 055022- Krishnendu Adhikary**

**14th February, 2025**

**Introduction**

Cardiovascular diseases remain a leading cause of mortality worldwide, with heart disease being one of the most prevalent conditions. This project aims to analyse various factors contributing to heart disease, utilizing a dataset originally sourced from Kaggle. However, due to insufficiencies in the original dataset, additional variables such as "Date of Diagnosis," "Work Type," and "City" were appended using hypothetical data. The primary goal is to derive meaningful insights that help in understanding heart disease risk factors and trends.

**Used Variables**

**The dataset includes the following variables:**

* **Demographics:** Age, Gender, Work Type, City
* **Health Indicators:** BMI, Blood Pressure, Diabetes Status, Stress Level, Sleep Hours
* **Lifestyle Factors:** Smoking, Alcohol Consumption, Exercise Frequency
* **Medical Indicators:** C-Reactive Protein (CRP), Homocysteine Level
* **Temporal Data:** Date of Diagnosis (Month, Year)

**Newly Appended Variables:**

* **Date of Diagnosis –** To analyse time-based trends in heart disease cases.
* **Work Type –** To study the impact of occupation on heart health.
* **City –** To compare the prevalence of heart disease across different urban and rural areas.

**Data Pre-processing**

Given the limited scope of the original dataset, additional variables were incorporated using hypothetical data to ensure a more comprehensive analysis. Data cleaning involved handling missing values, normalizing numerical values, and encoding categorical variables where necessary. Power BI was then used to create interactive visualizations to explore various problem statements.

**Objective of the Project**

**The key objectives of this project include:**

* Identifying the correlation between age and heart disease prevalence.
* Examining the impact of BMI on blood pressure and diabetes.
* Evaluating the effect of exercise on heart disease risk.
* Understanding how smoking and alcohol consumption influence heart disease.
* Analysing the effects of stress and sleep deprivation on heart health.
* Exploring geographic and time-based trends in heart disease cases.

**Problem Statements, Solutions, and Insights**

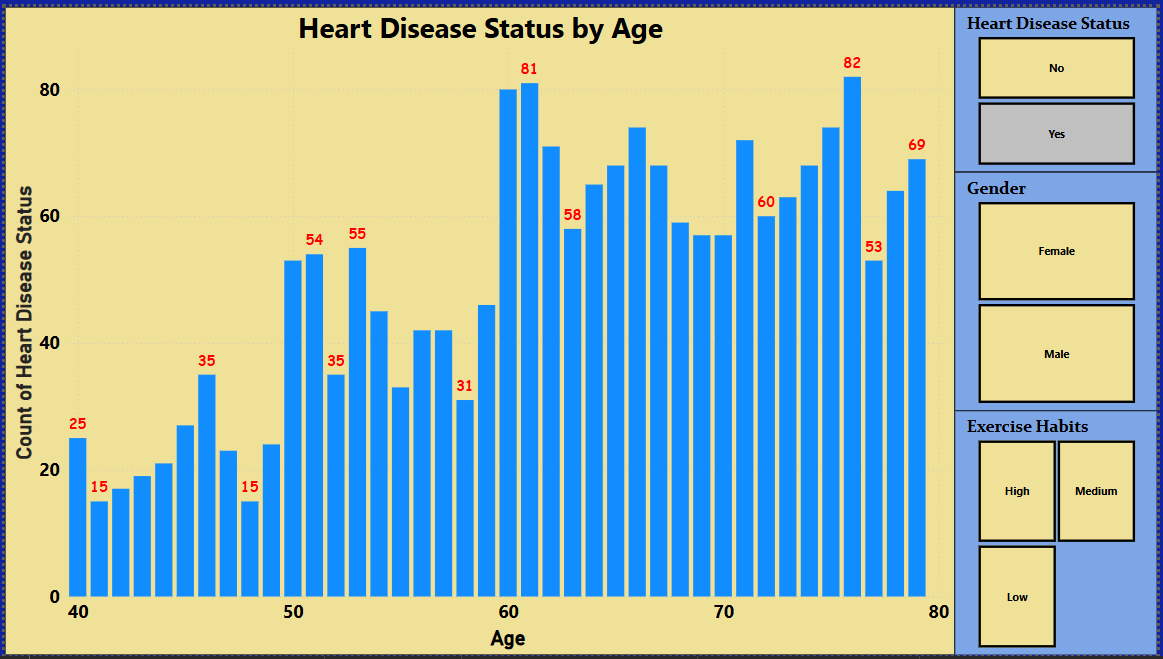
**Problem Statement 1:** How Does Age Affect Heart Disease Prevalence?

