Anupama Godara

+91 9717298623

[LinkedIn](#) | anupama.godara4@gmail.com

Data Engineer at S&P Global with a strong software engineering mindset. Skilled in Python, SQL, and Cloud technologies, with experience designing scalable systems, automating infrastructure, and optimizing data and ML pipelines through algorithmic thinking.

Work Experience (3+ yrs)

S&P Global – Associate Engineer
Jul 2023 - Present

Part of the Cloud Engineering and Shared Services team focused on building scalable cloud infrastructure and enabling data-driven innovation across business units.

- Designed and deployed cloud-native data pipelines on **AWS** (Glue, Lambda, S3, RDS) using **Python**, handling over **15M+ records per day** with fault tolerance, modular design, and point-in-time recovery.
- Automated infrastructure provisioning using Terraform, improving deployment reliability across multiple environments and reducing manual effort by 70%.
- Led **CI/CD** automation efforts using GitLab and **Terraform**, cutting deployment time by 40% and establishing review pipelines with team-wide adoption.
- Developed ML-powered cloud **cost forecasting** tools, applying algorithmic optimization techniques that reduced over-provisioning and saved **~$20K annually** in compute resources.
- Actively participated in code reviews, system design discussions, and cross-team initiatives to promote reusable infrastructure and data engineering standards.
- Built internal automation tooling for compliance tracking and reporting pipelines, enabling on-demand auditing and improving operational visibility for business stakeholders.
- Applied core data structures (e.g., hash maps, sets, stacks) and algorithmic techniques to optimize **ETL** job execution time by 30%, improving performance of time-critical data processing pipelines.

Captain Fresh – Business Analyst (I)
Sep 2022 – Jan 2023

A tech-first B2B platform leveraging data and automation to modernize supply chain operations. The company focuses on optimizing procurement and distribution through scalable, data-driven systems.

- Engineered Python scripts to automate complex data cleaning and pre-processing workflows, improving efficiency and reducing prep time by 30%.
- Developed and optimized **SQL** queries to aggregate and standardize large datasets from heterogeneous sources, enabling reliable insights for stakeholder reporting.
- Designed interactive Tableau dashboards to visualize KPIs and automate weekly reporting cycles, reducing ad-hoc data requests by 50%.
- Migrated the company's reporting ecosystem from Excel-based sheets to a centralized data visualization platform, improving data accuracy and enabling **real-time insights delivery**.

MiM - Essay – Business Analyst
Mar 2021 - Oct 2021

Consulting firm focused on analytics-driven user experience and data-informed business strategy.

- Designed SQL-based data pipelines to extract, analyze, and optimize web traffic and user engagement metrics — improving backend query performance and site responsiveness.
- Automated data workflows for bounce rate and conversion tracking using MySQL, contributing to data-informed UX decisions and reducing bounce rate by 24%.
- Applied structured query logic and basic system design principles to support SEO data infrastructure and traffic attribution models.
- Collaborated with marketing and engineering teams to develop and deploy scalable SEO strategies, increasing inbound traffic by 30%.

Education

Netaji Subhas Institute of Technology, Delhi
Aug 2015 -May 2019

CGPA – 8.1 (B.E. - Biotechnology)

Certification

- HashiCorp Certified: Terraform Associate (003)
- Introducing Generative AI with AWS – Udacity (Ongoing)

Skills

- **Programming Languages**: Python, Java, C++, SQL
- **Cloud & Infrastructure**: AWS (Lambda, Glue, S3, RDS, ECR, SageMaker AI), Azure
- **DevOps & Tooling**: Git, GitLab, Docker, Terraform, CI/CD Pipelines
- **Data Engineering & ETL**: Apache Airflow, PySpark, DBT, AWS Glue
- **Databases**: PostgreSQL, Snowflake, Amazon Redshift, DuckDB, NoSQL
- **Analytics & ML**: Scikit-learn, TensorFlow, Power BI (Advanced), Tableau
- **In Progress**: Exploring GCP (BigQuery, Cloud Functions, Dataflow) to expand multi-cloud proficiency