

Rishab Khatokar | Data Analyst

AZ, USA | +1 (520)-205-0384 | rishabkhatokar9@gmail.com | [LinkedIn](#)

SUMMARY

Data Analyst with 5 years of experience in healthcare, financial, and business analytics, specializing in **data engineering, predictive modeling, and business intelligence**. Adept at designing and optimizing **ETL pipelines in Azure and AWS**, building **machine learning models** for risk scoring and customer churn prediction, and creating **interactive dashboards using Tableau and Power BI**. Proven track record in delivering **actionable insights** that improve operational efficiency, reduce errors, and support strategic decision-making, with expertise in **HIPAA-compliant healthcare data, Medicare/Medicaid programs, and financial risk analysis**. Strong collaborator with cross-functional teams, skilled in **process optimization, stakeholder management, and data-driven strategy implementation**.

TECHNICAL SKILLS

Technical Skills: Python (Pandas, NumPy, Scikit-learn), SQL, PySpark, VBA/Macros, Azure, AWS, Snowflake, ETL pipeline development, data integration, Power BI, Tableau, Jupyter Notebook, Git, APIs, RESTful services

Machine Learning & Analytics: Predictive modeling, classification algorithms (Random Forest, Logistic Regression, XGBoost), feature engineering, model evaluation (Accuracy, Precision, Recall, ROC-AUC, F1 Score), anomaly detection, regression analysis, time series forecasting, clustering, NLP basics, A/B testing

Data Management & Reporting: Data cleaning, validation, transformation, workflow automation, KPI tracking, financial modeling, reporting efficiency, performance monitoring, dashboards, pivot tables, data reconciliation, data governance, metadata management

Healthcare Domain Expertise: Claims, clinical, and provider datasets, Medicare/Medicaid programs, HIPAA compliance, HEDIS, payment integrity, quality initiatives, population health analytics, risk adjustment, patient outcomes analysis, care management metrics

Financial Domain Expertise: Loan portfolio analysis, credit risk scoring, SME and vehicle finance, investment advisory, portfolio performance monitoring, risk evaluation, operational optimization, fraud detection, financial reporting, variance analysis, regulatory compliance

Business & Strategic Skills: Process improvement, operational efficiency, cross-functional collaboration, stakeholder management, actionable insights, decision support, project management, requirement gathering, business process analysis, strategic planning

WORK EXPERIENCE

Data Analyst | United Health Group | USA

Aug 2025 – Present

- Lead advanced analytics on **large-scale U.S. healthcare datasets**, including claims, clinical, and provider data, applying **predictive modeling and machine learning** to identify trends, improve patient outcomes, and ensure **HIPAA-compliant** data handling, which helped **reduce claim errors by 18%**.
- Design and maintain **interactive dashboards and reports** using **Tableau**, transforming complex **clinical and financial metrics** into actionable insights for non-technical stakeholders, enabling faster, data-informed decisions and **improving report accessibility by 25%**.
- Develop and optimize **Python-based data pipelines and SQL/PySpark queries** in **Azure**, ensuring high-quality, scalable **ETL processes** and seamless integration of sensitive healthcare datasets for accurate reporting and analysis.
- Collaborate with cross-functional teams, including data scientists, actuarial, and business units, to define analytical frameworks, implement **ML-driven models for risk prediction**, and support evidence-based strategies across Medicare and Medicaid programs, resulting in **10% improved predictive accuracy**.
- Support **payment integrity and quality initiatives**, including **Medicare Star Ratings** and claims accuracy, by performing thorough **data validation, root-cause analysis**, and applying **machine learning techniques** to detect anomalies and enhance compliance.
- Provide strategic guidance on **process improvements** and automation by identifying workflow inefficiencies, optimizing reporting pipelines, and introducing Python/SQL solutions that reduce manual effort and improve operational efficiency.

