DATA ANALYTICS WITH POWER BI PROJECT REPORT

(Project Semester October 2025 – January 2026)

Heart Diseases Prediction

Submitted by
Himanshu Yadav
Registration No: 12325588
Program P132: B. Tech Section: K23MR
Course Code: INT-374

Under the Guidance of

B Monica (32352)

Discipline of CSE/IT

Lovely School of Computer Science Engineering

Lovely Professional University, Phagwara

DECLARATION

I, Himanshu Yadav, student of B.Tech CSE under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date: 18.12.25 Signature

Registration No. 12325588 Himanshu Yadav

CERTIFICATE

This is to certify that Himanshu Yadav bearing Registration no. 12325588 has completed INT-374 project titled, **"*Heart Diseases Prediction*"** under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort and study.

Signature and Name of the Supervisor
Designation of the Supervisor
School of School of Computer Science Engineering
Lovely Professional University
Phagwara, Punjab.

Date: 18.12.25

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Lovely Professional University for providing the opportunity and platform to undertake this project.

I am especially thankful to Ma’am B Monica for her invaluable guidance, constant support, and insightful feedback throughout the course of this project. His mentorship played a crucial role in shaping the direction and success of my work. I also wish to thank my faculty mentors, peers, and friends for their encouragement and helpful suggestions during the project.

This experience has significantly enhanced my technical, analytical, and visualization skills, particularly in working with dataset. Finally, I extend my heartfelt thanks to everyone who directly or indirectly contributed to the successful completion of this project.

TABLE OF CONTENT

1. INTRODUCTION

Cardiovascular diseases, particularly heart disease, remain one of the leading causes of death worldwide. Rapid lifestyle changes, increased stress levels, unhealthy dietary habits, and lack of physical activity have significantly contributed to the rising prevalence of heart-related conditions across different age groups. Early identification of risk factors and timely diagnosis play a critical role in preventing severe cardiac events and improving patient outcomes.

Heart disease is influenced by multiple demographic and clinical factors such as age, gender, chest pain type, maximum heart rate, and other physiological indicators. Understanding how these variables interact and contribute to heart disease risk is essential for healthcare professionals, researchers, and policymakers. However, medical datasets often contain complex and high-dimensional data, making it difficult to extract meaningful insights using traditional tabular analysis.

In this context, data-driven analytics provides a powerful approach to explore patient health data and uncover hidden patterns associated with heart disease. Visual analytics, in particular, enables intuitive understanding of trends, correlations, and distributions that may not be apparent through raw data alone.

This project focuses on developing an interactive Heart Disease Prediction Dashboard using Microsoft Power BI to analyze patient health records. The dashboard transforms raw clinical data into clear and interactive visualizations, enabling users to explore heart disease patterns across age groups, gender, chest pain types, and heart rate levels. Rather than focusing on complex predictive modeling, this project emphasizes descriptive and diagnostic analytics to support better understanding of cardiovascular health indicators.

By leveraging Power BI's data modeling capabilities, DAX measures, and interactive visual design, the project demonstrates how business intelligence tools can assist in healthcare analytics, early risk identification, and data-informed medical decision-making.

2. PROBLEM STATEMENT

Despite the availability of large volumes of clinical and patient health data, healthcare professionals often face challenges in effectively analyzing and interpreting this information to identify heart disease risk patterns. Traditional analysis methods using static tables or spreadsheets make it difficult to compare multiple variables such as age, gender, heart condition, and chest pain types simultaneously.

There is a need for an interactive and user-friendly analytical solution that can consolidate patient data and visually represent key heart disease indicators in a meaningful way. Without such tools, critical insights related to age-wise risk distribution, gender differences, and symptom-based patterns may remain underutilized, limiting their impact on preventive healthcare strategies.

This project addresses this problem by designing an interactive Power BI dashboard that enables dynamic exploration of heart disease data, helping users quickly identify trends, correlations, and high-risk groups through visual analytics.

3. OBJECTIVES

The primary objective of this project is to analyze heart disease patterns using an interactive Power BI dashboard. The project aims to study how heart disease occurrence varies across key demographic and clinical factors such as age, gender, chest pain type, and maximum heart rate.

Another important objective is to compare patients with and without heart disease by visualizing their distribution and identifying differences across age groups and gender categories. This comparison helps in understanding high-risk population segments and common characteristics associated with heart disease.

The project also seeks to analyze age-based trends by grouping patients into age bins to identify age ranges with a higher prevalence of heart disease. This enables a clearer understanding of how heart disease risk changes across different stages of life.

