

Sneha Samita

Lead Data Engineer

Eskilstuna / Stockholm

+46764401327

snehasamita05@gmail.com

https://www.linkedin.com/in/sneha-samita/

Summary

A Data Engineering professional with over 14 years of experience specializing in Data Warehousing, ETL Development, and Database Development. Proven ability to design scalable **data architectures** and **data product frameworks** for secure and efficient data sharing in the **banking and financial sector**. Expertise in building reusable **data pipeline templates**, enforcing **data contracts**, and implementing **governance and supervision mechanisms** to ensure compliance and trust in data delivery.

- Designed and implemented reusable **data pipeline templates** and parameterized components to enforce **data contracts** aligned with data product principles, improving reusability and onboarding time by 30%.
- Contributed to the **Data Product Framework** for structured and secure data sharing across business units, enabling standardized access patterns and compliance checks.
- Provided architectural oversight and direction on **data integration and governance**, supporting supervised data access and lineage tracking across cloud-based data platforms.
- Led efforts to embed **metadata-driven frameworks** to automate schema validations and policy checks as part of the CI/CD pipeline, aligning with financial regulations.
- Collaborated with domain experts and business analysts to define **data product requirements**, ensuring that data services align with business needs and quality standards.

A **problem-solver with a strong analytical mindset**, I excel at meeting **tight deadlines**, adapting to **new technologies**, and driving **data-driven decision-making** in complex business environments.

Experience

Cognizant

Azure Data Engineer

April-2025

GSK – Data Modernization & Migration

- Designed and orchestrated **end-to-end data pipelines in Databricks**, creating multiple jobs and sequencing them through orchestration jobs to ensure smooth data flow.
- Developed processing logic to handle **Azure Data Explorer (ADX) time-series data**, enabling automated ingestion and preparation for downstream analytics and reporting.
- Implemented **data transformation workflows** for feature preparation, target variable creation, and reporting datasets, improving pipeline reliability and scalability.
- Collaborated with stakeholders to address integration challenges, including **Unity Catalog enablement**, **Azure ML registration automation**, and **enhanced transformation logic** for reporting and ADX datasets.
- Conducted **end-to-end testing** of pipelines to validate data quality, transformation logic, and reporting outcomes, ensuring consistency and trust in delivered datasets.
- Leveraged **Azure services (Databricks, ADX, ML, Data Lake)** to build scalable and reproducible data engineering solutions supporting the project.

Wells Fargo

Lead Data Engineer

November 2019 - October 2024

Country Risk and International Risk | Oct 2022 – Oct 2024

- Contributed to the development of a **metadata-driven data product framework** to ensure consistent data quality, lineage, and policy enforcement across cloud platforms.
- Designed **parameterized pipeline templates** in Azure Data Factory and Databricks, reducing redundancy and enabling faster delivery of compliant data products.
- Defined and enforced **data contracts** between producers and consumers to standardize schema validation and error handling.
- Provided **architectural guidance** for secure data sharing workflows aligned with organizational risk and compliance needs.

Centralized Analytics & Risk Aggregation Tool | Dec 2020 – Oct 2022

- Led MSBI ETL initiatives, designing 50+ Stored Procedures, SSIS packages, and SSRS reports.
- Optimized ETL pipelines using MSBI and Ab Initio, reducing data processing time by 40%.
- Developed interactive SSRS reports, enabling real-time drill-through analysis from Excel.
- Published & managed 50+ Tableau reports, ensuring 99.9% uptime and real-time data updates.
- Leveraged Ab Initio for scalable data integration, batch processing, and high-performance ETL jobs across large datasets

Escalation Audit Reporting | Dec 2020 – Sept 2021

- Collaborated with governance teams to embed **access control, RBAC**, and data usage policies into the audit dashboards.
- Implemented **data product standards** for visualization objects, ensuring audit consistency across reporting platforms.
- Developed **data sharing workflows** with strict contract enforcement to validate audit metrics from upstream systems.

Workforce Management Reporting | Jan 2021 – Dec 2021

- Created centralized **data models** and dashboard templates, enabling self-service analytics through governed data products.
- Enforced **data usage policies** and lineage tracking across Tableau to comply with internal audit and HR data protection rules.

Single-Counterparty Credit Limits (SCCL) | Nov 2019 – 2024

- Developed & optimized 50+ SSIS packages, 100+ Stored Procedures**, accelerating **data processing by 25%**.
- Implemented and automated SSRS reports**, improving executive reporting efficiency by 30%.
- Integrated Tableau dashboards** for SCCL compliance monitoring, reducing risk exposure by 15%.
- Strengthened data governance**, implementing **RBAC and security monitoring** for SCCL compliance.

