

Mariyam George

g.mariyam@gmail.com • [+16046006144](tel:+16046006144) • Vancouver, Canada • [linkedin.com/in/mariyam-rajeev-george-8944b870](https://www.linkedin.com/in/mariyam-rajeev-george-8944b870)

Professional Summary

Data Engineer with 9+ years of experience building scalable, low-latency data systems at companies like Meta and Cargill. Proven track record of reducing compute costs, accelerating developer velocity, and enabling experimentation through metric instrumentation. Skilled in Spark, Databricks, Presto, Airflow, Hive, and Hadoop.

Key Skills

- **Technologies:** Databricks, Delta Lake, Apache Spark, Python, Apache Airflow, SQL, Power BI, Presto, Hive, Hadoop
- **Data Engineering:** ETL/ELT processes, multidimensional data modeling, query optimization, data pipeline development, data migration, CI/CD implementation
- **Leadership:** Team leadership, cross-functional collaboration
- **Analytics:** Business intelligence dashboards

Professional Experience

Senior Data Engineer

Avanade @ Meta

Dec 2022 – Present

Vancouver, Canada

- Architected a multidimensional data aggregation framework powering 160+ widgets for Metaverse analytics, reducing whale query latency from 45s to <3s and saving ~\$60K/month in compute costs while processing 10TB+ of data daily.
- Designed and implemented a data model and pipeline for Meta's internal Metaverse dogfooding dashboard, enabling leadership to track employee usage against corporate VR adoption goals. Provided real-time visibility into individual and team-level compliance, influencing product adoption strategies across departments.
- Served as technical lead for a team of 5 data engineers, driving adoption of performance-driven ETL practices, improving data reliability SLAs, and aligning pipeline architecture with product experimentation needs.

Data Engineer

Avanade @ Meta

Jan 2021 – Nov 2023

Vancouver, Canada

- Led a cross-functional effort to productionize code review health signals, designing metric pipelines that enabled experimentation and adoption of the Overdue NudgeBot system. Enabled a 7% decrease in review time and a 10% reduction in churn through metrics instrumentation, A/B test enablement, and operational observability.
- Built a full-stack analytics suite for Meta's Codemod Services (Phabricator), tracking key KPIs like job runtime, reviewer cost, and stale diff abandonment. Used Apache Spark and Presto to reduce Sandcastle runtime by 15%, cut abandoned review effort by 30%, and define "healthy config" standards across hundreds of internal services.
- Built the metrics that powered and constructed dashboards with internal Meta tools to monitor codemod metrics (e.g., review cycle time, unique eyeball time, Sandcastle job counts), enabling data-driven decisions that reduced stale diffs and improved developer productivity by 20%.

- Engineered and delivered PSL performance metrics by analyzing 750K+ diffs and 200M+ data points, reducing code review times by 7.1%, cycle times by 6.8%, and long-running diffs by 11.8%, enabling scalable improvements in developer velocity and workflow efficiency.

Data Engineer

phData @ Cargill

Apr 2018 – Dec 2020

Minneapolis, MN, United States

- Built data migration pipelines using **Python** for a major SAP product, streamlining Cargill's financial data processing by 20% for actionable dashboard insights.
- Developed scalable ETL workflows with **Spark** for SAP financial data, enhancing data quality and enabling real-time supply chain analytics for Cargill.
- Optimized **SQL** queries for SAP-driven dashboards, reducing compute costs by 15% and improving financial reporting efficiency.

Educational Qualifications

- M.S., Data Science, University of St. Thomas, Saint Paul ,MN, USA • GPA: 3.8 • 2016–2018
- B.S., Information Technology, Anna University, India • 86% • 2006–2010

Certifications

- Databricks Certified Data Engineer Professional, Databricks, 2022
- Microsoft Azure Data Engineer Certified, Microsoft, 2024

Awards

- Growth Guru Award, Avanade, for being the first contingent worker from Avanade Canada in Meta and the primary driving factor in growing the team to a 31-member team
- AI Ambassador Award, Avanade, for spearheading and expanding a biweekly AI upskilling program impacting 30+ engineers. Developed comprehensive learning tracks from beginner to expert levels, organized a Gen AI hackathon, and facilitated GenAI use case demonstrations, resulting in 30+ engineers mastering Gen AI tools and 10+ engineers advancing to Azure AI 102 certification
- phDer Award, phData, for the best performing team for continually providing support activities for client