

HARSHIT SIWAL

India · Open to Relocation & Remote · +91 73514 87226 · siwal.work@gmail.com

[linkedin.com/in/HarshitSiwal](https://www.linkedin.com/in/HarshitSiwal) · github.com/Zorosec

Data Analyst with 18+ months of internship experience in banking and analytics. Skilled in SQL optimization, Python for data cleaning, automation and predictive modeling, and dashboards in Power BI, Tableau and Excel. Google Data Analytics and Microsoft PL-300 certified.

WORK EXPERIENCE

InnoPy Solutions

09/2025 - Present

Data Analyst Intern

- Automated data entry, cleaning and validation in Python (Pandas), cutting manual processing time by roughly 40%.
- Optimized report SQL using indexing, CTEs and query refactoring, reducing average execution time by about 30%.
- Built an Excel data-quality dashboard with Power Query and pivot tables for automated checks and weekly reporting.

SBM Bank

04/2025 - 05/2025

Data Analyst Intern (Security)

- Validated 500,000+ financial transactions with Python (Pandas, NumPy) and SQL, improving records for compliance reporting.
- Tuned SQL queries for fraud reviews, identifying 10+ high-risk anomalies for the security team.
- Supported transaction-data ETL and analyzed spending patterns with descriptive statistics to inform fraud checks.

Career break: full-time preparation for Indian government competitive examinations (2023 - early 2026).

EDUCATION

Bachelor of Computer Applications

IMS Noida · India · Score: 72%

2019 - 2022

Relevant Coursework: Database Management Systems, Statistics, C & C++, and Data Structures.

PROJECTS

AI Data Analyst with Claude Code

Built an agentic analytics system with 25 subagents and 45 Markdown skills, connecting Claude Code to data warehouses to answer business questions in plain English. Reduced ad-hoc analysis time from hours to minutes.

SKILLS

Core: PySpark, Python (Pandas, Scikit-Learn, Statsmodels), R, SQL

Machine Learning: Feature engineering; precision, recall, AUC-ROC; logistic regression, random forest, XGBoost, gradient boosting; K-Means and hierarchical clustering

Experimentation & Statistics: A/B, hypothesis and multivariate testing; power analysis

Product & Risk Analytics: Behavioral risk signals, cohort and funnel analysis, conversion optimization, pattern recognition, retention, risk modeling, user behavior and segmentation

Cloud & Data Engineering: AWS (S3, Glue, Athena, SageMaker), Azure, Databricks, dbt, Snowflake

Visualization: Looker, Power BI, Tableau

Certifications: Google Data Analytics Professional Certificate (Coursera); Microsoft Power BI Data Analyst Associate (PL-300); Python, C & C++ (IIT Bombay)