



NSBM Green University

Faculty of Computing

BSc (Hons) Data Science

DS403.3- Big Data Programming

Project Proposal

Group - 2

Student ID	Student Name
24490	MRK Karunathilaka
24572	MKIM Rohana
24614	GAAS Ganegoda

Module Lecturer: Mr. Adhil Rushdy

Health Insurance Analytics System Using Lambda Architecture on Azure

1. Problem Statement

- Most of the Sri Lanka Health insurance providers struggle to analyze both historical trends and real-time claims efficiently.
- This project aims to solve these issues using Lambda Architecture to process **1+ million historical records** and **100+ real-time claims/sec** simultaneously.

2. Data Consumers

Consumer	Use Case
Insurance Companies	Detect fraudulent claims, optimize pricing
Healthcare Providers	Analyze payment patterns by region
Government Agencies	Monitor Medicare spending trends

3. Data Sources

- **Data Source:** [CMS Medicare Provider Data](#)
- The **Centers for Medicare & Medicaid Services (CMS)** is a U.S. federal agency under the Department of Health and Human Services (HHS) that administers Medicare, Medicaid, and the Children's Health Insurance Program (CHIP), while overseeing healthcare provider enrollment, publishing public datasets on healthcare utilization, payments, and provider demographics, and maintaining systems like the National Provider Identifier (NPI) registry.
- **Policy Analysts:** Studying Medicare/Medicaid spending and fraud.
- **Researchers:** Investigating chronic disease management or care models.
- **Providers:** Optimizing billing and compliance with CMS programs.

4. Analysis Methods

Layer	Techniques	Tools
Batch (Historical)	Aggregation Feature Engineering Feature Extractions	Azure Databricks
Speed (Real-Time)	Sliding window analytics	Azure Stream Analytics
Serving	Unified SQL views (Synapse Serverless Pool) Dashboard integration (historical + live data)	Power BI, Azure synapse SQL Pool

5. Architecture Diagram