Additionally, the project aims to examine the relationship between age and maximum heart rate using scatter-based analysis, helping to identify patterns and correlations linked to heart health conditions.

A further objective of this project is to analyze different chest pain types and determine their association with heart disease occurrence, highlighting symptom patterns that may indicate higher risk.

Finally, the project aims to demonstrate the effective use of Power BI as a healthcare analytics and business intelligence tool. This includes data cleaning and transformation, creation of DAX measures, and the design of an interactive, professional dashboard that emphasizes clarity, usability, and data-driven insights for healthcare decision-making.

4. SOURCE OF DATASET

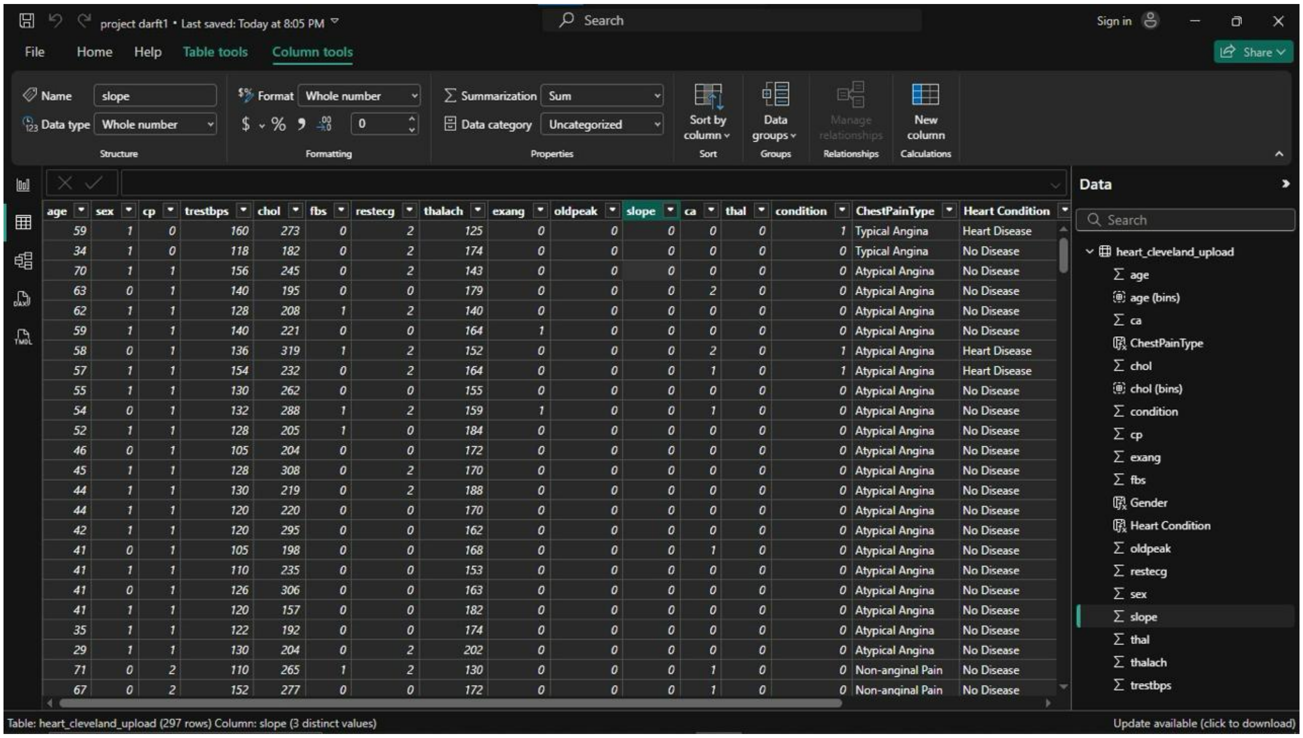
The dataset used for this project was obtained from UNData (undata.org), an open data platform maintained by the United Nations that provides reliable and standardized global datasets across various domains, including health and population statistics. The dataset contains structured records related to patient health indicators relevant to heart disease analysis.

The dataset includes both categorical and numerical variables, making it suitable for interactive dashboard development and healthcare analytics. Key attributes include patient age, gender, chest pain type, maximum heart rate, and heart disease condition status. These variables allow for multi-dimensional analysis of demographic factors and clinical indicators associated with heart disease.

Prior to analysis, the dataset was imported into Microsoft Power BI and checked for data quality, consistency, and completeness. Necessary data cleaning and transformation steps were performed using Power Query, including handling missing values and ensuring appropriate data types for accurate aggregation and visualization.

