SAIM MULANI

+91 9518576712 • Pune, Maharashtra • mulanisaim5010@gmail.com
linkedin.com/in/saimmulani-data • github.com/saim5010

OBJECTIVE

Detail-oriented Data Analyst and BI Developer with hands-on experience in Power BI, DAX, SQL, Python, and machine learning. Adept at building end-to-end data pipelines — from SQL analysis and ML modelling to live Streamlit deployment and executive dashboards. Seeking entry-level Data Analyst, BI Developer, or Business Analyst roles where I can contribute to data-driven decision-making and grow within a collaborative analytics environment.

EXPERIENCE

Data Analyst Intern | NMD Infotech, Pune *2025 – Present*

- Developed the Asia Development Pulse Power BI dashboard — a 6-page interactive report analysing 43 Asian countries using World Bank WDI data spanning 1990–2022
- Architected a star-schema data model and authored 29 DAX measures tracking KPIs across GDP growth, HDI, trade balance, energy consumption, and demographic indicators
- Published the completed report to Power BI Service and maintained the full project repository on GitHub (saim5010/Asia-Development-Pulse-PowerBI)
- Built a Supply Chain Risk Analytics solution featuring dual XGBoost models with SHAP explainability on a SQLite star-schema data model, delivered as a 5-page Power BI dashboard and published to GitHub and Power BI Service

Career Break – Civil Services Preparation | UPSC / MPSC *2022 – 2025*

- Devoted to full-time preparation for UPSC and MPSC examinations, cultivating rigorous analytical reasoning, research aptitude, and structured problem-solving capabilities

EDUCATION

Professional Certification in Data Science & AI Engineering — IIT Guwahati & AlmaBetter

2025

Bachelor of Engineering — Pune University

2018 – 2022 • CGPA: 8.27

SKILLS

BI & Visualisation: Power BI, DAX, Power Query, Data Modelling, Tableau, Excel, Dashboard Design, Data Storytelling

Programming: Python (Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, SHAP), SQL, Streamlit, EDA, Statistical Analysis

Machine Learning: Classification, Regression, Clustering, XGBoost, Random Forest, Gradient Boosting, Feature Engineering, SHAP Explainability, Threshold Tuning

Data Engineering: Star Schema Modelling, ETL, SQLite, PostgreSQL, Data Preprocessing, Time Series Analysis

Cloud & Tools: Databricks Fundamentals (Accredited, Oct 2025), Microsoft Azure Data Fundamentals (DP-900) — Completed, Power BI Service, GitHub, Streamlit Cloud

CERTIFICATIONS

Databricks Fundamentals Accreditation — Databricks Academy

Oct 2025

Microsoft Azure Data Fundamentals (DP-900) — Microsoft

Completed

PROJECTS

Supply Chain Risk Analytics — *SQL · Python · XGBoost · SHAP · Power BI*

- Built an end-to-end supply chain risk analytics solution on a SQLite star-schema data model, engineering dual XGBoost classification models to flag at-risk shipments and suppliers

- Applied SHAP explainability to surface the key drivers behind each risk prediction, translating model outputs into actionable business insight
- Delivered a 5-page Power BI dashboard visualising risk exposure, and published the complete project (code, models, dashboard) to GitHub and Power BI Service with a recruiter-facing README

IBM Telecom Customer Churn Prediction — *SQL · Python · XGBoost · Power BI · Streamlit*
Live App: Streamlit App | **Dashboard:** Power BI Report | **GitHub:** saim5010/telecom-churn-predictor

- Engineered a 6-phase pipeline on 7,043 IBM Telco customers: raw data ingestion into SQLite, 9 SQL business queries (CTEs, window functions), Python EDA with 3 custom features, XGBoost classification (AUC-ROC 0.843), 4-page Power BI dashboard, and live Streamlit deployment
- Identified key churn drivers: Month-to-month contracts (42.7% churn vs 2.8% for two-year), Electronic check payment (45.3%), and a high-risk segment of new high-value customers churning at 80.6%
- Implemented SHAP explainability and threshold tuning (0.42), achieving 67% recall on churners; built a What-If retention ROI parameter in Power BI showing $41.7K/month recoverable at 30% retention
- Deployed real-time churn predictor on Streamlit Community Cloud with gauge chart, SHAP waterfall explanation, and per-customer ROI calculator

BI & Data Analysis

Asia Development Pulse – Power BI — *NMD Infotech*

- 6-page interactive Power BI dashboard analysing 43 Asian countries using World Bank WDI data (1990–2022) with star-schema data model and 29 DAX measures across GDP, HDI, trade balance, and energy KPIs

CRM Sales Dashboard – Excel

- Designed a comprehensive sales performance dashboard analysing 8,800+ CRM records tracking over $10M in closed revenue with dynamic pivot tables, KPI scorecards, and trend charts

DataViz Challenge – Airbnb Dashboard — *Tableau & Power BI*

- Built an interactive dual-platform dashboard addressing 30+ business scenario questions; modelled data into fact/dimension tables, created DAX measures for pricing trends and host KPIs, and identified price-inelastic high-demand micro-markets

LEGO Set Explorer – Power BI

- 3-page interactive dashboard exploring 15,000+ LEGO sets with KPI cards, dynamic detail panel, and Decomposition Tree for category-to-theme-to-set drill-down; implemented DAX measures and multi-level interactive filters

Machine Learning

Retail Sales Prediction – Rossmann Stores — *Regression*

- Ensemble model (Random Forest, Gradient Boosting, XGBoost) forecasting daily sales across 1,115 stores, achieving Test R^2 of 0.915 and MAPE of 14.23%; automated pipeline reducing manual estimation by 80%, saving 4,000+ hours annually

Netflix Content Clustering — *Unsupervised ML*

- Clustered 7,787+ Netflix titles using PCA for dimensionality reduction, implementing K-Means, Agglomerative Clustering, and DBSCAN; achieved a silhouette score of 0.621

Data Engineering & SQL

Amazon Relational Schema Design — *SQL & PostgreSQL*

- Designed relational schema for 10 key e-commerce entities; wrote SQL queries for product performance, pricing trends, customer behaviour, and seller metrics; optimised retrieval via PostgreSQL indexing strategies