GAGANDEEP SINGH

gaga.analytics@protonmail.com | 647-947-4491 | Toronto | Github | Portfolio

Data Engineer/Data Leader specialising in data governance, metadata management, scalable data pipelines, ETL/ELT, and big data technologies to transform raw data into high-impact business solutions.

CERTIFICATIONS AND SKILLS

CERTIFICATIONS:

- AWS Data Engineering Associate   Sept 2025
- Certified Data Management Professionals   Preparing

SKILLS

- Technical: AWS, GCP, Python, SQL, Apache Airflow, dbt, CI/CD, Docker, PySpark, BigQuery, Microsoft Fabric, SQL Server, SSIS, SSAS, Tableau, Power BI, Jupyter Notebook, git, data modeling
- General: Leadership, Stakeholder Management, Storytelling, Critical Thinking, Communication, Business Acumen, Agile Methods

RELEVANT EXPERIENCE (Overall 10+ Years)

Sabbatical   **January 2025 – March 2025**

Sole Proprietorship   **February 2024 – December 2024**
Analytics Engineer   *Tallinn, Estonia*

- Engineered and deployed scalable batch data pipelines on AWS for a Spanish e-commerce marketplace, enabling advanced analytics and reporting.
- Designed and built a digital inventory management solution that improved product visibility, increased product availability and reduced waste by 30% for two grocery traders.
- Engineered end-to-end ML pipelines to prepare and process data for a custom generative model, enabling the creation of unique images for a designer.
- Developed and maintained data pipelines for Averto, a real-time alerting application for natural disasters, ensuring high-velocity data ingestion and low-latency processing.
- Built and maintained data pipelines to prepare and transform data for multiple machine learning applications, including customer segmentation, LTV (Lifetime Value), risk, and fraud detection for a financial startup.

Brightvision   **July 2023 – December 2023**
Analytics Consultant   *Tallinn, Estonia*

Led the design and implementation of a modern data warehouse on Google BigQuery using a Data Mesh architecture, while building ELT pipelines to integrate data from disparate sources. Deployed dbt to automate data transformations and ensure data quality, directly enabling key customer, marketing, and product analytics and powering BI dashboards for business stakeholders.

Estateguru   **March 2022 – May 2023**
Data Lead   *Tallinn, Estonia*

- Designed, deployed and managed the foundational data stack and infrastructure on AWS and GCP based on a Data Mesh architecture and star schema, which included both batch and streaming pipelines. This initiative established a framework for data quality, metadata management, lineage, documentation and data governance.
- Led the implementation of BI and reporting processes in Power BI and Tableau, providing stakeholders with critical insights for decision-making.
- Developed and deployed data science models for borrower risk, lead scoring, and SEO, which decreased loan approval time by 50% and reduced sales interactions by at least 40%.
- Improved web analytics and funnel analysis, leading to a 20% reduction in churn and a 10% increase in NPS by addressing key investor and borrower pain points.
- Hired, managed, and led a team of 6 data professionals (4 Data Analysts, 1 Data Engineer, 1 Data Scientist), fostering a culture of data excellence.

Admirals   August 2019 – December 2021
Tallinn, Estonia

Product Owner, Data Science   February 2020 – December 2021

- Developed streaming and batch data pipelines, data quality checks to integrating data from disparate sources for feature engineering and training various product, customer, marketing, risk and sales related high impact machine learning models. Also build production pipelines and APIs for close to real time inference from these models.
- Led the migration of all data services to Google Cloud and managed the adoption of BigQuery as the primary data lake and warehouse, ensuring data quality, metadata management and governance.
- Designed, built and deployed data pipelines for ELT from GA3(Universal Analytics) to Bigquery. Managed internal servers, BigQuery administration; deployment, access & upgradation of internal tools.
- Managed the data team for over 8 months using agile methods, ensuring project efficiency and successful delivery.

Data Analyst   August 2019 – January 2020

- Engineered data transformations using BigQuery SQL and dbt to prepare data for KPI-driven dashboards, ensuring accuracy and efficiency.
- Drove a 15% increase in revenue by designing and deploying over 10 custom BI dashboards in Power BI and Tableau, while also optimizing costs and marketing strategies through web analytics and advanced modelling.

Creditstar Group   March 2018 – July 2019
Data Analyst   *Tallinn, Estonia*

- ETL pipelines, data models and data transformations using python and SQL to integrate data from 10+ MySQL databases, web scraping and from external sources, ensuring data reliability and accessibility.
- Led the proposal and production implementation of Apache Airflow to orchestrate data workflows, significantly improving automation and efficiency.
- Developed and deployed machine learning models to detect financial fraud across 7 European markets, while also building a comprehensive customer analytics framework to drive business improvements and provide actionable insights.

IIT Mandi   Sept 2013 – May 2014
Senior Research Fellow   *Mandi, India*

- Engineered and deployed machine learning models for Text-to-Speech (TTS) systems in Indian languages, while managing the full infrastructure and codebase.
- Improved TTS system accuracy by 20% by creating a new font for the Rajasthani language and training a custom Optical Character Recognition (OCR) model to extract text from PDFs.

Amdocs   Aug 2010 – July2013
Subject Matter Expert   *Limassol, Cyprus*

I drove a 30% increase in team productivity by implementing new processes and providing targeted training. In my role, I was responsible for the full lifecycle of telecommunications applications, including the development and deployment of CRM, Billing, and PRM systems in C++ and Java. I also streamlined operations by automating configuration management and managed the administration of Oracle databases.

EDUCATION

Masters in Computer Science ([Thesis](#) Based)   June 2015 – June 2018
University of Tartu, Estonia