

Abhishek Saraf

Data Analyst / Senior Data Analyst

Dallas, TX | +1 (682) 391 2934 | abhisaraf7457@gmail.com

SUMMARY

Results-driven **Data Analyst** with 4+ years of experience in **data visualization, statistical analysis, and business intelligence** using **SQL, Python, Tableau, Power BI, and Excel**. Skilled in **ETL pipelines, data modeling, data validation, and data governance** to ensure accuracy, compliance, and consistency across datasets. Adept at performing **forecasting, KPI reporting, and A/B testing** to support data-driven decisions and performance optimization. Experienced in building **automated dashboards and predictive analytics solutions** that deliver actionable business insights. Strong expertise in **cloud platforms (AWS, Azure)** and modern data ecosystems including **Snowflake, Databricks, and BigQuery**. Collaborative communicator focused on translating complex data into strategic, business-ready insights.

SKILLS

Programming & Scripting: SQL (Oracle, SQL Server, PostgreSQL), Python (Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn), R

Data Warehousing & Cloud: Snowflake, Azure Synapse, AWS Redshift, Google BigQuery, Microsoft Azure, AWS, GCP

ETL & Workflow Orchestration Tools: Azure Data Factory, AWS Glue, Alteryx Designer/Server, SSIS, dbt, Informatica, Apache Airflow

Data Modeling & Databases: Star Schema, Snowflake Schema, Kimball Methodology, OLAP/OLTP Systems, PostgreSQL, MS SQL Server, MySQL, Oracle, MongoDB, Teradata

BI & Visualization Tools: Power BI (DAX, Power Query), Tableau, Excel (Pivot Tables, Advanced Formulas), Looker, Qlik Sense

Big Data Ecosystem: Databricks (PySpark, Delta Lake), Hadoop, Kafka, Flink, Hive, Apache Spark

Machine Learning & Analytics: Predictive Modeling, Logistic Regression, Random Forest, Clustering (K-means), Forecasting Models

CRM & Analytics Tools: Salesforce, Microsoft Dynamics 365, Google Analytics, Alteryx

Data Governance & Compliance: HIPAA, PCI DSS, Data Quality, Metadata Management, Validation Frameworks

Version Control & CI/CD: Git, GitHub, Docker, Azure DevOps, Jenkins

Collaboration Platforms & Office Tools: SharePoint, MS Office Suite (Word, Excel, PowerPoint), Google Workspace (Docs, Sheets, Slides), MS Teams, Slack

WORK EXPERIENCE

U.S. Bank, USA | Data Analyst

Jan 2024 - Current

- Designed and automated **ETL pipelines** using **SQL, Python, and Azure Data Factory**, improving data ingestion efficiency by **45%** and ensuring seamless integration across multiple financial systems.
- Developed and maintained **interactive Power BI and Tableau dashboards**, providing real-time insights into **credit portfolio performance, delinquency rates, and fraud metrics**, increasing executive reporting accuracy by **30%**.
- Conducted in-depth **data validation, cleansing, and quality checks** using **Python (pandas, NumPy)** and **SQL procedures**, reducing reporting errors by **25%** and ensuring compliance with **data governance** standards.
- Collaborated with cross-functional teams to establish **data governance frameworks** covering data lineage, ownership, and metadata management, improving data transparency and trust across departments.
- Created **forecasting and trend analysis models** using **regression, clustering, and time-series techniques** to predict customer churn and transaction behaviors, improving marketing campaign targeting by **20%**.
- Optimized **data models and schema designs** in **Snowflake** and **Azure Synapse**, enhancing query performance and reducing data retrieval time by **40%** across analytical workloads.
- Built **automated KPI reports** to monitor **loan performance, customer segmentation, and credit utilization trends**, enabling leadership to make faster and data-backed strategic decisions.
- Applied **A/B testing and statistical analysis** to evaluate new credit policies and promotional offers, improving customer acquisition efficiency by **15%** while reducing risk exposure.
- Partnered with the **data engineering** and **DevOps** teams to implement **CI/CD workflows (Git, Jenkins, Azure DevOps)** for analytics pipelines, reducing manual deployment time by **50%**.
- Leveraged **Databricks (PySpark)** to process high-volume transactional data and generate near real-time insights for **fraud detection and anomaly identification**.
- Ensured compliance with **HIPAA, PCI DSS**, and internal audit requirements by enforcing **data validation, encryption, and governance** best practices across analytical environments.
- Collaborated with stakeholders across finance, compliance, and operations to translate analytical findings into **business recommendations**, driving measurable improvements in operational performance and revenue growth.
- Implemented **data-driven process automation** using **Power BI APIs and Python scripts**, reducing manual data refresh cycles by **60%** and improving report delivery speed across business units.

- Developed and automated **ETL pipelines** using **SQL, Python, and AWS Glue**, streamlining data ingestion from ERP, CRM, and manufacturing systems, which improved data availability by **40%**.
- Designed and deployed **Power BI and Tableau dashboards** for monitoring **sales performance, inventory management, and supply chain KPIs**, reducing manual reporting time by **35%**.
- Conducted comprehensive **data validation, cleansing, and reconciliation** across multiple data sources, improving data accuracy by **28%** and aligning with corporate **data governance** policies.
- Built robust **data models (Star & Snowflake Schema)** in **Snowflake** and **Azure Synapse**, optimizing analytical workflows and enhancing query performance by **30%**.
- Partnered with cross-functional business units to define **data standards, lineage, and metadata documentation**, ensuring consistent and compliant data usage across departments.
- Utilized **Python (pandas, NumPy, scikit-learn)** to perform **statistical analysis and forecasting**, identifying key demand patterns that improved production planning efficiency by **22%**.
- Integrated **predictive analytics and machine learning** models to forecast product demand and detect anomalies in quality metrics, reducing supply chain disruptions by **15%**.
- Enhanced **data governance and compliance** by implementing **data validation rules, access controls, and audit logs**, ensuring adherence to **HIPAA** and **ISO data quality standards**.
- Implemented **process automation** using **SQL stored procedures and Power Query**, cutting repetitive reporting tasks by **25%** and improving analyst productivity.
- Collaborated with **data engineering and QA teams** to maintain version control and ensure continuous improvement of analytical workflows through **Git** and **Azure DevOps**.
- Presented actionable insights to senior management using **visual storytelling** and **KPI dashboards**, directly influencing pricing and product distribution strategies.
- Authored detailed **technical documentation, SOPs, and data dictionaries**, improving data literacy and ensuring knowledge transfer across analytics and business teams.

EDUCATION

Master's in Data Science – May 2024

The University of Texas at Arlington, Arlington, TX