

Wajih Arfaoui

Senior Data Engineer

Paris, France | +33 (0)7 58 67 88 27 | hello@wajiharfaoui.pro

linkedin.com/in/wajih-arfaoui-658aa2105/

Professional Summary

Senior Data Engineer building cloud data platforms across Azure, GCP, and AWS. Experienced in PostgreSQL data modeling, Python development, and analytics pipelines using dbt and Snowflake. Combines hands-on engineering with data governance, performance optimization, and mentoring, working closely with product and API teams to deliver reliable data solutions.

Technical Skills

Languages & development: Python, SQL, object-oriented design, reusable API integrations

Databases & modeling: PostgreSQL, MySQL, Snowflake, Amazon Redshift, BigQuery, OLTP and analytical modeling

Pipelines & processing: dbt, Dagster, Apache Airflow, Azure Data Factory, Apache Spark / PySpark, Databricks

Delivery & cloud: Flyway, Terraform, Git, Azure DevOps, GitHub Actions, Azure, GCP, AWS

Governance & BI: Data cataloging, lineage, data quality, monitoring, Power BI, Amazon QuickSight

Professional Experience

Senior Data Engineer | **EY France**

September 2023 – Present

Paris, France

- Designed PostgreSQL data models for OLTP applications, translating business requirements into normalized schemas, relational constraints, and indexing strategies.
- Implemented version-controlled schema migrations with Flyway and CI/CD workflows in Azure DevOps, incorporating automated validation, testing, and code review.
- Built reusable, object-oriented Python ingestion components for third-party APIs, standardizing authentication, pagination, and error handling.
- Developed Dagster-orchestrated ELT pipelines and dbt models in Snowflake to deliver reusable analytical datasets.
- Optimized PostgreSQL queries through indexing and SQL tuning, and refined Snowflake transformations and data models to improve performance and resource efficiency.
- Established data cataloging and governance practices covering metadata, ownership, lineage, and data quality standards.
- Coordinated with Scrum and backend/API teams to align data models, interface contracts, and dependencies with Agile feature delivery.
- Managed and mentored junior data engineers through code reviews and technical guidance; co-led the Data Community of Practice and defined monitoring and SLA/SLO standards.

Tech Stack: Python, SQL, PostgreSQL, Flyway, Dagster, dbt, Snowflake, Azure, GCP, Azure DevOps, Git.

Education

MSc in Big Data Analytics | IESEG School of Management

2021–2022

Lille, France. Coursework: data engineering, machine learning, and applied analytics.

BSc in Data Analytics & IT | Tunis Business School

2015–2019

Tunis, Tunisia. Thesis: evaluating customer satisfaction using web scraping and NLP.

Professional Experience (continued)

Data Engineer | Mangabey

January 2023 – September 2023

Paris, France

- Designed and delivered end-to-end data architectures on GCP for pharmaceutical clients, spanning ingestion, transformation, and analytics.
- Migrated single-process Python workflows to distributed PySpark pipelines to improve performance and scalability for datasets containing several billion rows.
- Orchestrated pipeline schedules and dependencies with Apache Airflow; developed SQL transformations and modular dbt models with automated tests and documentation.
- Automated deployments and environment management using GitHub Actions, and implemented data quality checks, monitoring, and failure alerts.
- Supervised and mentored a Data Engineer intern and a Data Analyst within the data quality sub-team, coordinating tasks and providing technical guidance.

Tech Stack: Python, SQL, dbt, Apache Spark / PySpark, GCP, Apache Airflow, Power BI, GitHub Actions.

BI Engineer | Amazon France

June 2022 – January 2023

Paris, France

- Managed and optimized Amazon Redshift clusters supporting large-scale logistics analytics and operational performance reporting.
- Built automated pipelines and SQL transformations using Python and dbt to deliver business-critical KPIs across European markets.
- Developed Amazon QuickSight dashboards and served as the primary contact for delivery and logistics KPIs, aligning reporting requirements and metric definitions with stakeholders.
- Presented performance trends and improvement opportunities during business reviews with operational teams and leadership.

Tech Stack: AWS, Amazon Redshift, Python, SQL, dbt, Amazon QuickSight.

BI Engineer | Edelman Intelligence

November 2019 – April 2021

Paris, France

- Built Python and SQL transformation pipelines in BigQuery to centralize and harmonize earned, owned, and paid media datasets.
- Developed Power BI dashboards and analytical reports to monitor media performance and support business decisions.
- Implemented data workflows to improve reporting accuracy, consistency, and timely delivery.

Tech Stack: Google BigQuery, Python, SQL, Power BI, Excel.

Certifications

Snowflake SnowPro Core Certification

Snowflake SnowPro Advanced: Data Engineer Certification

Google Cloud Professional Data Engineer Certification (*in progress*)

Languages

Arabic (5/5) | French (4/5) | English (4/5) | German (2/5)

Self-assessed proficiency ratings.

Interests

Football, kitesurfing, travel, and cultural exploration.