

Dev Sanjay Shah

Jersey City, NJ | (551) 358-6637 | dshah11@binghamton.edu | [LinkedIn](#) | [GitHub](#)

SUMMARY

Data Analyst with 2+ years turning operational and business data into decisions across research, consulting, and supply-chain settings. Strong in SQL, Python, Tableau, Power BI, and Excel, building interactive dashboards, ETL and data-modeling pipelines, and automated reporting that surface KPIs and process trends. Skilled at data cleaning, validation, and statistical analysis, translating complex datasets into clear business insights and actionable recommendations for cross-functional stakeholders and senior leadership.

TECHNICAL SKILLS

Programming & Analysis: Python (Pandas, NumPy, Matplotlib, Seaborn, scikit-learn), SQL (Joins, CTEs, Window Functions, Aggregations), R, JavaScript

Databases & Warehousing: MySQL, PostgreSQL, MongoDB, SQLite, DuckDB, Dimensional Modeling (Star Schema, SCD Type 2), dbt

Analytics & Visualization: Tableau, Power BI, Excel (Pivot Tables, VLOOKUP/XLOOKUP, Power Query), Plotly, Matplotlib, Seaborn, Streamlit

Data Processes: ETL/ELT, Data Cleaning, Data Modeling, KPI Reporting, Dashboard Development, Data Validation & QA, Statistical Analysis (Hypothesis & A/B Testing), Process Automation

Tools & Workflow: Jupyter Notebook, Git, GitHub, JIRA, PySpark, FastAPI, CI/CD

Certification: Google Data Analytics Professional Certificate

PROFESSIONAL EXPERIENCE

The Research Foundation for SUNY, Data & System Analyst | Binghamton, NY

December 2024 – May 2026

- Enabled data-driven decisions across multiple projects by building interactive dashboards tracking operational KPIs, throughput, and resource utilization in Tableau and Power BI.
- Cut manual reporting effort and improved accuracy by automating extract-validate-reconcile workflows with systematic data-quality checks in SQL and Python.
- Strengthened downstream analysis by cleaning, validating, and QA-ing high-volume system data and root-causing discrepancies in SQL and Python.
- Aligned reporting to business needs by partnering with cross-functional stakeholders to define requirements and standardize KPI definitions across projects using Tableau and Excel.

Mastek, Digital Solutions Analyst | Ahmedabad, India

January 2024 – May 2024

- Cut project delivery timelines by nearly two weeks by analyzing operational workflows to pinpoint bottlenecks and delivering recommendations to project and engineering teams, in SQL and Excel.
- Accelerated data extraction and reporting by writing optimized queries with advanced joins, CTEs, and indexing across large relational databases in SQL.
- Gave senior management real-time visibility across 5+ engagements by developing KPI dashboards and performance reports in Power BI and Excel.
- Improved deliverable accuracy and consistency by designing and implementing data validation and quality-control checks within reporting pipelines, in SQL.

Krishna Associates, Operations & Data Analyst | Mumbai, India

May 2020 – May 2021

- Reduced excess inventory by 18% by analyzing purchasing and inventory data to close demand-supply gaps and optimize stock allocation across construction projects in Excel.
- Decreased stock-out incidents by 25% by building material-usage and cost-trend reports that informed procurement decisions.
- Improved on-time material availability to 95% by developing recurring procurement-performance reports that aligned scheduling priorities across teams in Excel.
- Enhanced demand-forecasting accuracy and tightened cost control by consolidating multi-source inventory and procurement data into unified dashboards in Excel.

PROJECTS

Cohort & Ledger — SaaS Retention & Revenue Analytics Warehouse | *SQL, dbt, DuckDB, Python, FastAPI* | [GitHub](#)

- Computed cohort retention, MRR, NRR, and LTV correctly in one reusable SQL layer over 3,000 customers and 24K+ payments by building a star-schema warehouse with SCD Type 2 dimensions in dbt and DuckDB.
- Ensured every consumer reads identical metrics by modeling subscription edge cases — upgrades, reactivations, grandfathered pricing, and failed-payment churn — and serving the analytics views via FastAPI.

Surgical Complications & Mortality Benchmarking — Hospital Quality Dashboard | *Python, SQL, DuckDB, Streamlit* | [GitHub](#)

- Benchmarked hospital complication and mortality rates against national rates with a state-level choropleth dashboard and multi-measure outlier detection on real CMS data in Streamlit and DuckDB.
- Prevented a 10x metric overstatement by catching a per-1,000-vs-per-100 scaling error, then ran 17,000+ binomial tests with FDR correction and gradient-boosting models in scikit-learn.

EDUCATION

SUNY Binghamton (Binghamton University), Master of Science in Information Systems

May 2026