

# Sneha G G

## Lead Data Engineer

Exeter, UK • +44 - 7379473208 • sneha.g220691@gmail.com • <https://www.linkedin.com/in/sneha-gg>

*Lead Data Engineer with 7+ years architecting reliable, scalable data platforms (Snowflake, AWS) for healthcare systems processing >500M records monthly. Expert in building resilient ELT pipelines (dbt, Airflow, Python) that transform event-driven data into analytics-ready models for ML and BI. Proven track record delivering 99.9% availability while reducing costs by 40% in regulated environments. Passionate about leveraging data infrastructure to improve patient outcomes. UK authorized.*

### **AREAS OF EXPERTISE**

- Cloud - AWS, Azure
- Snowflake, Databricks
- Python, SQL, Pyspark
- ETL, ELT, dbt
- Acceptance Testing
- Reports, Dashboards
- Data Warehousing, Kimball
- Dimensional Data Modelling
- Data Lineage, Quality, Security
- Team Mentor & Collaborator
- CICD, Github Actions, Jenkins
- ML, Data Science
- Agile, Jira, Production Support
- Quick decision maker
- Solution-oriented, Focused

### **KEY ACHIEVEMENTS**

- **Leadership:** Mentored high-performing teams, fostering a culture of excellence and accountability
- **Operational Excellence:** Maintained 99.9% platform availability supporting high-volume data operations over 7+ years
- **Cost Optimization:** Reduced cloud operational costs by 40% through performance tuning and resource optimization
- **Data Quality Leadership:** Decreased downstream data issues by 50% by implementing best monitoring frameworks
- **Employee of the Quarter Award Recipient** - Recognised by Cigna Healthcare for innovative project delivery

### **PROFESSIONAL EXPERIENCE**

#### **GAPLABS., Inc** (PrimeVista LLC, US)

**UK (Remote)**

Client Engagement Manager

*Jan 2026–Present*

- Led stakeholder engagement and requirements translation for a U.S. construction SaaS platform, improving delivery speed and solution quality through tight collaboration with Product, UX/UI, and Engineering teams
- Managed stakeholder reviews, project milestones, and delivery timelines, accelerating project delivery by 2–3 weeks while maintaining alignment between business goals and technical constraints

#### **ACCENTURE** (Client: Cigna Health insurance, US)

**Chennai, India**

Associate Data Engineering Manager

*July 2017–Sept 2024*

- Owned the AWS data platform processing >500M records/month: architected ingestion, transformation and Snowflake-based dimensional models, improving report runtime by 35% and reducing data defects by 50%
- Designed and implemented ELT pipelines (dbt + Airflow + PySpark) for batch & near-real-time flows, cutting end-to-end latency by 40% for critical reports
- Optimised Snowflake compute and storage strategy generating a 40% reduction in cloud spend while maintaining 99.9% availability. Reduced end-to-end pipeline latency by 40% and data defects by 50% through best practices in monitoring, testing, and governance
- Led & mentored a team of 6 data engineers, instituted code review and testing standards, and partnered with Product & Data Governance to operationalize data security and compliance controls

*Tech stack: Snowflake, AWS (S3, Glue, Lambda, Redshift, Athena, Step Functions), dbt, Python, SQL, PySpark, Pytest, GitHub Actions*

#### **ANJANA SOFTWARE SOLUTIONS** (Client: Cigna Health insurance, US)

**Chennai, India**

Senior Software Engineer

*May 2013–July 2017*

- Delivered and supported production-grade data systems with 95%+ SLA compliance and rapid incident resolution
- Developed data integration services using web APIs, reducing processing time by 25% and improving data accuracy for downstream services
- Maintained a zero-defect delivery record across multiple releases through robust testing and version control practices

*Tech stack: Java, SQL, Web APIs, ETL, MVC, DB2*

## **EDUCATION**

---

**MSc in Data Science (1st Honors)**, University of Exeter, UK

*Sept 2024 - Sept 2025 (Completed)*

- **Dissertation:** Capri-CT Causal Analysis and Predictive Reasoning for Image Quality Optimisation in Computed Tomography: Implemented a causal deep learning framework employing VAEs and ensemble modelling to predict CT image quality (SNR) and optimise scan protocols, improving patient safety through reduced radiation exposure.
- **Modules:** Machine Learning, Data Analysis and Visualisation, Statistical Data Modelling, Social & text analysis

**MBA in Technology Management (1st Honors)**, Anna University, IN

*Feb 2015 - Aug 2017*

- **Modules:** Technology Strategy, Project Management, Business Analytics

**BE in Electronics & Communication (1st Honors)**, DMI College of Engineering, IN

*Aug 2008 - May 2012*

- **Modules:** Embedded Systems, Signal Processing, Networking

## **CERTIFICATIONS**

---

- **Udemy Certified** - The Project Manager Course, Beginner to PROject Manager, 2025
- **AWS Certified** - Data Analytics Specialty, 2024
- **Databricks Certified** - Data Engineer Associate, 2023

## **PUBLICATIONS & CONFERENCES**

---

*G.G., Sneha et al. "Capri-CT: Causal Analysis and Predictive Reasoning for Image Quality Optimisation in Computed Tomography"*

- Submitted to MICCAI 2025 Conference and BMC Medical Imaging Journal (under review).
- Preprint available: <https://arxiv.org/abs/2507.17420v1>