

Saurabh Badkas

Hyattsville, MD | (571) 201-0045 | saurabh.b@nextitjobs.com | [LinkedIn](#) | [GitHub](#)

SUMMARY

Detail-oriented Data Analyst with 4+ years of experience in healthcare, finance, and tech industries, specializing in SQL, Python, and cloud-based analytics. Proficient in building automated ETL pipelines, regulatory reporting, and predictive dashboards. Adept at delivering compliance-ready insights that enhance operational efficiency, clinical outcomes, and strategic business decisions.

WORK EXPERIENCE

Johnson & Johnson

New Brunswick, NJ

Data Analyst

Feb 2025 – Current

- Extracted insights from over 5 million EHR and insurance claim records using SQL and Python, which shaped clinical trial feasibility analysis and supported an 18% reduction in protocol revisions.
- Constructed end-to-end Power BI dashboards to monitor global supply chain metrics inventory, fulfillment rates, and delays resulting in a 22% boost in timely product distribution across 20+ facilities.
- Standardized clinical trial datasets in Snowflake using dbt and SQL to ensure GxP and HIPAA compliance this enhancement accelerated FDA review timelines by 30% and improved data traceability across trials.

Citi

New York City, NY

Data Analyst

May 2024 – Nov 2024

- Devised Power BI dashboards and scripted SQL workflows to track over \$500M in loan transactions, trimming reporting cycles by 40% and accelerating regulatory deliverables across lending operations.
- Diagnosed fraud-prone patterns by parsing cardholder activity using Python (pandas, NumPy), surfacing \$1.3M in suspicious behavior across high-risk cohorts and supporting proactive mitigation.
- Audited and reconciled KYC data pipelines from Snowflake and Hadoop ecosystems, uplifting data quality benchmarks by 25% and enhancing readiness for quarterly AML examinations.
- Curated stress-testing scenarios for CCAR using Alteryx and advanced SQL joins on market risk and trade exposure datasets, boosting forecast precision and cutting compliance prep time by 30%.

Zentek Infosoft

Jaipur, IN

Data Analyst

Jun 2020 – Jul 2023

- Executed over 150+ automated ETL pipelines using AWS Glue and Airflow, streamlining data ingestion from EHR systems and external APIs, resulting in a 45% reduction in processing delays.
- Leveraged PySpark and SQL on Snowflake to perform longitudinal analysis of patient records, uncovering trends that supported a 18% drop in hospital readmissions for chronic care patients.
- Flagged \$1.2M in anomalies through advanced pattern detection on Amazon EMR, enhancing accuracy in insurance claim verification and accelerating audit turnaround by 60%.
- Composed dynamic Power BI dashboards displaying real-time hospital KPIs patient inflow, treatment duration, and discharge rates helping leadership optimize staffing and bed allocation.
- Maintained 98%+ data quality across clinical datasets by applying dbt transformations and Python-based validation scripts, supporting regulatory readiness under HIPAA standards.
- Collaborated with analytics and DevOps teams to enable real-time KPI tracking, cutting performance reporting latency to under 10 minutes, improving responsiveness in decision-making.

PROJECTS

Capstone Project – Total Wine Inventory Planning

- Collaborated with leadership to identify product gaps and developed a predictive model (91% accuracy) to uncover high-revenue opportunities, driving data-informed assortment and inventory decisions across retail locations.

ML-Driven Fulfillment Optimization and Dashboard

- Delivered a predictive logistics solution on AWS with 88% F1-score, cutting shipment delays by 25% through automated ETL, Airflow-driven model retraining, and a Tableau dashboard tracking OTD percentage, delay factors, and high-risk orders.

SKILLS

Programming Languages: SQL, Python, R

Databases: SQL Server, MySQL, PostgreSQL, MongoDB, NoSQL, Snowflake, Azure SQL, AWS Redshift

Machine Learning & Statistics: Regression Analysis, Classification Models, A/B Testing, Forecasting, Hypothesis Testing

Visualization Tools: SSRS, Tableau, Power BI, Matplotlib, Seaborn, Plotly

Frameworks & Libraries: Django, NumPy, Pandas, Scikit-learn, BeautifulSoup, SciPy

Project Management Tools: Jira, Microsoft Visio, MS Project, MS SharePoint

Cloud Technologies & Version Control: AWS (S3, Redshift, EC2, Lambda, RDS), Azure, GCP, Git, GitHub, GitLab

Networking & Load Balancing: Advanced Excel (Power Query, Power Pivot), Google Analytics, Jupyter Notebooks, Mixpanel

EDUCATION

University of Maryland

College Park, MD

Master of Science in Business Analytics

Aug 2023 – Dec 2024

NMIMS University

Mumbai, IN

Bachelor of Technology in Data Science

Jul 2017 – Aug 2021