

Mohit Phulwani

San Jose, CA | (203) 727-8845 | mohitphulwani001@gmail.com | [LinkedIn](#) | [Website](#)

PROFESSIONAL SUMMARY

Data Engineer with a **Master's in Analytics (3.9 GPA)** and **15+ months** of experience in architecting and **orchestrating** production-grade data systems. Specialist in **Data Observability and Governance**, utilizing **SQL and Python** to conduct deep-dive **Root-Cause Analysis (RCA)** on **scalable distributed pipelines**. Expert in optimizing legacy database objects (**Stored Procedures, Views**) and **automating data validation** to eliminate manual auditing. Proven ability to bridge the gap between raw data infrastructure and executive-level BI reporting.

TECHNICAL SKILLS

- **Languages:** SQL (T-SQL, PostgreSQL, MySQL), Python, R, JavaScript (Node.js).
- **Data Engineering:** ETL/ELT Pipeline Development, Data Modeling (Schema Design), Data Quality (DQ) Testing, Query Optimization, Root Cause Analysis (RCA).
- **Tools & Cloud:** AWS (S3, Redshift), Supabase, Git/GitHub, VS Code, SSRS, Power BI, Excel (Advanced).
- **Libraries:** Pandas, NumPy, FastAPI, PyTorch, Scikit-learn, Streamlit.

WORK EXPERIENCE

Data Engineer (Graduate Assistant), Sacred Heart University, Connecticut

Oct 2024 – Dec 2025

- **Data Quality & Observability:** Owned the end-to-end troubleshooting lifecycle for institutional reporting; performed **Root-Cause Analysis (RCA)** on data discrepancies by auditing complex **T-SQL joins** and stored procedures, ensuring **100% data fidelity** for federal compliance.
- **Pipeline Optimization:** Refactored legacy SQL views and optimized indexing strategies within a **Microsoft SQL Server** environment, resulting in a **32% reduction in query execution time** for high-traffic dashboards.
- **Automated Validation:** Engineered **Python-based testing scripts (Pandas/NumPy)** to **automate Exploratory Data Analysis (EDA)** and validation checks, reducing manual data-reconciliation efforts by **14+ hours per week**.
- **Process Automation:** Spearheaded the migration of manual data entry workflows to **automated ETL pipelines**, handling **15,000+ student records** and eliminating human error in the "admissions-to-finance" data handoff.
- **System Integration:** Developed a modular suite of **50+ automated SSRS reports**, identifying and resolving bottlenecks between the **SQL database layer** and the **Visual Studio** front-end.
- **Technical Liaison:** Translated ambiguous business logic from Finance and Admissions stakeholders into **12 precise technical schemas**, reducing development rework by **18%**.
- **Financial Modeling:** Developed economic forecasting tools using Price Sensitivity modeling to simulate 10,000+ aid scenarios, identifying a **\$120K optimization** in institutional spending.

ENGINEERING PROJECTS

Orbit – Offline-First Data Sync Engine | React Native, FastAPI, PostgreSQL

- Architected a cross-platform data synchronization suite featuring a custom conflict-resolution engine and a local SQLite persistence layer to ensure **100% data availability** in offline environments.
- Implemented **JWT-based session management** and secure **REST API endpoints via FastAPI** to handle concurrent user synchronizations.

In-Cabin Seat Object Detection (SVIRO) | Python, YOLOv8, PyTorch

- Developed a high-performance **Computer Vision pipeline** for synthetic occupancy detection, achieving a **55.70% mAP@50** across 5 different model architectures.
- Engineered a **data augmentation pipeline** to resolve class imbalances, improving model accuracy in edge-case synthetic environments.

EDUCATION

- **Master of Science in Analytics (Data Science focus)**, Sacred Heart University, CT (GPA: 3.9/4.0) Aug 2024 – Dec 2025
- **Bachelor of Arts in Economics**, University of Mumbai Jan 2021 – Feb 2024
- **Diploma in Electronics and Communication Engineering**, VES Polytechnic, University of Mumbai

CERTIFICATIONS & ACHIEVEMENTS

- **IBM Data Science Professional Certificate:** Specialized in data visualization, machine learning, & structured query language.
- **Bloomberg Market Concepts (BMC):** Proficient in financial market data analysis and economic indicators.
- **Lead Researcher:** Engineered a high-performance AI-driven data pipeline using XGBoost and SHAP; automated the analysis of behavioral datasets to drive real-time intervention strategies.