Research Report on Healthcare Analytics

(Using BRFSS 2015 Dataset to Predict Heart Disease)

Introduction to Healthcare Analytics:-

Healthcare analytics involves the use of data analysis tools and techniques to extract meaningful insights from healthcare data. It helps identify trends, support clinical decisions, improve patient care, and reduce healthcare costs.

In this project, we analyze behavioral, demographic, and health data from the **Behavioral Risk Factor Surveillance System (BRFSS) 2015** to detect patterns that contribute to **heart disease**.

Objectives of This Project:-

- To understand **which health indicators** (e.g., smoking, BMI, high blood pressure) contribute most to heart disease
- To use **Python and data science tools** to explore and clean the dataset
- To build a **predictive model** that can classify individuals at risk of heart disease

---

About the Dataset:-

- **Source**: BRFSS 2015 (Behavioral Risk Factor Surveillance System)
- **Size**: Over 250,000 records
- **Features**: Includes indicators like smoking, drinking, BMI, physical activity, stroke history, mental and physical health, income level, and more
- **Target Variable**: HeartDiseaseorAttack – whether the individual has ever been told they had a heart attack or heart disease

---

How Healthcare Analytics Helps:-

1. **Risk Identification**: Analytics helps detect patterns in behavior (like smoking or lack of physical activity) that increase heart disease risk.
2. **Early Intervention**: Individuals at risk can be targeted for early screenings or lifestyle interventions.

3. **Personalized Care**: Insights allow health professionals to tailor treatments to specific risk profiles.
4. **Resource Optimization**: Hospitals and healthcare providers can better allocate resources by predicting patient needs.

Tools & Techniques Used:-

Area	Tool / Library
Data Cleaning	Pandas
Exploratory Analysis	Seaborn, Matplotlib
Feature Selection	Correlation, Heatmaps
Predictive Modeling	Logistic Regression, Random Forest (Scikit-learn)
Evaluation Metrics	Accuracy, ROC-AUC, Confusion Matrix

Example Insights You Can Extract:-

- Smokers and physically inactive individuals have a higher correlation with heart disease.
- Higher BMI and poor mental health indicators also show strong associations.
- Income levels and age groups reflect disparities in heart disease prevalence.

Challenges in Healthcare Analytics:-

- **Data Imbalance**: Heart disease occurrences may be underrepresented in general populations.
- **Missing or Inconsistent Values**: Need careful preprocessing.
- **Interpreting Causal vs. Correlational Data**: Requires domain understanding.
- **Ethical Use of Data**: Privacy and fairness must be maintained when working with patient data.

Conclusion:-

This project exemplifies how **healthcare analytics**, when applied using Python and real-world health data, can reveal crucial risk factors for heart disease. These insights not only guide better clinical decisions but also support public health interventions to reduce heart-related mortality.

With further modeling and refinement, this system could potentially act as a decision-support tool in preventive cardiology.