**Solution:**

* Older individuals (≥ 40 years) show an increasing trend of heart disease.
* The risk rises significantly after 60 years, indicating age is a major factor.

**Insights:**

* Visualization shows a sharp increase in cases among individuals aged 60+.
* Preventive measures should be targeted at individuals over 40.



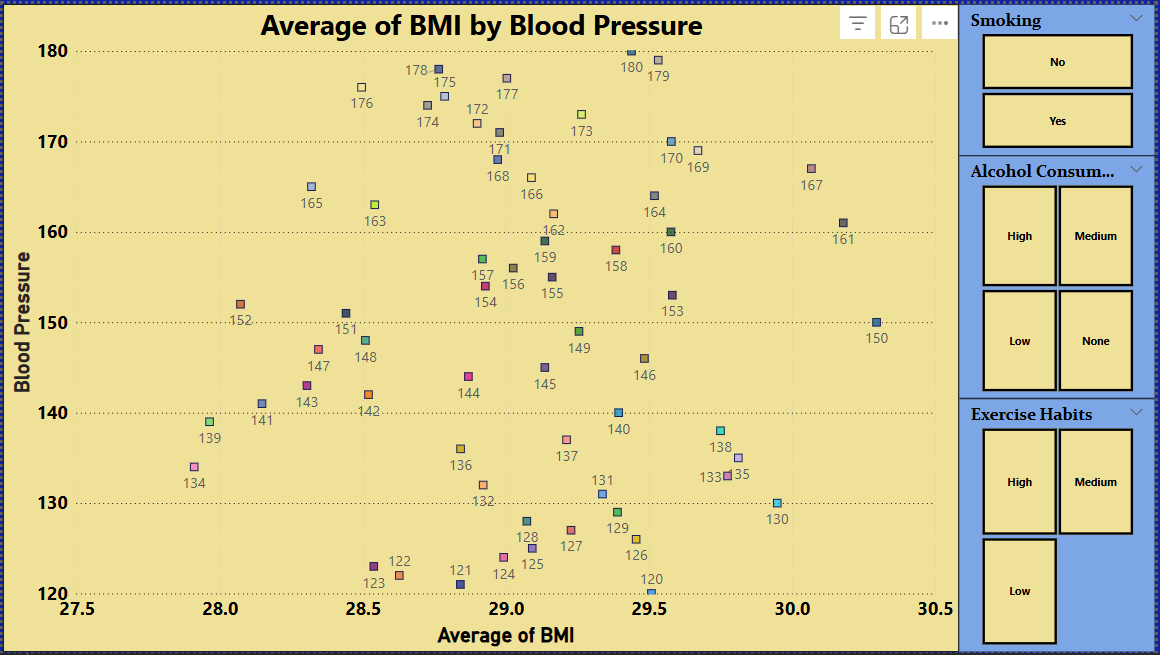
**Problem Statement 2:** How Does BMI Impact Blood Pressure and Diabetes?

**Solution:**

* Higher BMI (≥ 25) increases the likelihood of high blood pressure and diabetes.
* Obese individuals (BMI ≥ 30) have a significantly higher percentage of hypertension and diabetes-positive cases.

**Insights:**

* The data confirms a strong positive correlation between BMI, hypertension, and diabetes.
* Lifestyle interventions should focus on individuals with BMI ≥ 25 to mitigate risks.