Concentration Risk Reporting | Nov 2019 – Mar 2021

- Led MSBI ETL initiatives**, designing **50+ Stored Procedures, SSIS packages, and SSRS reports**.
- Optimized ETL pipelines**, reducing data processing time by 40%.
- Developed interactive SSRS reports**, enabling **real-time drill-through analysis** from Excel.
- Published & managed 50+ Tableau reports**, ensuring **99.9% uptime and real-time data updates**.

Tata Consultancy Services

Data Engineer

December 2010 - November 2019

Manpower Group RPO Business Unit | Manpower

May 2017 – Nov 2019

- Led end-to-end Data Warehouse development**, from database modeling to deployment, improving data processing time by 40%.
- Designed and implemented Fact and Dimension Tables** using **Star and Snowflake schemas**, optimizing reporting performance.
- Developed 50+ SSIS packages** for ETL automation, reducing data latency by 35%.
- Created 40+ optimized Stored Procedures, Views, and Functions**, enhancing SQL query performance by 30%.
- Managed Production deployments**, ensuring **99% ETL job success rate** and minimal downtime.
- Developed interactive Tableau dashboards** with Quick Filters, Parameters, and Calculated Fields, improving data insights.
- Optimized Tableau extract refresh times**, enhancing dashboard performance by 25%.
- Collaborated with stakeholders** to define reporting needs, document ETL workflows, and provide post-deployment support.

Customer Relationship Management (CRM) Security Model | Royal Bank of Scotland

Aug 2015 – Sept 2016

- ETL Enhancements:** Improved CRM data load efficiency by 25%, maintaining a **95% success rate** for ETL jobs.
- Performance Optimization:** Fine-tuned SQL scripts, reducing processing time by 30%.

Enterprise Referral Reporting | Bank of America

Mar 2014 – July 2015

- Automation & Optimization:** Automated revenue processing, reducing manual efforts by 80% and ensuring **100% success rate** for daily data loads.
- ETL Development:** Integrated new tables into **Data Mart**, improving reporting efficiency.

Capital Market Access Control System (CMACS) | Royal Bank of Canada

June 2012 – Feb 2014

- Led end-to-end ETL development**, automating data pipelines and reducing manual intervention by 60%.
- Resolved 100+ data and system issues**, ensuring **99.5% system uptime** and improved reliability.
- Optimized SQL queries and indexing strategies**, reducing report execution time by 25%.
- Automated daily ETL workflows**, minimizing manual monitoring efforts and achieving a **40% increase in operational efficiency**.
- Enhanced data validation processes**, improving data accuracy and reducing discrepancies by 30%.

Benefits Administration Services (BASE) | Bank of America

Dec 2010 – June 2012

- Data Warehouse Enhancements:** Improved Data Warehouse structures, boosting efficiency by 30%.
- Stored Procedure Development:** Created 20+ **optimized SQL objects**, enhancing database performance.

Skills	Technical Skills	
	<ul style="list-style-type: none"> • Cloud & Big Data: Azure Data Factory, Databricks, Synapse Analytics, Azure Data Lake Storage, Kafka, Apache Airflow, Docker • Databases & ETL: SQL Server, T-SQL, SSIS, Data Warehousing, ETL Development, Oracle, PL/SQL • Programming & Scripting: Python, PySpark, Unix • BI & Reporting: Power BI, Tableau, SSRS • DevOps & Deployment: CI/CD, Azure DevOps 	
	Soft Skills	
	<ul style="list-style-type: none"> • Data Governance & Compliance • Performance Optimization & Troubleshooting • Cross-functional Collaboration, Process Automation • Problem-solving & Analytical Thinking • Stakeholder Communication 	
Education	National Institute of Technology, Jamshedpur 9.14	2007-2010 Computer Science
	Birla Institute of Technology, Mesra, Ranchi 8.81	2004-2007 Computer Science
Certifications	Data Engineering on Microsoft Azure (DP-203) Microsoft Certified	
	Azure Fundamentals (AZ-900) Microsoft Certified	
	Data Engineer Associate Databricks Certified	
	Certified Tableau Desktop Specialist (Tableau 2020.2) Tableau	
	Professional in Querying Microsoft SQL Server 2012 (70-461) Microsoft Certified	
	Professional in Implementing a Data Warehouse with Microsoft SQL Server 2012 (70-463 MCP Exam) Microsoft Certified	
	1Z0-051 Oracle Database SQL Fundamentals 1 Oracle Certified Professional	
Languages	Languages	
	<ul style="list-style-type: none"> • English – Full Professional Proficiency • Hindi – Native / Bilingual Proficiency • Swedish – C1 Level (Advanced Proficiency) 	