# Cardiovascular Disease Risk Factor Analysis

#### An interactive dashboard analysis

#### of the key drivers of cardiovascular disease

|  |  |
| --- | --- |
|  |  |
|  |  |
|  | **DONE BY :** S T SIVARANCHANI |
|  |  |
|  |  |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

1. **Introduction & Dataset Details**

* This report outlines the findings from an analysis of a patient dataset aimed at identifying key risk factors for cardiovascular disease (CVD). The analysis was conducted using Tableau to create a series of interactive dashboards for data exploration and insight generation.
* The dataset, cardio\_train.csv, contains records for 70,000 patients. It includes a

mix of demographic, physiological, and lifestyle data points.

**Key Data Fields:**

* **Demographic:** id, age (in days), gender (1: Female, 2: Male)
* **Physiological Measurements:** height (cm), weight (kg), ap\_hi (Systolic Blood Pressure), ap\_lo (Diastolic Blood Pressure)
* **Blood Markers:** cholesterol (1: Normal, 2: Above Normal, 3: Well Above Normal), gluc (Glucose Levels)
* **Lifestyle Factors:** smoke (binary), alco (binary for alcohol consumption), active (binary for physical activity)
* **Target Variable:** cardio (1: Presence of CVD, 0: Absence of CVD)

**Data Preparation:** Initial data preparation involved creating calculated fields for **Age in Years**, **Body Mass Index (BMI)**, and descriptive labels for categorical data (e.g., changing '1' to 'Female'). Data quality issues, such as physiologically impossible BMI values, were identified and filtered out to ensure the accuracy of the analysis.

**2.0 Project Objective**

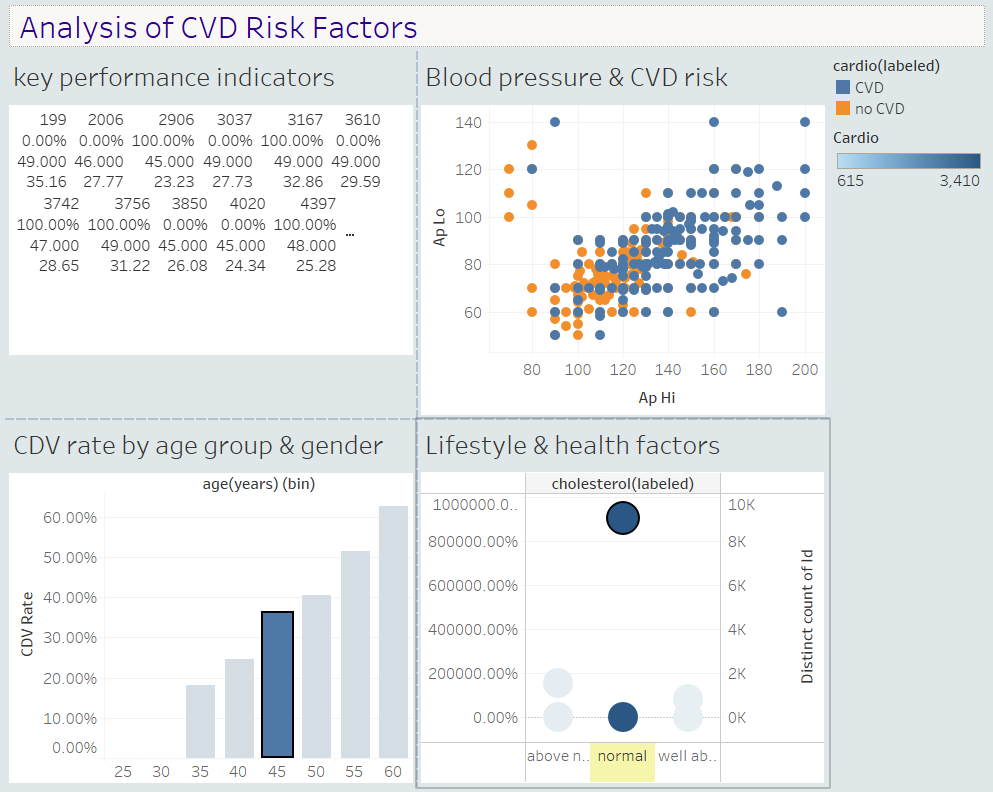
The primary objective of this project was to **identify and visualize the most significant risk factors contributing to cardiovascular disease**. The goal was to create a series of interactive dashboards in Tableau that would allow users to:

1. Gain a high-level overview of the patient population and CVD prevalence.
2. Explore the relationships between demographic, clinical, and lifestyle factors and the likelihood of having CVD.
3. Present these findings in a clear, compelling narrative through a Tableau Story.