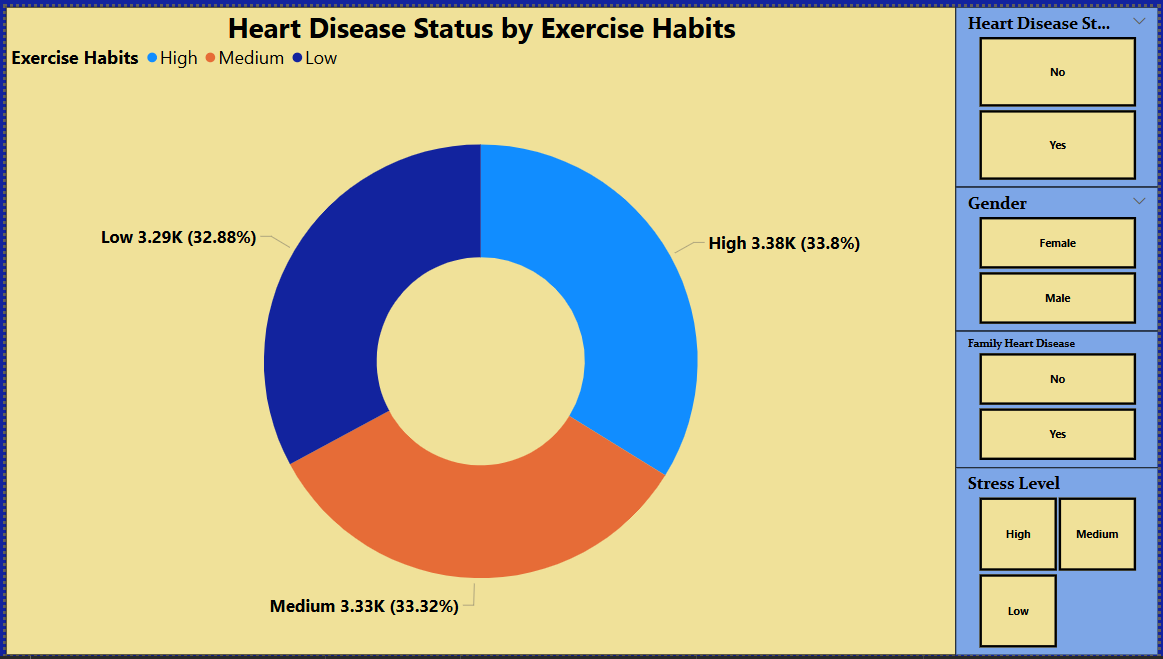
**Problem Statement 3:** What is the Impact of Exercise on Heart Disease?

**Solution:**

* Low exercise habits are directly linked to a higher percentage of heart disease cases.
* Individuals with high exercise habits show significantly lower rates of heart disease despite other risk factors.

**Insights:**

* Sedentary individuals exhibit a higher prevalence of heart disease.
* Encouraging physical activity can significantly reduce cardiovascular risks.



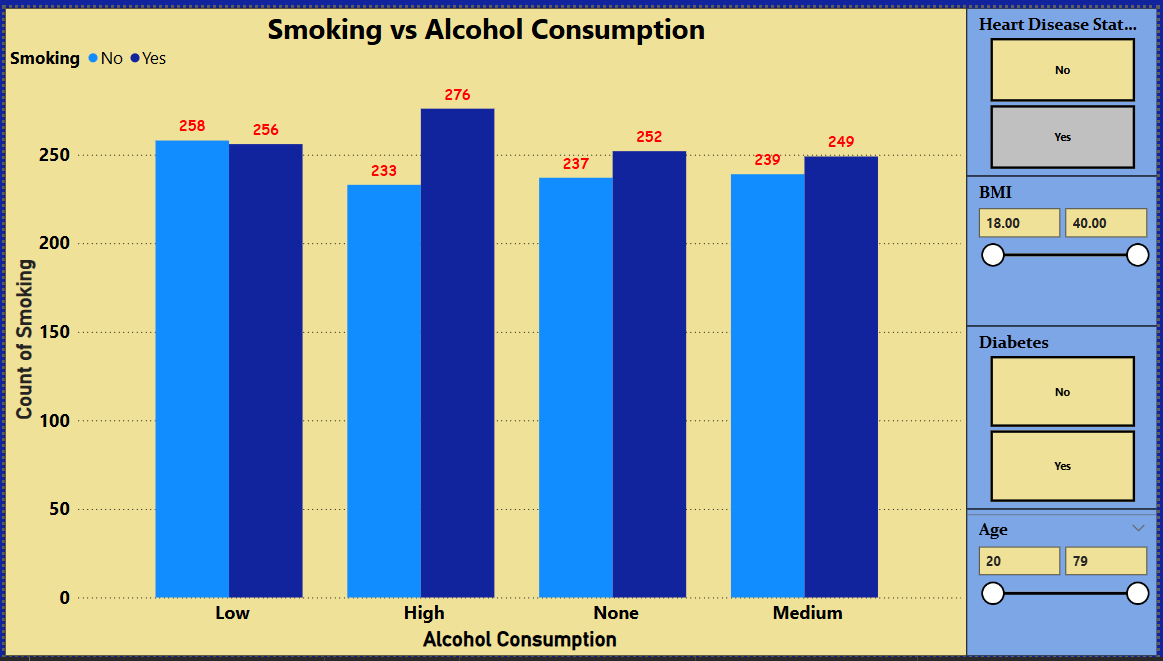
**Problem Statement 4:** Does Smoking and Alcohol Consumption Influence Heart Disease?

