Best Technology Talent For
Companies Around The World



ncube

Camila

Data Engineer

Summary

Senior Data Engineer and Data Architect with 7 years of experience designing and implementing robust data architectures, ETL/ELT pipelines, and cloud-based data platforms across GCP, AWS, and Snowflake. Brings a strong software engineering discipline to data work: CI/CD pipelines, Infrastructure as Code, Agile delivery, and rigorous data quality practices.

Experienced across both batch and streaming pipeline architectures using Python, PySpark, SQL, and Airflow, with deep expertise in data modelling, data lake design, and modern ELT tooling (dbt, Snowpipe, Dataform). Track record of migrating legacy on-premises systems to cloud-native architectures and delivering scalable, maintainable data infrastructure across Health Insurance, Reinsurance, Media, Entertainment, Retail, Education, and Marketing domains.

Collaborative and internationally experienced: worked with cross-functional teams across Europe, Latin America, and the US, partnering closely with Business, Data Science, BI, and Engineering stakeholders to translate requirements into production-grade data solutions.

TECHNICAL SKILLS:

- ETL/ELT pipelines, batch and streaming processing, data modelling, data quality, data cleansing, EDA
- Python, PySpark, SQL, CTEs, API integration, dbt
- Apache Airflow (Expert), Kafka, RabbitMQ
- GCP (BigQuery, Cloud Functions, Dataform), AWS (S3, Glue, Lambda, Redshift, DynamoDB, MWAA), Snowflake (Snowpipe)
- Docker, Kubernetes, Terraform, GitHub, GitLab, BitBucket, CI/CD pipelines
- Oracle ODI, Oracle Database, Alteryx, Jupyter Notebook, PowerBI
- Software Engineering best practices, Agile, Data Architecture

Professional Career

Senior Data Architect

December 2023 – Present | **Brazil | Porto, Portugal**

Project Info:

A global leader in designing, manufacturing, and supporting personal computers, servers, data storage devices, network switches, and information technology (IT) services. Marketing technologies.

Responsibilities:

- Strategic development and implementation of data architectures: researched, tested, and implemented robust data architectures and ETL/ELT pipelines tailored to project-specific requirements across Media and Entertainment and Education verticals
- Led migration of data from MySQL databases to Snowflake, using Snowpipe and dbt to build and orchestrate ELT pipelines deployed and managed under a CI/CD workflow via GitHub
- Managed ETL/ELT processes and data modelling, extracting data from diverse sources including files, databases, and APIs, using AWS (S3, Glue, Lambda, Redshift, DynamoDB, MWAA), dbt, PySpark, Python, SQL, and CTEs
- Operated within CI/CD pipelines and Agile methodologies, collaborating with international teams across Europe and the US to ensure seamless delivery and coordination

Software and technologies:

Snowflake, Snowpipe, dbt, AWS (S3, Glue, Lambda, Redshift, DynamoDB, MWAA), Python, PySpark, SQL, GitHub, CI/CD, Airflow

Data Engineer

July 2022 – December 2023 | **Rio de Janeiro, Brazil**

Project Info:

Media and Entertainment and Education industry data engineering projects.

Responsibilities:

- Researched, tested, and implemented data architectures and ETL/ELT pipelines tailored to project requirements across Media and Education domains
- Migrated data from MySQL databases to Snowflake using Snowpipe and dbt, orchestrating ELT pipelines within a CI/CD GitHub workflow for reliability and scalability
- Managed data modelling and ETL/ELT processes extracting from files, databases, and APIs, using AWS (S3, Glue, Lambda, Redshift, DynamoDB, MWAA), dbt, PySpark, Python, SQL, and CTEs

- Collaborated with international teams in Agile environments, ensuring effective communication and delivery across distributed, cross-functional groups

Software and technologies:

Snowflake, Snowpipe, dbt, AWS (S3, Glue, Lambda, Redshift, DynamoDB, MWAA), Python, PySpark, SQL, GitHub, CI/CD, Airflow

Data Engineer

September 2021 – July 2022 | **Rio de Janeiro, Brazil**

Project Info:

One of Brazil's largest cosmetics and personal care retail groups. Contributed to the design and implementation of Data Lake infrastructure and ETL pipelines on GCP, covering both batch and streaming operations across the retail business.

Responsibilities:

- Designed and implemented ETL processes for both batch and streaming operations on GCP, developing API extraction methods using Python and PySpark integrated with Cloud Functions and BigQuery
- Designed and implemented Airflow orchestrations to ensure efficiency, reliability, and scheduling precision across the full data pipeline
- Worked closely with Business, Data Science, and BI teams to understand requirements and translate them into effective, production-ready data pipelines
- Collaborated within an Agile team environment, contributing to a responsive and iterative data infrastructure development process

Software and technologies:

GCP (BigQuery, Cloud Functions), Apache Airflow, Python, PySpark, SQL, API integration, Docker, Kubernetes

Data Engineer Consultant

January 2019 – September 2021 | **Rio de Janeiro, Brazil**

Project Info:

A leading American multinational technology and consulting company specializes in computer hardware, software, cloud computing, artificial intelligence (AI), and quantum computing, operating in over 175 countries. Data engineering consultancy engagement across Health Insurance, Reinsurance, and Media and Entertainment industries. Covered Data Warehouse design, Oracle-to-cloud migration, ETL development, and BI reporting across multiple client projects.

Responsibilities:

- Designed and implemented Star Schema data modelling for Data Warehouse projects, crafting metrics and measures for OLAP systems to support PowerBI dashboards
- Developed batch ETL pipelines using Oracle ODI and Oracle Database, and contributed to cloud migration using GCP BigQuery and Dataform
- Implemented API data extraction using Python and processed diverse data formats (JSON, CSV) using normalisation and sanitisation techniques for Exploratory Data Analysis (EDA) in Jupyter Notebook with Python and PySpark
- Ingested data from CSV files and SQL databases using Alteryx and Oracle ODI, ensuring data integrity and processing efficiency
- Prepared comprehensive As-Is and To-Be documentation, created PowerBI reports, and participated in business unit meetings to align data solutions with stakeholder requirements

Software and technologies:

Oracle ODI, Oracle Database, GCP (BigQuery, Dataform), Python, PySpark, Alteryx, PowerBI, SQL, Jupyter Notebook, JSON/CSV

Education

2019 – 2023

Fluminense Federal University, Rio de Janeiro, Brazil
Bachelor's Degree in Computer Systems Technology

Certifications

Astronomer Certification for Apache Airflow 3 Fundamentals
Astronomer Certification DAG Authoring for Apache Airflow 3

Foreign Languages Skills

Portuguese: Native
English: Upper-Intermediate