**3.0 Work Performed: Dashboard Suite**

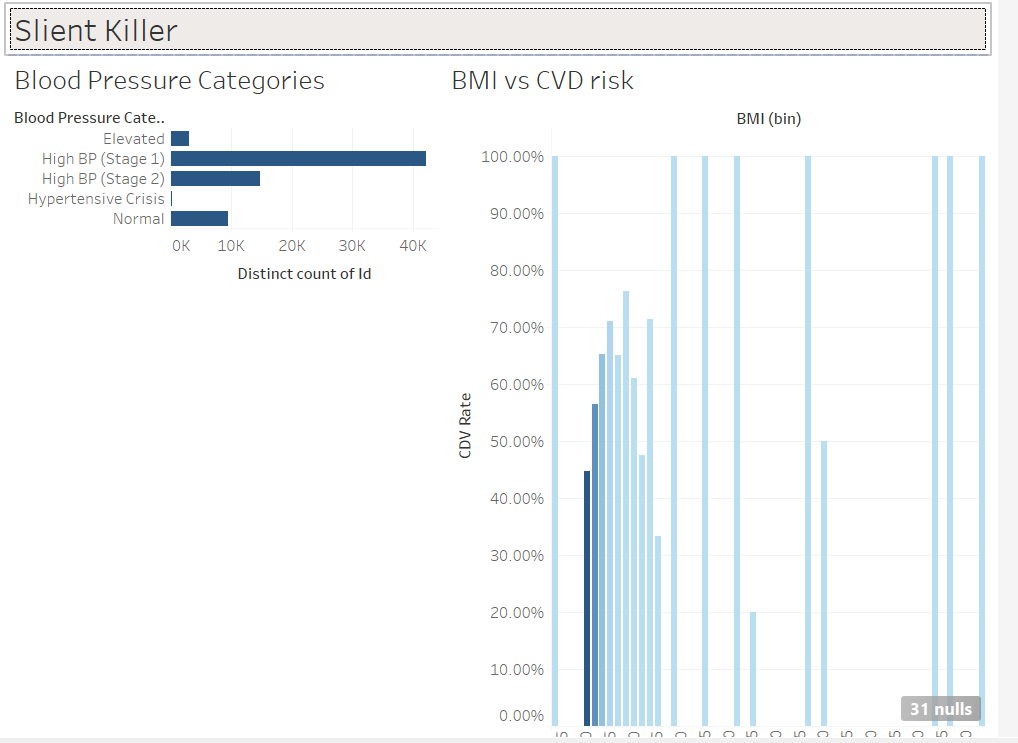
Three distinct dashboards were created to analyze the data from different perspectives, which were then compiled into a cohesive story.

* **Dashboard 1: Overall Risk Factor Analysis:** This served as the main dashboard, providing a comprehensive overview. It included KPIs for total patients and CVD rate, and charts visualizing the relationship between age, blood pressure, cholesterol, and CVD.



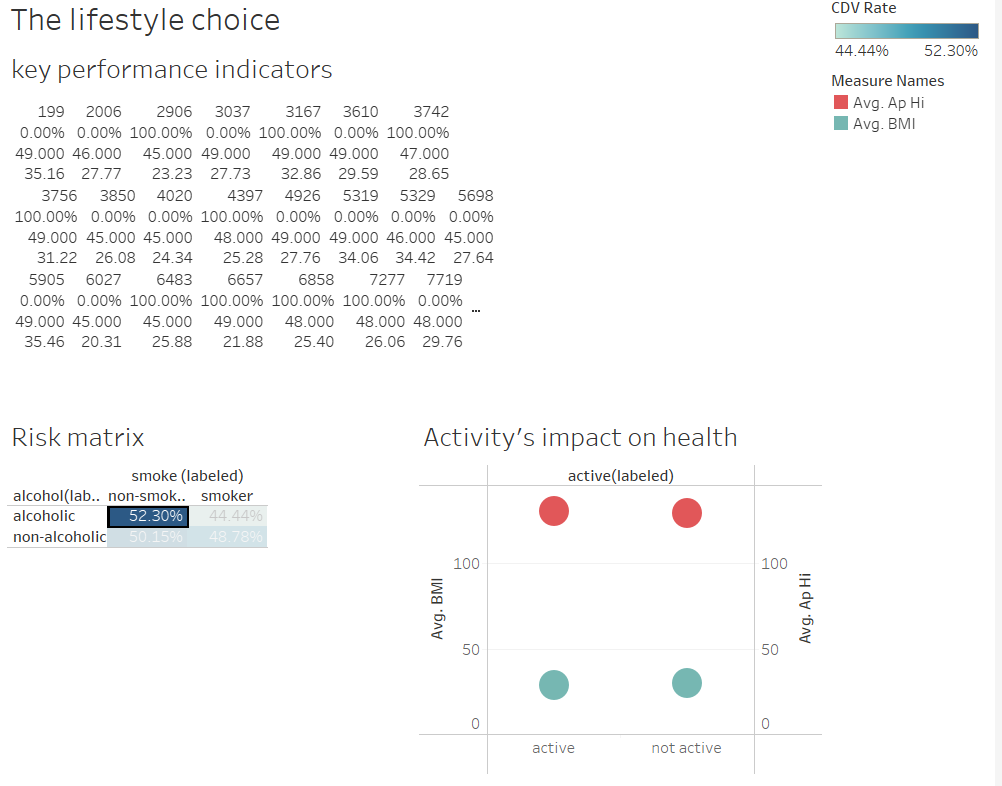
* **Dashboard 2: The "Silent Killers" Analysis:** This dashboard focused on