**Solution:**

* Smokers and heavy alcohol consumers have a significantly higher risk of heart disease.
* Heavy alcohol consumption (High) combined with smoking shows the worst-case scenario in heart disease prevalence.

**Insights:**

* Data confirms that individuals who engage in both smoking and heavy drinking have the highest incidence of heart disease.
* Awareness programs should focus on reducing these behaviours.



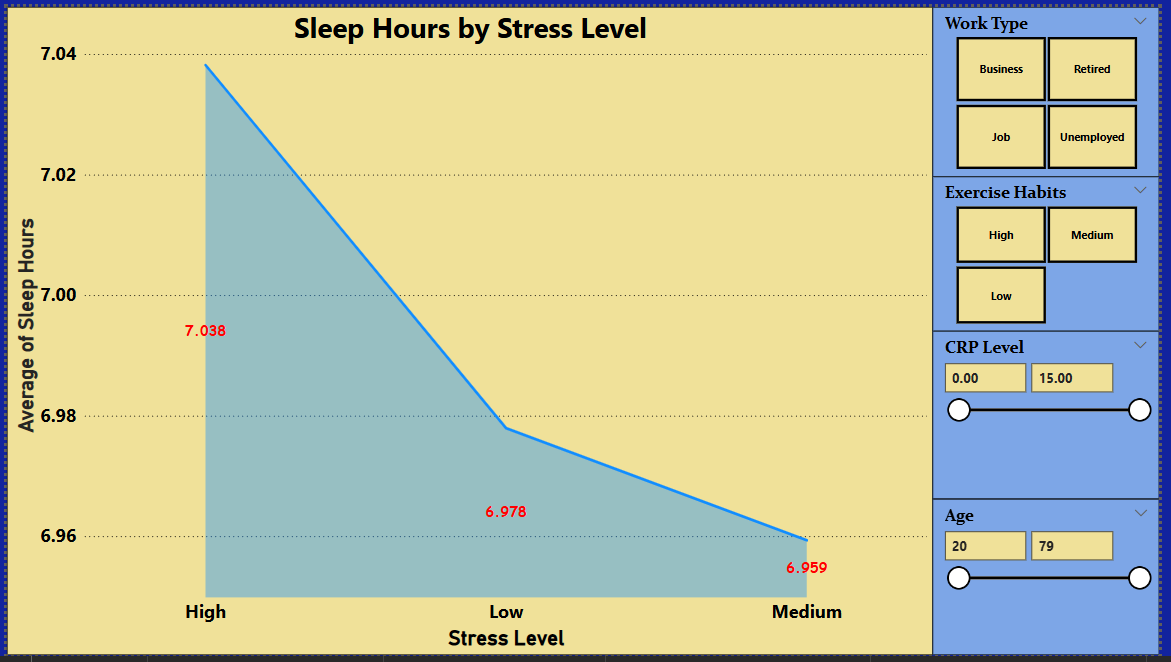
**Problem Statement 5:** How Do Stress and Sleep Affect Heart Disease?

