

GNANESHWAR YADLA

Data Engineer • AI Automation Specialist — Python, SQL, Claude API, Snowflake, dbt, Databricks, Power BI, AWS, Azure
Richmond, VA • +1-571-751-0012 • gnaneshwarreddy056@gmail.com • [linkedin.com/in/gnaneshwar-reddy-](https://www.linkedin.com/in/gnaneshwar-reddy-) • [yadla.dev](https://github.com/yadla)

ABOUT ME

I'm a Data Engineer and AI Automation specialist with 3+ years building production-grade data pipelines and agentic AI systems. At Wipro, I shipped 15+ ETL/ELT pipelines processing 500GB+ daily across Snowflake, AWS, and Azure — and was promoted from intern to full-time engineer within 3 months. During my time at NeuralSeek, I built multi-LLM workflows integrating 4+ AI APIs to automate data processing at scale. I also independently built an AI-powered job application pipeline that runs 200+ daily applications across LinkedIn, Dice, and Monster using Claude API, Ollama, and Airflow — for under \$14/month. I work at the intersection of data engineering and AI automation, building systems that are cost-efficient, practical, and production-ready. Currently completing my M.S. in Computer Science at VCU (GPA 3.8/4.0).

TECHNICAL SKILLS

Python • SQL • PySpark • Databricks • Snowflake • dbt • AWS (S3, Glue, Redshift, Lambda) • Azure (Data Factory, Synapse) • Claude API • Ollama • LLM Orchestration • ETL/ELT • Data Modeling • Power BI • Tableau • Flask • Airflow • Docker • Kafka • Git

PROFESSIONAL EXPERIENCE

Data Engineer / Data Analyst | Wipro Limited

Aug 2021 – Jul 2024

Promoted from Intern to Full-Time within 3 months based on outstanding performance.

- Designed and built 15+ ETL/ELT pipelines in Snowflake and Azure ingesting data from 15+ source systems, transforming with dbt and PySpark, processing 500GB+ daily across AWS S3, Glue, Redshift, Lambda, and Azure ADF/Synapse with 99.8% uptime.
- Optimized complex SQL queries in Snowflake improving performance by 45%; built Power BI and Tableau dashboards tracking 24 KPIs and business performance for 200+ stakeholders across 3 business units.
- Implemented dbt data quality frameworks and Airflow-scheduled 12 processes, reducing data incidents by 50%, saving 320 hrs/month and \$120K annually in operational costs.

Agentic AI Intern — Data & Analytics | NeuralSeek

Jun 2025 – Aug 2025

- Built Python AI workflows processing 50K+ records/day integrating 4+ LLM APIs (Claude, GPT, Gemini, Watsonx), designed prompt caching strategies across Claude Haiku/Sonnet Batch API reducing token costs by 35% on high-volume inference pipelines.

Data Analyst Intern | Presidency University

May 2021 – Jul 2021

- Analyzed 40K+ records and built 3 Tableau dashboards tracking 12 KPIs, delivering 5 trend analyses and analytical reports for 200+ stakeholders across sales and operations.

KEY PROJECTS

AI Job Application Automator — Python • Claude API • Ollama • Playwright • Flask • Airflow

- Architected end-to-end AI automation pipeline targeting 200+ applications/day across LinkedIn, Dice, and Monster using Python scrapers with Claude Haiku/Sonnet Batch API + Prompt Cache, total AI cost of \$13.30/month.
- Designed multi-agent workflow: Ollama qwen3.5 (free, local) for JD analysis, Claude Haiku for tailored resume/cover letter generation, and Claude Sonnet with computer use for intelligent form filling across 3 job boards.
- Deployed Flask monitoring dashboard with 3x daily Airflow-scheduled runs tracking per-application status, pipeline health metrics, and real-time AI cost breakdowns for full operational visibility.

Airbnb Analytics ELT Pipeline — Snowflake • dbt • AWS S3 • Python • Airflow • Power BI

- Architected a full ELT pipeline ingesting raw Airbnb listing and review data from AWS S3 into Snowflake, processing through layered dbt models (staging, intermediate, mart) with incremental materialization that reduced full-refresh runtime by 60%.
- Built dbt data quality tests at every pipeline layer, designed a star schema with 12+ dimensions, and orchestrated nightly Airflow DAGs with 99.5% uptime delivering Power BI dashboards tracking pricing trends and occupancy KPIs for 200+ stakeholders.

Uber Trip Analytics Platform — Databricks • PySpark • Delta Lake • Kafka • AWS S3 • Tableau

- Built a real-time and batch analytics platform processing 100K+ daily Uber trip events using Kafka for streaming ingestion and PySpark on Databricks for micro-batch transformations, computing 40+ KPIs with sub-60-second data latency.
- Implemented Delta Lake for ACID-compliant storage with 6-month time-travel and schema enforcement, designed PySpark window aggregations for rolling demand forecasts, and delivered Tableau dashboards for 150+ business stakeholders.

Multi-Source Sales & Operations Dashboard — Snowflake • SQL • Python • dbt • Power BI • Excel

- Consolidated 10+ data sources (CRM, ERP, flat files) into a unified Snowflake analytics layer using Python ingestion scripts and 18+ dbt transformation models, building a star schema data mart for sales and operations reporting.
- Created 6 Power BI dashboards tracking 24 KPIs for quota attainment, pipeline conversion, and go-to-market analytics used by 200+ stakeholders, with automated data quality validation reducing reporting incidents by 50%.

EDUCATION

M.S. Computer Science, Virginia Commonwealth University | GPA: 3.8/4.0

Aug 2024 – Dec 2025

B.Sc. Computer Science, Presidency University, Bangalore, India

Aug 2019 – May 2022