Executive Summary

1. Introduction

This report summarizes the key activities, findings, and outcomes of the Healthcare Analytics Internship conducted at **Curvelogics**, with a real-world analytics project executed for the client **CareNova**, a healthcare services provider.

The primary objective of the internship was to transform raw healthcare data into actionable insights through a structured analytics workflow involving **Excel, SQL Server, Python, Power Query, and Power BI**.

The project focused on understanding patient treatment patterns, costs, outcomes, and operational performance to support data-driven decision-making for CareNova.

2. Project Scope & Objectives

The internship aimed to deliver an end-to-end analytics solution based on the following goals:

Key Objectives

- Clean and prepare multi-source healthcare treatment data.
- Build a structured SQL model for aggregations, KPIs, and advanced analysis.
- Develop interactive dashboards showcasing trends & insights.
- Identify operational inefficiencies, cost variations, and performance gaps.
- Present findings that support CareNova's clinical & administrative decision-making

Business Need for CareNova

CareNova required a centralized analytics framework to:

- Track patient outcomes
- Monitor costs across treatments & departments
- Identify inefficiencies in treatment duration
- Analyze doctor & department performance
- Improve patient satisfaction and operational throughput

3. **Data Overview**

The analysis was conducted using three key datasets:

- **PatientInfo** – demographics & region
- **DoctorDetails** – doctors, specialties, experience
- **TreatmentRecords** – dates, outcomes, costs, duration, feedback

Data Challenges

- Missing or incomplete treatment fields

- Inconsistent date formats

- Duplicate patient/treatment entries

- Unstandardized categorical values

- Variations in treatment duration calculation

Cleaning and transformation were completed using **Power Query** and **SQL-based logic**.

4. **Methodology**

A structured, phase-wise approach was used:

Phase 1 – Understanding Requirements

Reviewed datasets, business needs, KPI definitions, stakeholder expectations and formulated an initial plan for execution.

Phase 2 – Data Cleaning

Fixed missing values, corrected formats, standardized categories and removed duplicates.

Phase 3 – Data Transformation

Created calculated columns for:

- Age groups

- Cost groups

- Treatment duration

- Recovery indicators

- Monthly/yearly trends

Phase 4 – <u>Exploratory Analysis</u>

Identified early patterns related to:

- High-cost treatments
- Outcome distribution
- Admission cycles
- Department performance

Phase 5 – <u>SQL Analytics</u>

Built complex queries using:

- CTEs for multi-level aggregations
- Treatment duration comparisons
- Readmission & cost variation checks
- Trend-based KPI models

Phase 6 – <u>Dashboard Development</u>

Created Power BI dashboards featuring:

- Admissions & outcome trends
- Cost analytics
- Doctor & department performance
- Patient satisfaction scoring
- Duration-based cluster analysis

Phase 7 – <u>Validation</u>

Cross-checked accuracy, reconciled values, and corrected inconsistencies.

Phase 8 – <u>Reporting</u>

Compiled insights, summaries, and visual dashboards for stakeholders.

5. Key Insights (Actionable Findings)

1. ## Variation in Treatment Costs

    - Significant cost variation between departments and treatment types.

    - High-cost procedures often did not correlate with better outcomes.

    - **Action:** Standardize cost guidelines; monitor outliers monthly.

2. ## Outcome Performance Gaps

    - Some specialties recorded consistently higher recovery rates.

    - Certain departments showed lower satisfaction despite shorter durations.

    - **Action:** Investigate operational gaps, workflows, and doctor allocations.

3. ## Treatment Duration as a Performance Driver

    - Longer treatment durations were linked to lower satisfaction scores.

    - Certain treatments deviated heavily from the average expected duration.

    - **Action:** Introduce duration benchmarking and time-bound protocols.

4. ## Seasonality in Admissions

    - Monthly admissions showed noticeable seasonal peaks.

    - **Action:** Adjust staffing and resource allocation in high-traffic months.

5. ## Department & Doctor-Level Variation

    - A few doctors had significantly higher throughput with similar outcomes.

    - **Action:** Share best practices across teams or departments.

6. ## Patient Satisfaction Drivers

    Higher satisfaction was strongly linked to:

    - Shorter duration
    - Lower cost
    - Successful outcomes
    - **Action:** Develop a patient-experience improvement framework.

6. Impact & Business Value

The analytics model developed during the internship provides multiple benefits:

- **Clear visibility** of operational performance across departments
- **Standardized KPIs** for cost, outcome, and duration
- **Better decision-making** using interactive dashboards
- **Proactive identification** of inefficiencies
- **Improved forecasting** using trend and seasonal data
- **Foundation for predictive analytics**, such as readmission likelihood

For CareNova, these insights can support:

- Optimized resource planning
- Reduced unnecessary cost variation
- Enhanced patient outcomes
- Improved satisfaction and service quality

7. Skills Developed

- **Technical:** SQL Server, Python, Power BI, Power Query, DAX, Data modelling.
- **Analytical:** KPI formulation, cluster analysis, trend analysis
- **Professional:** Reporting, documentation, communication, presentation
- **Domain:** Understanding of healthcare metrics, treatment workflows

8. Conclusion

This internship provided practical exposure to real-world healthcare analytics. The project helped transform raw hospital data into meaningful insights that can guide strategic improvements for CareNova.

The combination of SQL modelling, Power BI dashboards, and structured analytics strengthened both technical and analytical expertise while contributing measurable value to the client.

The outcomes of this project lay a strong foundation for deeper predictive and prescriptive analytics in the future.