The structured nature of the dataset supports comprehensive analytical exploration, enabling comparisons across age groups, gender categories, and clinical conditions. This makes the dataset well-suited for descriptive and diagnostic analysis of heart disease patterns using interactive Power BI dashboards.

5. DATA PREPARATION AND TRANSFORMATION



age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	condition	ChestPainType	Heart Condition
59	1	0	160	273	0	2	125	0	0	0	0	0	1	Typical Angina	Heart Disease
34	1	0	118	182	0	2	174	0	0	0	0	0	0	Typical Angina	No Disease
70	1	1	156	245	0	2	143	0	0	0	0	0	0	Atypical Angina	No Disease
63	0	1	140	195	0	0	179	0	0	0	2	0	0	Atypical Angina	No Disease
62	1	1	128	208	1	2	140	0	0	0	0	0	0	Atypical Angina	No Disease
59	1	1	140	221	0	0	164	1	0	0	0	0	0	Atypical Angina	No Disease
58	0	1	136	319	1	2	152	0	0	0	2	0	1	Atypical Angina	Heart Disease
57	1	1	154	232	0	2	164	0	0	0	1	0	1	Atypical Angina	Heart Disease
55	1	1	130	262	0	0	155	0	0	0	0	0	0	Atypical Angina	No Disease
54	0	1	132	288	1	2	159	1	0	0	1	0	0	Atypical Angina	No Disease
52	1	1	128	205	1	0	184	0	0	0	0	0	0	Atypical Angina	No Disease
46	0	1	105	204	0	0	172	0	0	0	0	0	0	Atypical Angina	No Disease
45	1	1	128	308	0	2	170	0	0	0	0	0	0	Atypical Angina	No Disease
44	1	1	130	219	0	2	188	0	0	0	0	0	0	Atypical Angina	No Disease
44	1	1	120	220	0	0	170	0	0	0	0	0	0	Atypical Angina	No Disease
42	1	1	120	295	0	0	162	0	0	0	0	0	0	Atypical Angina	No Disease
41	0	1	105	198	0	0	168	0	0	0	1	0	0	Atypical Angina	No Disease
41	1	1	110	235	0	0	153	0	0	0	0	0	0	Atypical Angina	No Disease
41	0	1	126	306	0	0	163	0	0	0	0	0	0	Atypical Angina	No Disease
41	1	1	120	157	0	0	182	0	0	0	0	0	0	Atypical Angina	No Disease
35	1	1	122	192	0	0	174	0	0	0	0	0	0	Atypical Angina	No Disease
29	1	1	130	204	0	2	202	0	0	0	0	0	0	Atypical Angina	No Disease
71	0	2	110	265	1	2	130	0	0	0	1	0	0	Non-anginal Pain	No Disease
67	0	2	152	277	0	0	172	0	0	0	1	0	0	Non-anginal Pain	No Disease

Data preparation and transformation were carried out entirely within Microsoft Power BI using the Power Query Editor. The initial step involved examining the dataset to understand column structures, data types, and the presence of missing or inconsistent values. Columns that were not relevant to the analytical objectives were removed to simplify the data model and improve dashboard performance.

