ANIKET SAKHARKAR

+1 857-230-5126 | aniketsakharkar4@gmail.com | LinkedIn | Portfolio | GitHub | Kaggle

SUMMARY

Data professional with experience across analytics, data engineering, and applied machine learning. Skilled in transforming raw data into actionable insights, building automated reporting pipelines, and developing predictive models using Python, SQL, and cloud technologies. Strong background in dashboard development, ETL workflows, and statistical analysis to support strategic, operational, and financial decision-making. Skilled at collaborating with cross-functional teams, enhancing data quality, and delivering scalable end-to-end data solutions

EDUCATION

Master of Science in Data Analytics Engineering (GPA 3.9), *Northeastern University, Boston, USA*   Jan 2022 - Dec 2023

TECHNICAL SKILLS

Programming & Analytics: Python (Pandas, NumPy, Scikit-learn), SQL (MySQL, Oracle, SQL Server), Data Wrangling, Statistical Modeling, A/B Testing, Feature Engineering
BI & Visualization: Power BI, Tableau, QuickSight, Advanced Excel, Data Storytelling, DAX
ETL & Data Engineering: ETL/ELT Pipelines, Airflow (concepts), Snowflake, Databricks, API Integration, Data Quality Checks
Machine Learning: Classification, Regression, Model Evaluation (Precision, Recall, F1), Predictive Analytics
Cloud & Databases: AWS (S3, Lambda, DynamoDB, API Gateway, ECR), Azure, GCP BigQuery
Business Skills: KPI Development, Requirements Gathering, Root Cause Analysis, Documentation, Cross-Functional Collaboration

PROFESSIONAL EXPERIENCE

Data Analyst, *KGS Technology, USA*   Feb 2025 - Aug 2025
- Designed **Power BI dashboards** to track service quality, regional performance, AOV, and delivery efficiency across 4 regions and 17 product categories
- Analyzed **500K+ operational transactions** using SQL, increasing **data accuracy by 30%** and reducing reporting discrepancies
- Identified performance gaps through trend analysis and delivered insights that improved decision-making **cycles by 12%**
- **Documented KPI** definitions, data flows, and logic enabling consistent reporting **across operations and finance teams**
- Partnered with cross-functional teams to highlight **root causes in margin and performance variation**, helping drive operational improvements

Data Analyst, *Rebecca Everlene Trust Company*, *USA*   Feb 2024 - Jan 2025
- Built **operational dashboards in Power BI** to monitor program engagement, completion rates, and satisfaction metrics for 1,083 students
- Consolidated CRM + attendance data (10K+ rows/month) using SQL, **improving data consistency by 35%** and reducing manual cleanup
- Conducted trend analysis to identify performance drop-offs, boosting **overall attendance and user experience by 8%**
- Automated reporting pipelines; increased refresh frequency from monthly to weekly and **reduced manual work by 60%**
- Created comprehensive documentation **for reporting standards, KPI logic**, and data quality processes

Data Analyst Intern, *State Farm Insurance, USA*   May 2023 - Aug 2023
- Enhanced health policy analysis **efficiency by 15%** through advanced data extraction and analysis of prior year's health data using DBeaver, delivering actionable insights to stakeholders and improving decision-making processes
- **Automated SQL query** execution and report generation, increasing **accuracy by 64%** and reducing manual intervention, while integrating **DBeaver with QuickSight** to streamline **SAS-to-Excel** processes for faster monthly health reporting
- Led a Health Analytics Proof of Concept (POC) that achieved an **85% seamless data migration** to AWS, while developing statistical algorithms for pricing function analysis, improving model accuracy, and **overall efficiency by 25%**

ACADEMIC PROJECTS

Supply Chain ML Inference Application |**Project Link**   Sep 2025 - Nov 2025
- Built a real-time ML inference system using **Snowflake, Databricks, Docker/ECR, AWS Lambda, and API Gateway** to predict shipment ETA delays

AWS Serverless ETL Pipeline |**Project Link**   Jun 2025 - Jul 2025
- Developed a fully serverless ETL pipeline using **S3 event triggers, AWS Lambda (Python), and DynamoDB** for automated CSV ingestion and transformation

Serverless ML Prediction API |Project Link   Mar 2025 - Apr 2025
- Deployed a real-time inference API using **AWS Lambda, API Gateway, and a scikit-learn Lambda Layer** for low-latency prediction serving

Databricks Retail ELT Pipeline |Project Link   Nov 2024 - Jan 2025
- Developed an end-to-end PySpark ELT pipeline on **Databricks** using **Delta Lake** (Bronze–Silver–Gold), generating retail **KPIs** and **dashboards** for sales and customer insights