



CLARA TRUJILLO SANTOS-OLMO

Senior Data Scientist · Experimentation & causal measurement · Pricing, promotion & retention
Madrid, Spain · +34 605 010 581 · claratrujillosantosolmo@gmail.com ·
<https://www.linkedin.com/in/clara-trujillo-santos-olmo-5966aa120/>
English [C1] · Spanish (native) · EU citizen — open to relocation within the EU

PROFILE

Senior Data Scientist with a background in Economics and 3 years in analytics consulting at NTT DATA, specialised in commercial decision models: Bayesian experimentation, forecasting and customer analytics across insurance and FMCG — taken end to end, from model design to the committee decision it supports.

EXPERIENCE

Senior II Data Scientist NTT DATA — Madrid

Jul 2026- Present

Data Scientist, Oct 2023 – Jun 2026 — promoted two grades in 12 months.

Embedded with insurance and consumer-goods clients. Coordinate a team of 3 analysts.

Demand Forecasting & Segmentation — eB2B platform, FMCG

- Built a **daily demand forecasting model at distributor–outlet–SKU level**, combining trend features (EWMA, moving averages, most-similar-week volume) with weather and event signals (league fixtures, public holidays); backtested over four months of orders.
- Solved the **intermittent demand problem** with linear interpolation and a statistical zero-assignment rule, correctly calling **78% of no-purchase weeks with 1.3% false zeros**.
- Designed the **behavioural segmentation of a 23,500-outlet network**: six clustering variables (recency, frequency, RFM, discount frequency, purchase profile, unique SKUs) plus a **peer sub-clustering layer** built on business type, environment, climate, geography and revenue percentile.

Promotional Efficiency & Customer Lifecycle — eB2B platform

- **Showed that promotional spend was returning nothing**: a net –€0.1 per € invested, as forward buying on the first coupon order (+€2) was cancelled out by cannibalisation on the following ones (–€0.5 each). **Bayesian Ridge regression over 417,000 orders, on an €883k discount spend**.
- Defined a **five-segment customer lifecycle** across 17,950 outlets, deriving every threshold analytically rather than by rule of thumb **including the 75% coupon-share cut-off where ROI differences between groups stabilise**.
- **Turned it into a per-segment activation plan**, each with its own KPI and measurement window, plus a KNN look-alike model to prioritise the offline outlets most likely to migrate online. Taken to committee as recommendations on promotional investment and pricing, with A/B tests designed to validate each action

Customer Experience Analytics — Insurance

- Built a **predictive detractor model** (Random Forest on authorisation data, **~75% recall**) and a **CSAT logistic regression model**, plus a **province–specialty friction scoring model** (Ridge, K-fold) ranking priority areas for service intervention.
- Diagnosed the NPS gap against market benchmarks on a cohort of **541 detractors**, showing **42%** stemmed from operational issues rather than coverage gaps; **delivered via committee summaries and Power BI dashboards**.

Data Platform & Team

- Designed a scalable **ETL pipeline in Databricks and Azure Data Lake** processing **5M records daily**, with automated data-quality validations that improved reliability by 50%.
- Work in **Git/GitHub and VS Code** with branch-based review and use **Claude Code** to automate refactoring, test scaffolding and documentation across recurring analytical work.

Junior Financial Auditor Ernst & Young — Madrid

Jan 2023- Sept 2023

- **IFRS-based financial analysis and audit testing** for international clients — data integrity validation, risk identification and reporting to stakeholders.

SKILLS

Languages & tools	Python (pandas, scikit-learn, NumPy, statsmodels), SQL, Git / GitHub, VS Code, Claude Code
Statistics & ML	Bayesian Ridge and linear regression, A/B testing and experimental design, demand forecasting, clustering (K-means, KNN), gradient boosting (LightGBM, XGBoost), Random Forest, recommender systems
Platforms & delivery	Databricks, Azure Data Lake, Apache Spark, ETL design · Power BI, Tableau, Streamlit (advanced)
Certifications	Azure AI Fundamentals (AI-900), Microsoft · Advanced Streamlit

EDUCATION

MSc Big Data, Data Science & BI Universidad Complutense de Madrid

Oct 2022 – Sep 2023

BSc Economics and Business Universidad de Castilla-La Mancha

Sep 2018 – Jun 2022