

ARYAN MEENA

Boston, MA | (617) 861-1307 | araj7042@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

SUMMARY

Data analyst who turns messy multi-source data into reporting people decide from. Wrote the SQL behind 4.5M+ transactional records, automated a reporting cycle down 40%, and found the data-quality issues quietly distorting two headline numbers before anyone acted on them.

SKILLS

SQL and Databases: SQL (window functions, CTEs, subqueries, optimization), PostgreSQL, star schemas, ETL
BI and Excel: Power BI, DAX, Power BI Service, Tableau, Excel (pivot tables, XLOOKUP, Power Query)
Python and R: pandas, NumPy, scikit-learn, XGBoost, Matplotlib; R (tidyverse, caret); A/B testing
Statistics and Cloud: ANOVA, regression, forecasting, cohort and RFM analysis, PySpark, GCP, Docker, Git

EXPERIENCE

Boston University — Boston, MA

Sep 2025 – May 2026

Graduate Research Assistant, Quantitative Research

- Sourced and validated 3,932 trading days of data across 7 correlated financial series into reusable Python modules
- Tested a pre-specified 504-configuration grid and found the cost level, 1.56 basis points, where the margin disappears
- Reported that the best-performing result failed statistical correction, isolated the one component that held up, and wrote that up as the finding rather than the headline

Unified Mentor — Remote

Oct 2024 – Jan 2025

Data Analyst Intern

- Wrote and tuned the SQL behind 4.5M+ transactional records held across systems that did not agree
- Built self-service Power BI and Tableau dashboards over \$577K in revenue and 4.9K orders
- Automated the recurring reporting cycle for a 40% manual-effort cut, documenting every query and metric
- Found data-quality issues quietly distorting 2 headline numbers, and built the fraud classifier at 0.96 precision on 284K transactions behind the reporting

Space Applications Centre (SAC), ISRO — Ahmedabad, India

Jun 2022 – Jul 2022

Machine Learning Research Intern

- Trained U-Net and DeepLabV3+ road-segmentation models in PyTorch (ResNet-50, Dice loss, IoU) on 6,723 aerial images, with QGIS in the geospatial loop
- Found that benchmark-trained models did not transfer to Indian satellite imagery, and traced why

PROJECTS

RetailPulse: KPI Reporting and Dashboards | PostgreSQL · SQL · Power BI

[GitHub](#) · [Live Dashboard](#)

- Modeled 9 tables and 100K orders into a PostgreSQL star schema, writing cohort-retention and RFM SQL with window functions and NTILE quintiles
- Published a 4-page Power BI report with DAX measures; late deliveries cut review scores 4.29 to 2.57 stars

Hospital Readmission Risk and Cost Model | Python · scikit-learn · SHAP

[GitHub](#)

- Benchmarked 10 algorithms on 30K records at a 7:1 imbalance, tuning the threshold on cost and projecting \$1.77M annual savings per hospital

Learning Difficulty Screening: NHIS 2024 | R · caret · XGBoost

[GitHub](#)

- Benchmarked 36 models across 9 algorithm families and 4 balancing schemes on 353 survey variables, then resampled to separate stable variable importance from single-split noise

EDUCATION

Boston University, Metropolitan College — M.S., Applied Data Analytics | GPA: 3.75/4.0

Jan 2025 – May 2026

Indian Institute of Space Science and Technology — B.Tech., Electronics and Communication

Aug 2019 – May 2023

PUBLICATIONS

- Meena, A., et al. *WikiFlow: Predicting Wikipedia Editor Dropout at Scale*. EAI CSECS 2026, Springer LNICST
- *Accelerated CNN Training with Genetic Algorithm Optimization*. IEEE IATMSI 2024, IEEE Xplore; 1 under review
- *Enhanced Scene Text Recognition for Geocoding and Accessibility*. ICADIE 2024. **First author on all 3**