Data Analyst | Boost fincrop group | India

Aug 2022 - Jul 2024

- Developed and maintained **financial models** to forecast loan performance and evaluate investment opportunities, helping the team **reduce non-performing assets by 12%** while minimizing operational risks and supporting strategic decisions.
- Monitored **KPIs and portfolio performance**, providing real-time **insights to senior management**, enabling faster decision-making and **improving overall portfolio efficiency by 15%** across vehicle finance, SME loans, and investment products.
- Built and managed **ETL pipelines** using **Python and SQL** to integrate data from multiple financial platforms, ensuring **data accuracy, integrity, and regulatory compliance** across all reporting processes.
- Created **interactive dashboards and reports** using **Power BI**, translating complex **market trends, credit scores, and portfolio data** into actionable strategies that guided lending and investment decisions.
- Automated **data collection and reporting workflows** using **Python, SQL, and Excel VBA**, reducing manual effort by **30%** and significantly improving operational efficiency, data consistency, and reporting speed.

Data Analyst | Adons Softech | India

Sep 2020 - Jul 2022

- Helped clients improve **operations** and make better **decisions** by collecting and analyzing **information** from multiple sources, identifying **trends**, and providing actionable **recommendations** that resulted in **20% improvement in process efficiency**.
- Supported **business decision-making** by transforming large volumes of **data** into clear **insights**, explaining **patterns** and **performance metrics** to managers, which **led to a 15% increase in on-time project delivery** across teams.
- Collected, organized, and managed data** using **Excel and SQL**, ensuring **accuracy, consistency, and proper structure** to enable reliable **analysis, reporting, and business planning** for multiple projects.
- Analyzed business processes** and conducted **market research** to identify **inefficiencies** and uncover **opportunities**, delivering actionable **suggestions** that helped **reduce operational errors by 18%**.
- Established **dashboards and reports** using **Power BI**, transforming complex **information** into visual **insights** that allowed **stakeholders** to monitor **performance** and make informed, timely **decisions**.

PROJECTS

Customer Churn Prediction (Python, Scikit-learn, Tableau)

- Built a **machine learning model** using **Python** and **Scikit-learn** to predict customer churn by analyzing historical transactional and behavioral data, enabling proactive retention strategies that **reduced churn risk by 12%**.
- Engineered features, performed **data cleaning**, and applied classification algorithms including **Random Forest** and **Logistic Regression** to identify key churn indicators, improving **prediction accuracy by 15%** for actionable business insights.
- Evaluated model performance using metrics such as **accuracy, precision, recall, and ROC-AUC**, and fine-tuned hyperparameters to optimize predictions for data-driven decision-making across multiple customer segments.
- Created interactive **Tableau dashboards** to visualize churn probabilities, customer segments, and risk factors, enabling stakeholders to monitor trends and implement targeted **retention initiatives** effectively.

Financial Risk Scoring System (SQL, Power BI)

- Enhanced a **financial risk scoring system** to assess loan and investment risks by analyzing historical financial and transactional data, enabling more accurate **credit decision-making** for high-value portfolios.
- Engineered complex **SQL queries and data models** to calculate risk scores, segment customers, and identify high-risk accounts, improving the **accuracy and reliability** of risk assessments.
- Planned interactive **Power BI dashboards** to visualize risk distributions, portfolio exposure, and key performance metrics, allowing stakeholders to **monitor trends in real-time** and take proactive actions.

Healthcare Data Pipeline (AWS, Snowflake, Python)

- Designed and implemented an end-to-end **healthcare data pipeline** to ingest, process, and store large-scale clinical, claims, and provider datasets, ensuring **HIPAA compliance** and high-quality data for analytics and reporting.
- Built automated **ETL workflows using Python**, transforming raw healthcare data into structured formats and enabling seamless integration across **AWS and Snowflake** for real-time analytics.
- Developed **Power BI dashboards** to visualize key healthcare KPIs, claim trends, and patient outcomes, providing actionable insights that **reduced reporting errors by 20%** and boosted stakeholder decision-making.

EDUCATION

Master of Science (MS) in Information Science/Machine Learning | University of Arizona, USA

Bachelor of Technology (B.Tech) in Electronics and communication | Dayananda Sagar University, India