CarePulse HDHI – Executive Summary

Title:

CarePulse HDHI
A Comprehensive Data Analytics System for Patient Outcomes and Environmental Risk at Hero DMC Heart Institute

Executive Overview

Background

Hero DMC Heart Institute (HDHI) stands as a key cardiac institution in India. As the healthcare landscape modernizes, there is a pressing need to leverage internal patient data to improve care quality, risk forecasting, and clinical planning.

Project Objective

To design a high-impact, full-stack data analytics system that combines demographic, clinical, and environmental information to:

- Monitor patient inflow and health trends
- Predict high-risk and long-stay patients
- Assess the impact of environmental exposure (e.g., pollution)
- Support evidence-based hospital decisions

Tools & Technologies

- Power BI, Python, Pandas, DAX, SQL, Sklearn, XGBoost

Data Foundation

- Real data from HDHI including admissions, lab results, mortality, and pollution
- Primary datasets used:
    - master_hospital_data.csv

- predicted_outcomes.csv

Methodology & Execution

Data Engineering & Processing (Python)

- Cleaned and enriched patient data
- Calculated key variables: duration_of_stay, age_bucket, pollution overlays
- Created flags for comorbidities, lab anomalies, and admission types

Machine Learning Workflow

- Developed predictive models using XGBoost
- Predicted:
  - Probability of extended stay
  - Mortality risk
- Clustered patient segments using KMeans
- Built a recommendation engine using cosine similarity for care path suggestions

Power BI Reporting

Built an interactive, 3-page dashboard:

1. **Patient Overview & Demographics** – Admissions, gender/age splits, LOS
2. **Risk & Outcome Insights** – LOS flags, model results, risk trends
3. **Clinical & Environmental Impact** – Pollution effects, lab indicators, mortality

---

Insights & Results

Admission Trends

- Over 16,000 patient records
- 61–80 age group accounted for 40%+ of admissions
- Seasonal peaks linked to winter months and respiratory patterns

Risk & Prediction Outcomes

- ~29% flagged as Likely Long Stay
- Heart failure, low hemoglobin, and high creatinine associated with higher mortality
- Extended LOS correlated with higher glucose and platelet fluctuations

Environmental Risk Findings

- PM2.5 and NO2 levels showed strong positive correlation with patient mortality
- Rural patients had longer LOS and higher predicted risks
- Air quality degradation aligned with spikes in emergency admissions

---

Strategic Recommendations

For Hospital Leadership

- Allocate additional ICU or ward resources for high-risk patient windows
- Integrate this dashboard into monthly operations review
- Use pollution alerts as indirect signals to prepare for admission surges

For Clinical Decision-Makers

- Prioritize lab tests like hemoglobin, glucose, and platelets upon admission
- Monitor elderly and rural patients using model-based early warning flags

- Extend chronic care planning for heart failure and kidney disease cases

For Public Health & Policy

- Create region-wise interventions during high-pollution periods
- Promote cross-hospital adoption of data-driven care models
- Fund environmental-health outcome research based on findings

---