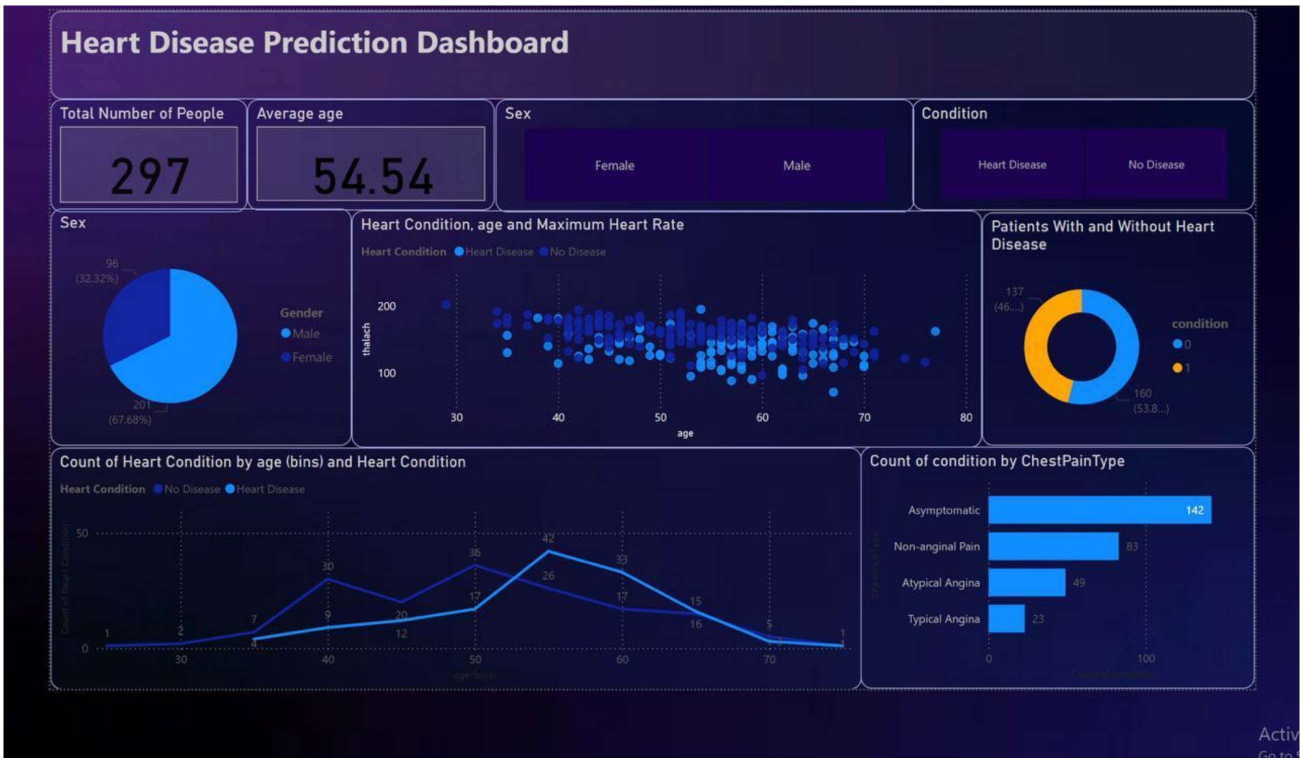
Missing values in critical fields such as age and maximum heart rate were reviewed and handled appropriately to ensure accurate analysis and aggregation. Data types were explicitly assigned, with numerical variables treated as numeric values and categorical attributes such as gender, chest pain type, and heart disease condition formatted as text.

Column names were standardized to improve readability and ensure consistency across visuals and DAX calculations. Text-based fields were cleaned using trimming and formatting operations to eliminate inconsistencies caused by spacing or formatting variations.

Once the data cleaning process was completed, the dataset was loaded into the Power BI data model. Calculated measures were created using DAX to compute key metrics such as total number of patients, average age, and condition-wise patient counts. This clean and structured dataset formed a reliable foundation for building an interactive and responsive heart disease analytics dashboard.

6. DASHBOARD ANALYSIS AND IMPLEMENTATION

Overall Heart Disease Analysis Overview



The primary deliverable of this project is an interactive and professionally designed Power BI dashboard developed to analyze heart disease patterns within patient data. The dashboard serves as a centralized analytical interface that consolidates clinical and demographic information into a clear and visually intuitive format. Its design focuses on enabling users to explore heart disease indicators efficiently while maintaining clarity and ease of interpretation across all analytical views.

The dashboard was created with the objective of transforming raw patient health data into meaningful insights that support healthcare understanding and data-driven analysis. Instead of relying on static tables or spreadsheets, the dashboard uses interactive visualizations to reveal patterns related to age distribution, gender differences, heart condition status, chest pain types, and maximum heart rate behavior.

Rather than focusing on complex predictive or machine learning models, the dashboard emphasizes descriptive and diagnostic analytics to clearly explain how heart disease occurrence varies across different demographic and clinical factors. By analyzing historical patient data, the dashboard answers practical questions such as which age groups are more affected, how heart disease differs between genders, and which chest pain types are commonly associated with heart conditions. This ensures that insights remain transparent, interpretable, and applicable in real-world healthcare analytics contexts.

KPI CARDS

The dashboard follows a structured and hierarchical layout that prioritizes readability, consistency, and ease of interaction. Visual elements are arranged logically to guide users from high-level summaries to more detailed analytical views. This structured design allows users to quickly understand the overall health dataset before exploring specific patterns and relationships.

At the top of the dashboard, Key Performance Indicator (KPI) cards provide a concise overview of the dataset. These KPIs display total number of patients, average patient age, and the distribution of patients with and without heart disease. By presenting these metrics prominently, the dashboard offers an immediate snapshot of the overall health profile without requiring detailed chart interpretation.

