

Mizuki Nakano

Vancouver, BC | 236-978-2639 | mizukif1187@gmail.com | [LinkedIn](#) | [Github](#)

ABOUT

Data engineer with 3+ years of experience designing and optimizing data platforms across startups and global enterprises. Skilled in building scalable data pipelines, modeling analytics-ready datasets, and enforcing data governance standards to drive impactful business decisions. Proficient in Python, SQL, Airflow, PySpark, and cloud/data platforms including GCP, AWS, Databricks, and Snowflake.

EXPERIENCE

Freelance Data Analyst

January 2025 – April 2025

Booklistia

Tokyo, Japan (Hybrid)

- Built a statistical model to forecast revenue impact from discontinuing the customer loyalty program, supporting strategic planning and risk assessment.
- Conducted what-if analysis to model cost and profit scenarios across marketing budget reallocations, evaluating ROI trade-offs between initiatives.
- Analyzed customer purchase behavior across multiple segments to identify patterns associated with high customer lifetime value, supporting data-driven marketing strategies and customer retention initiatives.
- Developed interactive dashboards in Metabase by integrating sales, product, and customer data to support the marketing team's promotion strategy.

Key Tech Stack: Python, Snowpark, SQL, Snowflake, Google Sheet, Metabase, GitHub

Data Architect

October 2023 – January 2025

Accenture

Tokyo, Japan (Hybrid)

- Built scalable data marts using PySpark running on AWS EKS, processing large volumes of marketing data to drive data-informed decisions.
- Architected data marts and ensured on-time delivery by translating business requirements into technical specs, assigning tasks to engineers, and overseeing execution.
- Monitored 20+ DAGs and 50+ data sources, ensuring pipeline reliability and fast incident recovery.
- Created a centralized data catalog to improve data governance, lineage tracking, and accelerate incident resolution.
- Standardized development workflows by introducing CI/CD workflows, modular code structure, and code review guidelines across the team.
- Represented the data engineering team in a cross-team Databricks PoC, testing performance gains and cost efficiency for ETL workloads in comparison to existing PySpark on EKS pipelines.

Key Tech Stack: Python, pySpark, SQL, AWS, EKS, Kubernetes, Grafana, Airflow, GitLab

Junior Data Engineer

June 2022 – September 2023

Panalyt

Tokyo, Japan (Remote)

- Developed data pipelines using Python and SQL for multi-source data ingestion, transformation, and data modeling into a centralized BigQuery data warehouse, utilizing both in-warehouse SQL and external processing with pandas.
- Leveraged Apache Airflow for data orchestration, allowing scheduled and dependency-aware executions of ETL pipelines.
- Integrated data validation and quality checks into ETL workflows to support data governance and ensure data integrity.
- Built real-time monitoring dashboards using Looker to track pipeline health, detect failures proactively, and ensure SLA compliance.
- Maintained CI/CD pipelines using GitHub Actions to support automated deployment and testing.
- Utilized Terraform to manage and provision infrastructure following infrastructure-as-code practices.

Key Tech Stack: Python, Pandas, SQL, GCP, Big Query, PostgreSQL, Airflow, Terraform, Docker, GitHub

EDUCATION

Cornerstone International Community College, BC, Canada(May 2025 - Present) : Data Science Program

Le Wagon Toyko, Tokyo, Japan(Jan 2022 - Mar 2022) : Data Science - Batch 796

University of Oregon, Oregon, US(Sep 2018 - Dec 2020) : BS in Marketing

Green River College, Washington, US(Sep 2016 - Jun 2018) : AA in Business