

Problem Statement Worksheet (Hypothesis Formation)

By the end of 2025, determine the impact of age, gender, and BMI on the likelihood of developing heart disease, focusing on individuals within the age brackets of 20–29, 30–39, 40–49, 50–59, and 60–69. **H**

The goal is to identify key risk factors across these age groups and reduce heart disease incidence by 10%, particularly in individuals aged 50 and above.

1 Context

Heart disease remains a global leading cause of death, necessitating the identification of risk factors for targeted interventions. This study explores how age, gender, and BMI affect heart disease risk in individuals aged 20 to 69, with a focus on those 50 and above. By examining patterns across age groups, the goal is to inform strategies that reduce heart disease incidence by 10% by the end of 2025.

2 Criteria for success

- Identification of significant correlations between age, gender, and BMI with heart disease.
- Development of a predictive model for heart disease risk in individuals aged 50 and above.
- A reduction in heart disease incidence by 10% in the target population by the end of 2025.
- Actionable insights for healthcare practitioners to implement preventative measures.

3 Scope of solution space

- Exploratory data analysis (EDA) of age, gender, and BMI to uncover patterns linked to heart disease.
- Machine learning models for predictive risk analysis.
- Statistical validation to confirm significance of identified factors.
- Application of insights to develop targeted intervention strategies for at-risk populations.

4 Constraints within solution space

1. Data completeness: Ensure sufficient data quality, addressing missing values or inconsistencies.
2. The results may be limited to the dataset population and require further validation for wider applicability.
3. Time constraints: All analysis and insights must be ready by the end of 2025.
4. Ethical considerations: Protecting personal health information (PHI) in data handling.

5 Stakeholders to provide key insight

1. Healthcare professionals: Provide clinical insights on heart disease risk factors.
2. Data Analysts and Data scientists: Offer expertise in robust data and statistical analysis, predictive modeling.
3. Public health officials: Assist in translating findings into policy or community-based interventions.
4. Patients and advocacy groups: Provide qualitative data on risk behaviors and lifestyles.

6 Key data sources

- diabetes_prediction_dataset