Atish Dhamala

✉ atish.dhamala@gmail.com — 📞 (605)-910-2476 — 📍 Vermillion, SD — 🔗 LinkedIn.com/in/atishdhamala — ⭘ github.com/Adh101

SUMMARY

Data/Analytics Engineer with experience designing scalable data pipelines, warehouses, and analytics platforms using Python, SQL, Spark, Airflow, cloud tools etc. Adept at improving pipeline efficiency, ensuring data quality, and applying machine learning for actionable insights and automation in data-driven teams.

EDUCATION

Masters in Computer Science - Artificial Intelligence   *Aug. 2024 - Dec. 2026*
University of South Dakota   Vermillion, South Dakota, US

Bachelor of Engineering in Information Technology   *Sept. 2016 – Aug. 2022*
Pokhara University   Nepal

SKILLS

Programming Languages: Python, SQL, R, Javascript, HTML/CSS
Databases: PostgreSQL, MySQL, Snowflake
Cloud & DevOps: AWS (EC2, S3, Glue, Athena), GCP, Docker
Data & ML Tools: Apache Spark, Apache Flink, Apache Kafka, Apache Airflow, Databricks, Pandas, Scikit-learn, PyTorch, TensorFlow, Tableau, PowerBI, Seaborn, Numpy, Pandas, Matplotlib, Seaborn, MS Excel
Monitoring & Version Control: Logging, Airflow Email Alerts, GitHub Actions (CI/CD), Git, Github
Knowledge Domains: Data Modeling, ETL Pipelines, Software Engineering Principles, Data Warehousing, Machine Learning, Computer Vision

WORK EXPERIENCE

SuperAI Brain Inc.   *July 2023 – Aug. 2024*
Data Engineer - Remote   San Francisco, United States
- Designed and implemented ETL pipelines using AWS Glue, Databricks, and Apache Spark that processed over 2M records, reducing runtime by 40% and achieving 25% cost savings.
- Developed a Python-based data validation layer to flag 95% of anomalies, reducing manual efforts by over 20 hours per week while ensuring robust data quality.
- Built and managed Snowflake data warehouses with integrated testing frameworks including unit tests, integration tests, and data quality validation checks, reducing production issues by 85% across 20+ data sources.
- Integrated business intelligence tools such as Power BI and Tableau with an in-house LLM to enable natural language-based ad hoc enterprise data analysis.

Ingrails Pvt. Ltd.   *Mar. 2022 - Feb. 2024*
Research Analyst - Data   Lalitpur, Nepal
- Constructed data pipelines in Python and SQL to ingest, clean, and transform multi-source marketing data, enhancing data quality and consistency for downstream applications.
- Analyzed customer behavior trends across 100K+ records, assisting the product team in identifying retention risks and reducing churn by 30%.
- Created Tableau dashboards and automated reporting tools for cross-functional teams, reducing manual reporting time by 40%.

TechAxis Nepal   *Aug. 2023 - Dec. 2023*
Teaching Instructor   Lalitpur, Nepal
- Courses taught: Data Analysis with Python, Machine Learning with Python

Freshktm Pvt. Ltd.   *Apr. 2021 - Feb. 2022*
Data Analyst   Bhaktapur, Nepal
- Improved supply chain efficiency by 20% by building SQL data pipelines integrating ERP and inventory data.
- Developed and maintained Tableau dashboards with weekly refresh, reducing executive reporting time by 30%.
- Forecasted market demand using ARIMA and Prophet models, improving prediction accuracy by 25% and reducing stockouts by 15%.

PROJECTS

End-to-end Modern Data Warehouse & BI Automation   *2025*
PostgreSQL, Python, gspread,Tableau, Github, Notion
- Designed Medallion-structured Data Warehouse (Bronze–Silver–Gold) integrating ERP & CRM data into PostgreSQL.
- Developed a fully automated analytics pipeline using Python and Google Sheets API to feed a dynamic, interactive Tableau Public dashboard from PostgreSQL data warehouse, visualizing KPIs such as Total Sales, AOV, and regional/category performance.

Real-Time Stock Market Data Pipeline for Scalable Analytics   *2025*
Python, SQL, Apache Kafka, AWS EC2, AWS S3, Glue, Athena, Linux, Jupyter

- Developed a real-time data pipeline to support stock market analytics use cases such as price movement monitoring, anomaly detection, and trade signal generation.
- Enabled scalable ingestion of live stock data using Kafka producers and consumers, reducing latency in data delivery and eliminating batch processing delays.
- Streamlined data storage by automating structured JSON exports to AWS S3, ensuring durability and availability for downstream analytics and machine learning workflows.
- Solved infrastructure-level challenges (e.g., IAM permission conflicts, broker connection failures, and EC2 service misconfigs) to ensure reliable data streaming and fault tolerance.

Real-Time Currency Exchange ETL Pipeline *2024*

Python, PostgreSQL, REST API, GitHub Actions, launchd
- Built automated ETL pipeline to extract daily exchange rates via REST API, transform using Pandas, and load into PostgreSQL.
- Scheduled daily runs using GitHub Actions and secured credentials with GitHub Secrets for production-like automation.

Brain MRI 3D Image Segmentation using Deep Learning *2024*

Python, TensorFlow, Keras, VSCode
- Implemented 3D brain MRI segmentation models using U-Net and Attention U-Net architectures to segment brain tumors and structures from the BraTS 2020 dataset, achieving high precision in identifying tumor cores, edema, and enhancing regions with Dice Coefficients and IoU metrics.
- Developed end-to-end training pipeline leveraging TensorFlow and Keras to preprocess and augment 3D MRI data, fine-tune models using Adam optimizer with Total Loss, and evaluate segmentation performance, resulting in improved segmentation accuracy using IoU and Dice Coefficient, achieving accurate segmentation (Final IoU: 0.72 training, 0.63 validation).

VOLUNTEERING

- **AIESEC in Nepal** *May. 2018 - Mar. 2019*
  Team Leader   Kathmandu, Nepal
- **AIESEC in Nepal** *Oct. 2017 - May. 2018*
  Team Member   Kathmandu, Nepal

Certifications

- **Data Scientist with Python** *2020*

- **SQL Fundamentals Track** *2021*

- **Associate Data Engineer Track** *2025*

- **Datacamp Data Fellowship, Code For Nepal** *2020-2025*

- **AWS Data Engineer Associate Certification** *In Progress*