

João Pereira dos Santos

joaopereiradsantos@gmail.com | [linkedin.com/in/joaopereiradsantos](https://www.linkedin.com/in/joaopereiradsantos) | [joaopereiradsantos.com](https://www.joaopereiradsantos.com)

EXPERIENCE

Senior Data Engineer — Autonomous Procurement

Transporeon

Dec. 2024 – Present

Barcelona, Spain (Full Remote)

Business Intelligence Engineer II — Transportation & Supply Chain

Amazon

Feb. 2023 – Nov. 2024

Barcelona, Spain (Hybrid)

EU SC FC Launch & Transfer Analytics Team

- Led the EU SC Network Planning Tech MBR, coordinating 5 subteams across 25 tech projects. Collaborated closely with 5 tech members and 5 business stakeholders, orchestrating the successful delivery of AWS solutions
- Developed and deployed the Pod Selection Algorithm (PSA) DasBoard UI, enhancing accessibility and functionality. Created a formalized approach to the production model, democratizing accessibility for non-technical users. DasBoard automates internal application productionization for developers and is deployed on AWS Fargate
- Rolled out EXTREME: Excel to Redshift Migration Engine to 20k+ EU Central Flow users. (EXTREME is centered around a micro-service orchestration workflow using AWS Step Functions)
- Optimized QuickSight by descheduling 50+ unused datasets, reducing Redshift processing by 6+ hours weekly. Leveraged CloudTrail logs and Python for automated analysis

Business Intelligence Engineer — Transportation & Supply Chain

Amazon

Nov. 2021 – Feb. 2023

London, United Kingdom (Hybrid)

EU SC Data Engineering Team

- Facilitated data access for 300+ users. Managed 4 AWS accounts and 2 Redshift clusters (+1 PB & +2k tables)
- Developed a native AWS solution for seamless conversion of .xlsx files to Redshift tables (EXTREME: Excel to Redshift Migration Engine), supporting 20+ EU Supply Chain active users with ~ 5 daily service calls
- Developed an Amazon CloudWatch dashboard, averting slowness with automated alerts and enhancing proactive Redshift cluster issue response. Viewed by 100+ users ~ 75 times weekly
- Deployed 3 AWS solutions (AutoTag, AutoScheduler, AutoStop) to automate best practices and achieve +70% cost savings on EC2 and SageMaker instances. AutoTag enhances cost allocation via Lambda and CloudWatch. AutoScheduler reduces operational costs by scheduling EC2 and RDS instances. SageMaker AutoStop achieved a 98% decrease YoY in costs

EU SC FC Launch & Transfer Analytics Team

- Guided 10+ interns and junior BI professionals via Agile methodologies, ensuring software development best practices. Mentored and managed technical projects, and conducted interviews to recruit for the organization
- Developed an algorithm utilizing linear programming for optimized pod selection in a new robotic fulfillment center. Achieved a 10% improvement in volume per pod and 15% decrease in units per pod. In 2023, facilitated the transfer of 43MM units to 3 sites, totaling 3MM ft³
- Engineered a Python script utilizing an internal API for pre-launch outbound transshipments control, focusing on processing capacity thresholds. Deployed on Amazon EC2
- Developed 2 comprehensive dashboards for tracking new fulfillment centers ramp-up metrics, including the creation and management of data pipelines. 1k+ views from over 60 users. Used ~ 30 times weekly

Business Intelligence Intern — Transportation & Supply Chain

Amazon

Feb. 2021 – Jul. 2021

Luxembourg, Luxembourg (Hybrid)

EU SC Inbound Cross Dock Team

- Created a new metric to measure the bias in volume distribution across fulfillment centers (spread bias)
- Explored hard constraints in optimization models and impact on placement
- Analyzed the influence of sort share on the selection of unique items and cross-border fulfillment costs
- Created 2 dashboards to monitor and enhance tote utilization, as well as optimize fluid loading for improved productivity, sustainability, and cost savings

EDUCATION

Faculty of Engineering - University of Porto

MEng in Chemical Engineering - Processes and Product

Sep. 2015 – July 2020

Porto, Portugal

PROJECTS

Tennis (Consensus) Betting Bot | *Python (BeautifulSoup, Selenium), AWS*

Aug. 2023

- Implemented a betting bot focusing on tennis matches, utilizing Python and AWS services
- Employed web scraping techniques, upgrading data acquisition methods for comprehensive coverage
- Integrated consensus probability for fair betting decisions based on bookmakers' odds
- Conducted simulations using different betting strategies, assessing performance and yield

Tennis Betting Bot | *Python (TensorFlow, PyTorch), AWS*

May 2022

- Developed a machine learning model (neural network) to predict tennis match outcomes
- Trained the neural network for different court surfaces (clay, hard, grass) to enhance prediction accuracy
- Evaluated six betting strategies, including favoring favorites, underdogs, and value-based betting
- Incorporated email automation to notify about valuable bets meeting specific criteria (% threshold)

TECHNICAL SKILLS

Languages: Python, SQL (Redshift, Spark), HTML/CSS, R

Developer Tools: AWS, Git, Poetry, Pyenv, Docker, VS Code, PyCharm, Vim, Asana, Slack

Libraries: asyncio, awscli, boto3, BeautifulSoup, boto3, dash, flask, json, numpy, pandas, plotly, psycopg2, pytorch, requests, scikit-learn, selenium, statsmodels, tensorflow, toml, xpress, yaml

Orchestration: Apache Airflow, AWS Step Functions

Certifications & MOOC: AWS Certified Cloud Practitioner, Data Analyst w/ SQL Server, Interactive Python Dashboards with Plotly and Dash, Celonis Process Mining Expert, 365 Complete Data Science Bootcamp

LANGUAGES

Portuguese (Native), **English** (Fluent), **French** (Fluent), **Spanish** (Proficient)