**Solution:**

* Higher stress levels (High) and lower sleep hours (< 6 hrs) significantly contribute to heart disease.
* People with medium/high stress and poor sleep quality show higher CRP & Homocysteine levels (markers of heart risk).

**Insights:**

* Stress management programs and sleep hygiene education are crucial in heart disease prevention.
* People with chronic stress should be closely monitored for early intervention.



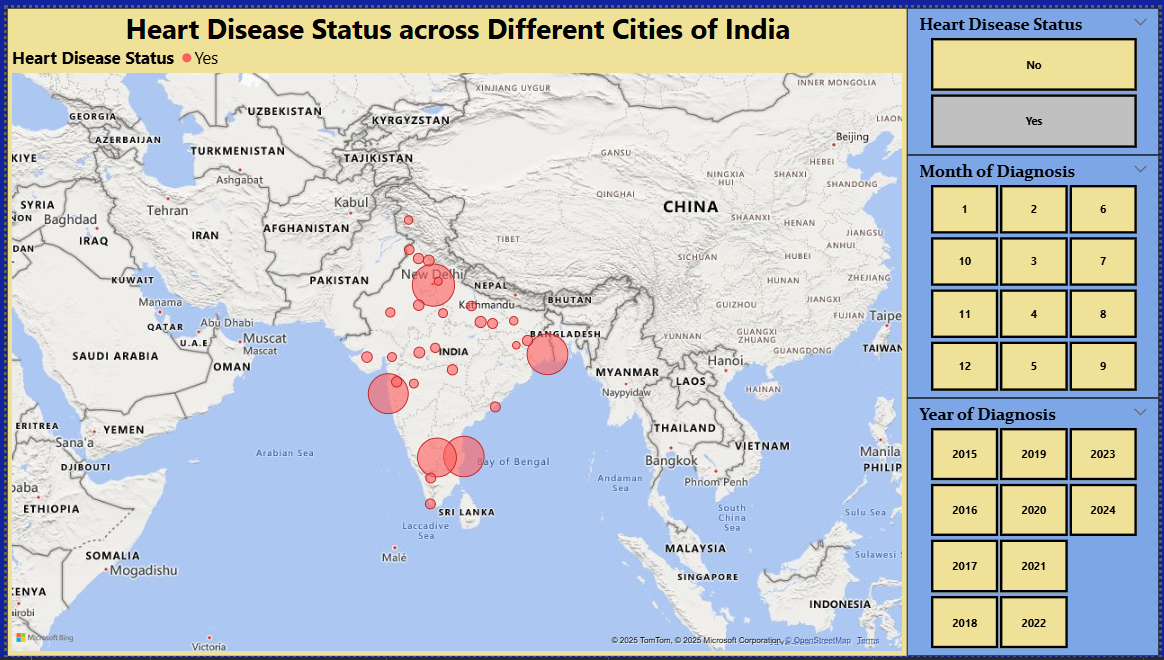
**Problem Statement 6:** How Does Heart Disease Prevalence Vary Across Different Cities in India?

**Solution:**

* Regional hotspots – Identification of cities with a high prevalence of heart disease.
* Time-based patterns – Trends in heart disease cases over months and years.
* Urban vs Rural impact – Comparison between metro cities and smaller towns.

**Insights:**

* Metro cities show a higher prevalence of heart disease, possibly due to lifestyle factors and pollution.
* Seasonal trends suggest higher cases in winter months, indicating climate as a potential contributing factor.
* Policies should target urban populations with awareness and screening programs.



**Conclusion**

The analysis of heart disease data in Power BI provided valuable insights into key risk factors such as age, BMI, lifestyle habits, stress, and geographic trends. The findings emphasize the importance of early intervention and preventive measures, especially for high-risk groups such as elderly individuals, smokers, obese individuals, and those with sedentary lifestyles.

By leveraging data-driven insights, healthcare providers and policymakers can design targeted interventions to reduce heart disease prevalence and improve public health outcomes. The study also highlights the need for further research using real-time and larger datasets to refine these observations further.