





|   |  |  |
|---|--|--|
| <div><div>Sneha Samita</div><div>Lead Data Engineer</div><div><div><div><div></div><div>Eskilstuna / Stockholm</div></div><div><div></div><div>+46764401327</div></div><div><div></div><div>snehasamita05@gmail.com</div></div><div><div></div><div>https://www.linkedin.com/in/sneha-samita/</div></div></div></div></div> |  |  |
| Summary   | <p>A Data Engineering professional with over <b>14 years of experience</b>, specializing in <b>Data Warehousing, ETL Development, Cloud Data Platforms, and Database Development</b>. Proven ability to design <b>scalable data architectures, Lakehouse solutions, and data product frameworks</b> for secure and efficient data sharing. Expertise in <b>Databricks, Unity Catalog, Delta Lake, and Azure</b>, with strong focus on <b>data governance, compliance, and production-grade pipelines</b>.</p> <p><b>Key Achievements &amp; Contributions:</b></p> <ul style="list-style-type: none"><li>Designed and implemented <b>reusable, parameterized ETL pipelines</b> and data workflows in <b>Databricks</b>, enforcing <b>data contracts</b> and <b>governance principles</b>, improving reusability and onboarding time by <b>30%</b>.</li><li>Contributed to <b>Data Product Frameworks</b> for secure and standardized data sharing across business units, enabling compliance checks and structured access patterns.</li><li>Provided <b>architectural oversight</b> for <b>cloud-based data platforms</b>, ensuring <b>supervised access, lineage tracking, and metadata-driven governance</b>.</li><li>Embedded <b>metadata-driven automation</b> for <b>schema validation, policy enforcement, and CI/CD integration</b>, aligning with regulatory requirements.</li><li>Collaborated with <b>business analysts and domain experts</b> to define data product requirements, ensuring data services meet <b>quality standards, compliance, and business needs</b>.</li><li>Mentored teams and promoted <b>best practices in modern data engineering</b>, including <b>Lakehouse architecture, Delta Live Tables, and Unity Catalog governance</b>.</li></ul> <p>A <b>problem-solver</b> with a strong analytical mindset, adept at <b>adapting to emerging technologies</b>, meeting tight deadlines, and enabling <b>data-driven decision-making</b> in complex business environments.</p>  |  |
| Experience  | <div><div>Cognizant</div><div>Azure Data Engineer</div><div>GSK – Data Modernization &amp; Migration</div><div><ul style="list-style-type: none"><li>Designed and orchestrated <b>end-to-end data pipelines in Databricks</b>, creating multiple jobs and sequencing them through orchestration jobs to ensure smooth data flow.</li><li>Developed processing logic to handle <b>Azure Data Explorer (ADX) time-series data</b>, enabling automated ingestion and preparation for downstream analytics and reporting.</li><li>Implemented <b>data transformation workflows</b> for feature preparation, target variable creation, and reporting datasets, improving pipeline reliability and scalability.</li><li>Collaborated with stakeholders to address integration challenges, including <b>Unity Catalog enablement, Azure ML registration automation, and enhanced transformation logic</b> for reporting and ADX datasets.</li><li>Conducted <b>end-to-end testing</b> of pipelines to validate data quality, transformation logic, and reporting outcomes, ensuring consistency and trust in delivered datasets.</li><li>Leveraged <b>Azure services (Databricks, ADX, ML, Data Lake)</b> to build scalable and reproducible data engineering solutions supporting the project.</li></ul></div><div>Wells Fargo</div><div>Lead Data Engineer</div><div>Country Risk and International Risk   Oct 2022 – Oct 2024</div><div><ul style="list-style-type: none"><li>Contributed to the development of a <b>metadata-driven data product framework</b> to ensure consistent data quality, lineage, and policy enforcement across cloud platforms.</li><li>Designed <b>parameterized pipeline templates</b> in Azure Data Factory and Databricks, reducing redundancy and enabling faster delivery of compliant data products.</li><li>Defined and enforced <b>data contracts</b> between producers and consumers to standardize schema validation and error handling.</li><li>Provided <b>architectural guidance</b> for secure data sharing workflows aligned with organizational risk and compliance needs.</li></ul></div><div>Centralized Analytics &amp; Risk Aggregation Tool   Dec 2020 – Oct 2022</div><div><ul style="list-style-type: none"><li>Led MSBI ETL initiatives, designing 50+ Stored Procedures, SSIS packages, and SSRS reports.</li><li>Optimized ETL pipelines using MSBI and Ab Initio, reducing data processing time by 40%.</li><li>Developed interactive SSRS reports, enabling real-time drill-through analysis from Excel.</li><li>Published &amp; managed 50+ Tableau reports, ensuring 99.9% uptime and real-time data updates.</li><li>Leveraged Ab Initio for scalable data integration, batch processing, and high-performance ETL jobs across large datasets</li></ul></div><div>Escalation Audit Reporting   Dec 2020 – Sept 2021</div><div><ul style="list-style-type: none"><li>Collaborated with governance teams to embed <b>access control, RBAC</b>, and data usage policies into the audit dashboards.