clinical measurements that are often asymptomatic but are strong predictors of CVD. It featured visualizations on the impact of Body Mass Index (BMI) and detailed Blood Pressure categories.



* **Dashboard 3: Lifestyle Choices & Their Impact:** This dashboard focused

on controllable lifestyle factors. It included a risk matrix for smoking and alcohol consumption and an analysis of how physical activity correlates with other health metrics.

* 

**4.0 Analysis & Inferences from Dashboards**

**Dashboard 1: Overall Risk Factor Analysis**

* **Age is a Primary Factor:** The analysis showed a clear and strong positive correlation between age and the prevalence of CVD. The CVD rate begins to increase significantly after the age of 45 and continues to rise in older age groups.
* **Blood Pressure is a Critical Indicator:** The scatter plot of systolic vs. diastolic blood pressure revealed a dense cluster of patients with CVD at higher blood pressure levels. This confirms that hypertension is a major risk factor.
* **Cholesterol's Clear Impact:** Patients with "Above Normal" and "Well Above Normal" cholesterol levels exhibited a substantially higher rate of cardiovascular disease compared to those with normal cholesterol.

**Dashboard 2: The "Silent Killers"**

* **BMI Correlates with Risk:** The analysis of BMI bins demonstrated a direct relationship between increasing BMI and a higher CVD rate. After filtering out data errors, the trend showed that overweight and obese individuals (BMI > 25) are at a significantly higher risk.
* **Hypertension Stages Show Stark Differences:** By categorizing blood pressure into clinical stages (Normal, Elevated, High BP Stage 1 & 2), it became evident that the proportion of patients with CVD increases dramatically with each successive stage of hypertension.

**Dashboard 3: Lifestyle Choices**

* **Smoking's Influence:** Smokers consistently showed a higher CVD rate than non-smokers across different demographic groups. The risk matrix suggested that this risk is compounded when combined with other factors like alcohol consumption.
* **Physical Inactivity is Detrimental:** The data indicated that inactive individuals, on average, have a higher BMI and higher systolic blood pressure compared to those who are physically active, linking a sedentary lifestyle to other known risk factors.

**5.0 Conclusion & Key Takeaways**

The analysis conducted in Tableau successfully visualized the primary drivers of cardiovascular disease within this dataset. The findings strongly align with established medical knowledge.

**The key conclusions are as follows:**

1. **Non-Modifiable Risks:** Age is the most significant non-modifiable risk factor, with CVD prevalence rising sharply in middle age.
2. **Clinical Indicators are Paramount:** High blood pressure, elevated cholesterol, and a high BMI are the strongest clinical predictors of cardiovascular disease. The dashboards effectively demonstrated that as these metrics worsen, the risk of CVD increases dramatically.
3. **Lifestyle Matters:** Controllable lifestyle choices, particularly smoking and physical inactivity, are significant contributing factors that exacerbate the risk posed by physiological conditions.

This project underscores the power of interactive data visualization to not only confirm existing hypotheses but also to explore the complex interplay between various health factors. The created dashboards serve as an effective tool for understanding and communicating the multifaceted nature of cardiovascular disease risk.