Pradyumn K. Pottapatri

United States | 404-493-8533 | pottapatri@gmail.com | LinkedIn | GitHub | Leetcode

Data Engineer: 5 years in data and AI, sole owner of the data platform for a 250-person supply chain procurement org. Built its scheduled batch Snowflake + dbt pipelines and recovered $37M of contracted inventory with FP&A, Engineering.

Experience

Stanley Black & Decker Towson, MD
Sourcing Analyst/Data Engineer *05/2023 – 10/2026*

- Delivered same-day reporting (was 3 days) to executives and finance by owning a 3-layer batch ELT pipeline for 5M+ SAP transactions/day and contract manufacturer (CM), distributor feeds, using Snowflake, dbt, SQL, Python.
- Kept per-part inventory accurate for executives by catching null-row records (dbt tests on 7 models) and failed loads (alerts before each Power BI refresh), and documenting every pipeline, using dbt, Snowflake.
- Shortened Power BI refresh from 15–20 minutes (with frequent timeouts) to under 1 minute for 51 users on two VP-led teams by replacing nested views with incremental dbt models, using dbt, Snowflake, SQL.
- Reduced contracted inventory by $25.6M with FP&A, ending about $1M/yr in holding fees, by exposing $30–40M never drawn down (about $12M at one supplier, escalated to executives), using Power BI, Snowflake, SQL.
- Accelerated report retrieval 98% (to 30 min) for executives and grew chatbot users from 7 to 50+ by piloting a RAG chatbot on the 6 feasible of 11 interview-surfaced needs, then scaling it with the Gen AI team on Azure AI Foundry.
- Secured $11.8M of contracted inventory absorbed by 5+ CMs by segmenting 1,000 SKUs by usage to expose alternates bought outside the approved parts list, adopted by FP&A and Engineering, using Snowflake, SQL.
- Reached 85% tool-selection and 92% answer accuracy on a text-to-SQL and semantic-search routing agent piloted with a 7-person sourcing team, via LLM-as-judge evaluation over 110 questions, using Snowflake Cortex Agents.
- Completed pre-production review of 50+ system designs in 6 months on the IT Architecture Review Board, flagging a loading-table change that would have broken every downstream view and pipeline.

ADP Atlanta, GA
Data Science Intern *05/2022 – 08/2022*

- Cut Marketplace DevOps alerts 85% by replacing a static threshold firing every 15 minutes on 100K+ API calls with Kalman-filter and Exponential Smoothing thresholds, using Splunk MLTK, SPL, Python.

Georgia State University Atlanta, GA
Graduate Data Scientist & Researcher *08/2021 – 03/2023*

- Lifted F1 on a long-form contract classifier for a GSU research lab from 0.67 to 0.82, beating earlier teams, by redesigning the annotation quality framework and leading 3 people, using LSTM with GloVe, BERT.
- Attained F1 0.79 on a responsible-AI classifier that helps a GSU lab find stronger ethical agreements in corporate ethics documents, leading 6 NLP students, using PyTorch, TensorFlow.

Projects

ERP-to-App Pipeline, HVAC Distribution Case Study | *Python*, *SQL*, *DuckDB*, *Streamlit* 2026

- Proved dollar-exact reconciliation of 1.06M invoice lines across three marts by building an ingestion-to-mart pipeline from 10 raw ERP extracts, using Python, SQL, DuckDB.

Customer Churn Prediction | *XGBoost*, *Optuna*, *SHAP*, *MLflow*, *FastAPI*, *GitHub Actions*, *AWS ECS/Fargate*

- Served churn predictions in under 10 ms (92.5% recall) via a FastAPI service with GitHub Actions CI/CD on AWS ECS/Fargate, using Python, XGBoost, SQL.

Technical Skills

Languages: Python, SQL (Snowflake, MySQL, PostgreSQL), R
Data Platform: Snowflake, dbt (incremental, tests), Airbyte, ELT, DuckDB, SAP data, role-based access
Modeling & Quality: OLAP modeling, query tuning, freshness checks, idempotent reruns, failed-load alerting
Cloud & Engineering: AWS (S3, Lambda, ECS/Fargate), Git, GitHub Actions CI/CD, FastAPI, Docker, Spark
BI & AI: Power BI, Streamlit, RAG (Azure AI Foundry), LLM evaluation (RAGAS, LLM-as-judge)

Education and Certifications

Georgia State University Master of Science in Data Science and Analytics | GPA: 3.89
ABV-IIITM Gwalior Bachelor of Science in Computer Science | GPA: 3.70
Certifications: AWS Certified Cloud Practitioner, DataCamp AI Engineer, Machine Learning, Deep Learning.ai