All KPI values are fully dynamic and interactive, updating automatically based on user-selected filters such as gender and heart disease condition. This interactivity ensures that the KPIs remain contextually accurate during exploratory analysis and allow users to compare different patient segments in real time. Overall Heart Disease Analysis Overview

The primary deliverable of this project is an interactive and professionally designed Power BI dashboard developed to analyze heart disease patterns within patient data. The dashboard serves as a centralized analytical interface that consolidates clinical and demographic information into a clear and visually intuitive format. Its design focuses on enabling users to explore heart disease indicators efficiently while maintaining clarity and ease of interpretation across all analytical views.

The dashboard was created with the objective of transforming raw patient health data into meaningful insights that support healthcare understanding and data-driven analysis. Instead of relying on static tables or spreadsheets, the dashboard uses interactive visualizations to reveal patterns related to age distribution, gender differences, heart condition status, chest pain types, and maximum heart rate behavior.

Rather than focusing on complex predictive or machine learning models, the dashboard emphasizes descriptive and diagnostic analytics to clearly explain how heart disease occurrence varies across different demographic and clinical factors. By analyzing historical patient data, the dashboard answers practical questions such as which age groups are more affected, how heart disease differs between genders, and which chest pain types are commonly associated with heart conditions. This ensures that insights remain transparent, interpretable, and applicable in real-world healthcare analytics contexts.

7. KEY INSIGHTS & FINDINGS

1. Age as a Significant Risk Factor

The analysis reveals that heart disease prevalence increases with age, particularly in middle-aged and older patient groups. Age-based trends clearly show higher concentrations of heart disease cases in higher age bins.

2. Gender-wise Differences

The dashboard indicates a higher proportion of heart disease cases among male patients compared to female patients, highlighting gender as an important demographic factor in heart disease analysis.

3. Chest Pain Type Association

Certain chest pain types are more frequently associated with heart disease cases, suggesting that chest pain classification plays a critical role in identifying potential cardiac risk.

4. Maximum Heart Rate Patterns

Scatter plot analysis between age and maximum heart rate shows noticeable variation between patients with and without heart disease. Patients diagnosed with heart disease tend to achieve lower maximum heart rates, especially in older age groups.

5. Condition-wise Distribution

The comparison between patients with and without heart disease provides a clear understanding of overall disease prevalence within the dataset and supports high-level risk assessment.

6. Effectiveness of Interactive Filters

Dynamic slicers allow focused analysis of specific patient segments, enabling quick identification of patterns across gender and condition categories.

8. CONCLUSION

This project successfully demonstrates the use of Power BI as an effective healthcare analytics tool for analyzing heart disease patterns using patient health data. By transforming raw clinical records into interactive visualizations, the dashboard enables clear understanding of how demographic and physiological factors are associated with heart disease occurrence.

The analysis highlights the importance of age, gender, chest pain type, and maximum heart rate in understanding heart disease risk. Through descriptive and diagnostic analytics, the dashboard allows users to explore condition-wise patterns without relying on complex predictive models. The interactive nature of the dashboard ensures that insights can be dynamically explored, making the analysis more intuitive and impactful compared to traditional static reports.

The project also showcases strong data preparation practices, including data cleaning, transformation, and DAX-based measure creation. The structured layout, dynamic KPI cards, and well-organized visuals ensure clarity, usability, and professional presentation. Overall, this dashboard serves as a practical example of how business intelligence tools can support healthcare analysis, early risk identification, and data-driven understanding of cardiovascular health.

9. FUTURE SCOPE

1. **Integration of Predictive Models**

Machine learning algorithms such as Logistic Regression, Decision Trees, or Random Forest can be integrated to predict heart disease risk and enhance the dashboard from descriptive analytics to predictive analytics.

2. **Inclusion of Additional Clinical Features**

Incorporating variables such as blood pressure, cholesterol levels, blood sugar, BMI, and lifestyle indicators (smoking, physical activity) would provide deeper and more accurate insights.

3. **Real-Time Data Integration**

The dashboard can be extended to support real-time or near real-time patient data from hospital systems or wearable health devices for continuous monitoring.

4. **Doctor-Focused Decision Support**

Advanced drill-through features can be added to create patient-level views that assist healthcare professionals in clinical decision-making.

5. **Geographical and Population Analysis**

If location-based data is available, regional analysis can help identify high-risk populations and support public health planning.

6. **Deployment on Power BI Service**

Publishing the dashboard to Power BI Service with role-based access control would enable secure sharing among healthcare analysts and decision-makers.

7. **Scalability to Other Diseases**

The analytical framework developed in this project can be reused for studying other diseases such as diabetes, hypertension, or neurological disorders.

10. LINKS

Linkedin- https://www.linkedin.com/posts/himanshu-yadav

Github- https://github.com/Himanshukalaliya/heart-desease-project-power-bi