



Data Engineer

Location: Argentina (remote)

About the Role

We are seeking a talented and motivated Data Engineer to join our data engineering team. In this role, you will design, build, and maintain the data pipelines and data models that power our freight and parcel audit, analytics, and AI solutions. You'll collaborate with software engineers, analysts, data scientists, and business teams to deliver reliable, scalable, and well-governed data, in the logistics contract negotiation and optimization and audit business.

What You'll Do

- Design, build, and maintain scalable data pipelines to ingest, transform, and load data from carriers, clients, and internal systems.
- Develop and maintain data models, data warehouse, and data lake structures that support reporting, analytics, and AI initiatives.
- Write clean, efficient, and well-documented SQL and Python code.
- Implement data quality checks, validation, and monitoring to ensure data accuracy and reliability.
- Optimize pipeline and query performance for large, high-volume datasets.
- Participate in code reviews and contribute to team best practices and data engineering standards.
- Collaborate with product managers, internal customers, analysts, data scientists, and software engineers to define and deliver data solutions.
- Troubleshoot, debug, and resolve pipeline failures and data issues, working with Data Operations on root cause and prevention.
- Support data security, access control, and governance practices.
- Stay current with emerging data technologies and industry trends, including AI tools that improve engineering productivity.

Required

- Bachelor's degree in Computer Science, Information Systems, Engineering, or a related field (or equivalent experience).
- Advanced SQL experience, including complex joins, CTEs, window functions, and performance tuning.
- Databases: MSSQL and PostgreSQL
- Languages: Python (pandas)
- Experience designing and building ETL/ELT pipelines for high-volume data.



- Experience with workflow orchestration and integration tools (e.g., Airflow)
- Understanding of data modeling concepts (dimensional modeling, normalization, data warehouse and lakehouse design) and performance/optimization implications
- Knowledge of containerization tools (e.g., Docker, Kubernetes)
- Experience with cloud data platforms (Azure preferred; e.g., Azure Data Lake, Synapse).
- Familiarity with CI/CD pipelines, Git, and DevOps practices.

Preferred

- Pyspark knowledge is a plus
- Experience with Azure ML Studio,
- Experience with freight, parcel, invoice, or other transactional data is a plus.