

# João | Data Engineer

## SUMMARY

He is a Tech Lead and Data Engineer with over 8 years of experience designing data pipelines, building dashboards, and delivering business intelligence solutions. He currently leads a team of 10 people and works night shifts at his company. His career began with data analysis, building dashboards and APIs, and later transitioned into data engineering roles.

He has worked across industries such as marketing and government, developing end-to-end data workflows and embedding analytics into business systems. His tech stack includes Python (9+ years), Java (3 years), and TypeScript, with extensive use of tools like Airflow (5 years), DBT (2 years), PySpark, Docker, and Apache Spark. He also has experience with unit testing using Pytest, CI/CD pipelines via GitLab and GitHub Actions, and NLP (3 years), along with limited exposure to Generative AI and augmentation methods.

In terms of cloud experience, he has worked with AWS for 7–8 years, leveraging services like S3, Lambda, Glue, Redshift, Athena, and EC2, and had limited exposure to Azure through one client engagement. He also brings strong BI experience with Power BI and Tableau (both for 6 years), and has worked with multiple databases including PostgreSQL, MySQL, SQL Server, MariaDB, MongoDB, Cassandra, and Redshift. He used Snowflake extensively until 2022 (almost 3 years of experience).

Although he has not worked with Kubernetes, he has a solid background in infrastructure as code, having used Terraform in past projects. He also currently uses AI tools like GitHub Copilot in his workflow. His broad skill set, leadership capabilities, and cross-functional experience enable him to bridge backend development and data strategy effectively.

## CORE SKILLS

- |                      |              |                   |
|----------------------|--------------|-------------------|
| • ETL/ELT Pipelines  | • MongoDB    | • Apache Superset |
| • Data Modeling      | • Redshift   | • FastAPI         |
| • API Development    | • Python     | • Flask           |
| • Embedded Analytics | • JavaScript | • Next.js         |
| • PostgreSQL         | • Go         | • Fastify         |
| • SQL Server         | • Docker     | • AWS             |
| • MariaDB            | • Git        | • Vercel          |
|                      | • CI/CD      |                   |

## WORK HISTORY

### XXXXXX

Tech Lead

*Apr 2023 – Present*

- Leads a team of 12 engineers, focusing on collaboration, technical excellence, and delivering solutions aligned with both business and public-sector needs.
- Designed and implemented a Business Intelligence (BI) system embedded in the company's ERP platform, used by government clients across finance, healthcare, education, and environmental sectors.
- Built and maintained full ETL pipelines to extract data from ERP databases, process it, and load it into isolated storage environments using AWS Redshift or PostgreSQL, depending on the municipality's infrastructure.
- Created interactive dashboards with Apache Superset, embedded directly into the ERP.
- Developed a mobile notification system to deliver real-time dashboards and key performance metrics to mayors and department heads, enabling timely, data-driven decision-making.
- Demonstrated strong technical leadership in system architecture, code reviews, problem-solving, and team mentoring.
- Key technologies: Python, TypeScript, .NET Framework, Go, REST APIs, RabbitMQ, CI/CD, Docker, Redshift, PostgreSQL, Apache Superset, HTML5, OOP, System Design.

### EL Producoes de Software

Data Engineer

*Mar 2021 – Jul 2023*

- Contributed to extending the company's ERP platform through BI dashboards, APIs, and mobile apps to help municipal governments make informed decisions.
- Designed and maintained ETL pipelines extracting data from ERP systems hosted on AWS and on-premise environments, transforming it for analytics in Amazon Redshift.
- Collaborated with data analysts to define KPIs and build intuitive dashboards embedded in the ERP and accessible via mobile apps.
- Key technologies: Python, SQL, Amazon Redshift, Amazon S3, FastAPI, Flask, Next.js, Docker, Git, React.js, Linux, XML, CI/CD.

### Tegrus – Vitória, Espírito Santo, Brazil

Data Engineer

*Jan 2020 – Aug 2021*

- Designed and maintained robust ETL processes and data pipelines using tools such as DBT, Airflow, and AWS services (S3, Lambda, Glue, Redshift).
- Supported data infrastructure and architecture, ensuring scalability and performance.

- Ensured data reliability and consistency across the entire pipeline to enable accurate reporting and decision-making.
- Collaborated with analytics and product teams to monitor dashboards and track key business KPIs.
- Contributed to the continuous improvement and stability of the data platform.
- Tech stack: Python, Docker, Airflow, Rocketry, Go, AWS, DBT.

## **Tegrus – Vitória, Espírito Santo, Brazil**

Data Engineer

*Jan 2020 – Aug 2021*

- Built BI dashboards and developed ETL processes to integrate and transform data from multiple sources.
- Worked closely with Product Managers to define and track strategic KPIs, delivering actionable insights.
- Ensured data quality, reliability, and usability in reports and visualizations for business and product teams.
- Contributed to the creation of performance indicators aligned with organizational objectives.
- Tech stack: Python, Docker, AWS (S3, Redshift), Git, XML, BI tools

## **Haus Software House**

Backend Developer

*2015-01-01 - 2017-12-31*

- Built web platforms, mobile apps, and BI tools for a variety of clients across different industries.
- Led a key project for FarmlyClub, a coffee export platform that had previously failed to launch under another provider.
- Took over a broken system and rebuilt it from scratch, making it fully operational and reliable.
- Enabled the client to manage inventory, track sales, and generate reports through the platform.
- Developed Power BI dashboards connected to the production database to monitor platform performance and growth in real time.

## **EL Producoes de Software**

Data Analyst

*Jun 2016 – Dec 2017*

- Worked as a Business Intelligence Analyst, designing and implementing ETL processes (Extract, Transform, Load) for client-facing dashboards.
- Developed data integration workflows to segment and structure information according to business context.

- Built dashboards in Tableau to support key stakeholders such as the Mayor and the City Health Secretary, enabling monitoring of city-wide KPIs.
- Managed structured and unstructured data using SQL, Microsoft SQL Server, and PostgreSQL.
- Applied knowledge of multithreading, data structures, and algorithms to optimize data handling and processing.
- Used Amazon S3 for data storage and Git for version control.
- Contributed to software development efforts with a strong foundation in programming and software engineering best practices.
- Technologies & Tools: Tableau, ETL, SQL, PostgreSQL, Microsoft SQL Server, Amazon S3, Git, Multithreading, Data Structures, Algorithms

## **Tegrus**

### **Data Engineer**

*Jan 2015 – Dec 2016*

- Designed and implemented data pipelines and dashboards for multiple consulting projects.
- Key contributor to V1, a ride-hailing app that became the first in its market to reach breakeven.
- Collaborated directly with the CTO and CEO to analyze vehicle distribution across city regions, prior to implementing dispatch rules.
- Led data-driven recommendations that resulted in a 20–25% fleet reduction, optimizing operational costs and improving availability in high-demand zones

## **EDUCATION**

### **Federal University of Espírito Santo (Official)**

**Degree:** Bachelor's Degree in Statistics

**Academic Discipline:** Time Series

**Period of Study / Expected Graduation:** 2018 – 2023

### **USP – University of São Paulo**

**Degree:** Statistics

**Grade:** Not Completed

**Period of Study:** 2015 – 2017

## **LANGUAGES**

Portuguese: Native

English: B2+