</li><li>Implemented <b>data product standards</b> for visualization objects, ensuring audit consistency across reporting platforms.</li><li>Developed <b>data sharing workflows</b> with strict contract enforcement to validate audit metrics from upstream systems.</li></ul></div><div>Workforce Management Reporting   Jan 2021 – Dec 2021</div><div><ul style="list-style-type: none"><li>Created centralized <b>data models</b> and dashboard templates, enabling self-service analytics through governed data products.</li><li>Enforced <b>data usage policies</b> and lineage tracking across Tableau to comply with internal audit and HR data protection rules.</li></ul></div><div>Single-Counterparty Credit Limits (SCCL)   Nov 2019 – 2024</div><div><ul style="list-style-type: none"><li><b>Developed &amp; optimized 50+ SSIS packages, 100+ Stored Procedures</b>, accelerating <b>data processing by 25%</b>.</li><li><b>Implemented and automated SSRS reports</b>, improving executive reporting efficiency by <b>30%</b>.</li><li><b>Integrated Tableau dashboards</b> for SCCL compliance monitoring, reducing risk exposure by <b>15%</b>.</li><li><b>Strengthened data governance</b>, implementing <b>RBAC and security monitoring</b> for <b>SCCL compliance</b>.</li></ul></div><div>Concentration Risk Reporting   Nov 2019 – Mar 2021</div><div><ul style="list-style-type: none"><li><b>Led MSBI ETL initiatives</b>, designing <b>50+ Stored Procedures, SSIS packages, and SSRS reports</b>.</li><li><b>Optimized ETL pipelines</b>, reducing data processing time by <b>40%</b>.</li><li><b>Developed interactive SSRS reports</b>, enabling <b>real-time drill-through analysis</b> from Excel.</li><li><b>Published &amp; managed 50+ Tableau reports</b>, ensuring <b>99.9% uptime and real-time data updates</b>.</li></ul></div><div>Tata Consultancy Services</div><div>Data Engineer</div><div>December 2010 - November 2019</div><div>Manpower Group RPO Business Unit   Manpower</div><div>May 2017 – Nov 2019</div><div><ul style="list-style-type: none"><li><b>Led end-to-end Data Warehouse development</b>, from database modeling to deployment, improving data processing time by <b>40%</b>.</li><li><b>Designed and implemented Fact and Dimension Tables</b> using <b>Star and Snowflake schemas</b>, optimizing reporting performance.</li><li><b>Developed 50+ SSIS packages</b> for ETL automation, reducing data latency by <b>35%</b>.</li><li><b>Created 40+ optimized Stored Procedures, Views, and Functions</b>, enhancing SQL query performance by <b>30%</b>.</li><li><b>Managed Production deployments</b>, ensuring <b>99% ETL job success rate</b> and minimal downtime.</li><li><b>Developed interactive Tableau dashboards</b> with Quick Filters, Parameters, and Calculated Fields, improving data insights.</li><li><b>Optimized Tableau extract refresh times</b>, enhancing dashboard performance by <b>25%</b>.</li><li><b>Collaborated with stakeholders</b> to define reporting needs, document ETL workflows, and provide post-deployment support.</li></ul></div><div>Customer Relationship Management (CRM) Security Model   Royal Bank of Scotland</div><div>Aug 2015 – Sept 2016</div><div><ul style="list-style-type: none"><li><b>ETL Enhancements:</b> Improved CRM data load efficiency by <b>25%</b>, maintaining a <b>95% success rate</b> for ETL jobs.</li><li><b>Performance Optimization:</b> Fine-tuned SQL scripts, reducing processing time by <b>30%</b>.</li></ul></div><div>Enterprise Referral Reporting   Bank of America</div><div>Mar 2014 – July 2015</div><div><ul style="list-style-type: none"><li><b>Automation &amp; Optimization:</b> Automated revenue processing, reducing manual efforts by <b>80%</b> and ensuring <b>100% success rate</b> for daily data loads.</li><li><b>ETL Development:</b> Integrated new tables into <b>Data Mart</b>, improving reporting efficiency.</li></ul></div><div>Capital Market Access Control System (CMACS)   Royal Bank of Canada</div><div>June 2012 – Feb 2014</div><div><ul style="list-style-type: none"><li><b>Led end-to-end ETL development</b>, automating data pipelines and reducing manual intervention by <b>60%</b>.</li><li><b>Resolved 100+ data and system issues</b>, ensuring <b>99.5% system uptime</b> and improved reliability.</li><li><b>Optimized SQL queries and indexing strategies</b>, reducing report execution time by <b>25%</b>.</li><li><b>Automated daily ETL workflows</b>, minimizing manual monitoring efforts and achieving a <b>40% increase in operational efficiency</b>.</li><li><b>Enhanced data validation processes</b>, improving data accuracy and reducing discrepancies by <b>30%</b>.</li></ul></div><div>Benefits Administration Services (BASE)   Bank of America</div><div>Dec 2010 – June 2012</div><div><ul style="list-style-type: none"><li><b>Data Warehouse Enhancements:</b> Improved Data Warehouse structures, boosting efficiency by <b>30%</b>.</li><li><b>Stored Procedure Development:</b> Created <b>20+ optimized SQL objects</b>, enhancing database performance.</li></ul></div></div> |  |

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