Healthcare Risk Prediction - Business Report

1. Executive Summary

This report outlines the findings and outcomes of a healthcare capstone project aimed at predicting patient health risks using data-driven insights. The project leveraged historical healthcare data to develop a risk stratification model and visualize critical metrics through an interactive Power BI dashboard. The ultimate goal is to improve patient care and optimize resource allocation.

2. Project Objective

The primary objective was to predict patients at high risk of developing critical health conditions and identify patterns contributing to elevated health risks. By understanding these factors, healthcare providers can take proactive measures to reduce complications, hospital readmissions, and associated costs.

3. Data Overview

- **Source**: The dataset comprises patient demographics, visit history, chronic disease flags, medication usage, and health outcomes.
- **Volume**: Tens of thousands of patient records with multiple associated features.
- **Key Features**:
    - Age, Gender
    - Number of doctor visits
    - Number of chronic diseases
    - Risk level (target variable)
    - Medical spending and utilization

4. Methodology

- **Data Preparation**: Performed using SQL for data cleaning, normalization, and feature engineering.
- **Risk Classification**: Patients were classified into three categories — **Low Risk**, **Medium Risk**, and **High Risk** based on clinical indicators and historical health usage.
- **Visualization**: Power BI dashboards were used to derive actionable insights with interactive visuals covering:
    - Risk distribution
    - Age and gender demographics
    - Cost analysis by risk group

- Visit patterns
- Chronic condition prevalence

5. Key Findings

➤ Risk Stratification

- **High-risk group** represents **15.9%** of total patients but contributes to **over 42%** of the total medical expenditure.
- **Low-risk group** comprises **57%** of the population but accounts for only **18%** of the cost.

➤ Age Distribution

- Risk level increases significantly with age:
  - Patients aged **60+** are disproportionately in the **High-Risk** category.
  - **Younger patients (below 40)** dominate the **Low-Risk** category.

➤ Chronic Conditions

- Number of chronic conditions is a major driver of risk:
  - Patients with **3+ chronic diseases** have an 80% chance of being in the **High-Risk** category.

➤ Visit and Cost Correlation

- Frequent visits correlate with higher risk, especially in patients with chronic illness.
- A small fraction of patients with very high visit counts drive a large share of costs.

➤ Gender Insights

- Gender disparity is minimal in risk levels, although **males** showed slightly higher representation in the high-risk segment.

6. Business Insights & Recommendations

- **Preventive Care Programs**
  Introduce targeted preventive programs for patients over 50 and those with early signs of chronic conditions to reduce future hospitalization risks.
- **Chronic Disease Management**
  Establish specialized care plans for patients with multiple chronic diseases—especially those with 3 or more—since they have the highest predicted risk.
- **Cost Optimization Strategies**
  Focus on optimizing resource allocation toward the high-risk 15.9%, as they contribute to over 40% of healthcare costs.

- **Patient Education & Engagement**
  Increase engagement efforts in the medium-risk population, who may transition to high-risk without intervention.
- **Predictive Monitoring Integration**
  Integrate risk scores with EHR systems to provide real-time alerts for clinicians during patient visits.

7. Conclusion

The healthcare risk prediction model successfully identifies high-risk individuals using demographic and clinical features. By leveraging these insights through Power BI dashboards, healthcare administrators can make data-driven decisions to enhance patient outcomes, reduce costs, and allocate resources more efficiently.