

Vanessa Schemes Martins Pinto

Senior Quantitative Researcher & Data Scientist | Experimentation & Decision Science (Statistician)

Portugal / Remote | +351 922 004 270 | vanessa.smartinsp@gmail.com | [linkedin.com/in/vanessasmp](https://www.linkedin.com/in/vanessasmp) | Portfolio: vsmmp.pages.dev

PROFESSIONAL SUMMARY

Statistician with extensive experience as a Senior Quantitative Researcher and Data Scientist, specializing in experimentation, causal inference, and applied econometric modeling for product and strategic business decisions, focusing on highly interpretable glass-box models.

Combines Python, R, and advanced SQL with rigorous probabilistic foundations to translate complex revenue, demand, and behavioral challenges into testable hypotheses, impact metrics, and counterfactual simulations. Experienced in robust experimental design (D-Optimal, laboratory, and quasi-experiments), discrete choice modeling, and attribute prioritization, with an emphasis on executive communication grounded in uncertainty bounds, methodological boundaries, and actionable decision criteria.

PROFESSIONAL EXPERIENCE

- **Quantitative Researcher & Data Scientist** | Chameleon Tecnologia | Remote | Feb 2023 – Present
 - **Technical Leadership & End-to-End Ownership:** Led quantitative research methodologies, sample design, and statistical validation end-to-end across all high-complexity studies. Served as the quantitative technical lead, conducting peer reviews, standardizing sampling protocols, and mentoring analysts and interns.
 - **Strategic Diagnostics & Key Drivers:** Conducted advanced quantitative intelligence studies leveraging Generalized Linear Models (GLMs), hypothesis testing, Key Driver Analysis (Shapley Value decomposition), Multiple Correspondence Analysis (MCA), and asymmetric impact grids (Kano / IPA) to isolate core drivers of retention and perceived brand equity.
 - **Discrete Choice Modeling & Demand Simulators:** Engineered analytical engines and shelf-choice simulators based on Latent Class Multinomial Logit (LC-MNL) and numerical optimization, enabling clients to simulate counterfactual pricing scenarios, product launches, and cross-elasticity portfolio dynamics.
 - **Optimal Experimental Design & Sampling:** Structured efficient experimental designs (D-Optimal Designs and balanced blocks) and A/B/quasi-experimental frameworks; implemented variance reduction methods (CUPED, post-stratification) and power calculations to maximize test sensitivity while minimizing collection costs and exposure time.
 - **Methodological Rigor & Econometric Diagnostics:** Addressed and corrected severe assumption violations in observational and experimental datasets — including diagnosis and remediation of multicollinearity (decomposition, regularization), heteroskedasticity (sandwich/robust standard errors, transformations), and censored/truncated data to mitigate selection bias and ensure consistent parameter estimation.
 - **Analytics Engineering & Interactive Interfaces:** Built modular data pipelines in Python, R, and advanced SQL, alongside interactive analytics interfaces (Streamlit), translating probabilistic models and analytical uncertainty into intuitive decision-support simulators.

Quantitative Research & Data Analyst Intern | Chameleon Tecnologia | Remote | Jun 2022 – Feb 2023

- **Processing Pipelines:** Built automated routines for microdata cleaning, transformation, and statistical consistency auditing in Python and R, reducing data preparation latency for complex studies.
- **Executive Reporting:** Assisted in quantitative analyses and technical report compilation, delivering final client deliverables through structured executive slide decks and written methodologies.

IT Infrastructure & Systems Support Analyst | ISATEC | Curitiba, Brazil | Mar 2013 – Feb 2018

- End-user technical support, local network infrastructure administration, server management, and incident resolution.

AREAS OF QUANTITATIVE EXPERTISE

- **Experimentation & Causal Decision Science:** Optimal experimental design (D-Optimal), A/B/n testing, sequential testing, variance reduction (CUPED), Bayesian inference (ROPE/HDI), Generalized Synthetic Control, Difference-in-Differences (DiD), Propensity Score Matching (PSM), and Double Machine Learning (DML).
- **Econometric Rigor & Bias Control:** Assumption diagnostics and non-spherical errors, heteroskedasticity-consistent sandwich covariance matrices (HC0–HC3, Newey-West), multicollinearity diagnostics (VIF, BKW spectral decomposition), censored and truncated data modeling (Tobit models, Heckman Two-Step), sample reweighting (Iterative Proportional Fitting / Raking), and survival analysis.
- **Product, Demand & Pricing Strategy:** Conjoint Analysis (CBC), MaxDiff, Latent Class Multinomial Logit (LC-MNL), utility decomposition (Part-Worths), cross-elasticity matrices, and constrained numerical optimization.
- **Marketing Science & Attribution:** Marketing Mix Modeling (MMM), Shapley decomposition, Adstock transformations (Weibull/Geometric), diminishing returns/saturation curves, and Multiple Correspondence Analysis (MCA).
- **Customer Experience & LTV:** Derived vs. Stated Importance, Importance-Performance Analysis (IPA), Penalty-Reward Analysis, longitudinal modeling, duration/survival models (Kaplan-Meier, Cox Proportional Hazards, AFT Models), and churn prevention.
- **Time Series & Forecasting:** State Space Models (SSM), Kalman Filtering (RTS), Unobserved Component Models (UCM), damped trends, trigonometric seasonality, and the ARIMA/SARIMAX family.
- **Machine Learning & Model Explainability:** Tree-based ensembles (XGBoost, Random Forest), probability calibration (Platt Scaling, Brier Score), threshold tuning for asymmetric business costs, and Explainable AI (SHAP, PDP).
- **Natural Language Processing (NLP):** Text mining, vector embeddings, semantic search, Named Entity Recognition (NER), topic modeling, and structured data extraction.

ANALYTICAL STACK & TOOLS

Languages & Scientific Computing: Python (Pandas, NumPy, SciPy, Scikit-learn, Statsmodels, PyMC, Streamlit, Plotly, spaCy), R (Tidyverse, ggplot2, lme4, survival, forecast, fable, sandwich).

Data Engineering & Cloud Data Warehouses: Advanced SQL (CTEs, Window Functions, query performance tuning), Snowflake, Google BigQuery, NoSQL (JSON), and Git.

AI-Assisted Development & Engineering: VS Code, Cursor, GitHub Copilot, and LLM-assisted workflows (code acceleration, refactoring, and automated testing).

Visualization & Executive Communication: Power BI (Star Schema modeling, advanced DAX), Tableau, and quantitative storytelling for technical and business leadership.

EDUCATION

Postgraduate Specialization in Natural Language Processing

Federal University of Goiás (UFG) | 2025 – 2026

Focus on unstructured data modeling, neural architectures, and optimization in information retrieval pipelines.

B.Sc. in Statistics (5-Year Honors Degree Equivalent)

Federal University of Paraná (UFPR) | Curitiba, Brazil | 2018 – 2023

Completed at a top-tier Brazilian public research university. Rigorous curriculum emphasizing Mathematical Statistics, Measure/Probability Theory, Classical and Bayesian Inference, Stochastic Processes, and Advanced Experimental Design. Undergraduate Thesis: Applied quantitative research on non-linear regime-switching time series (STARMAX models), integrating econometric theory and computational simulation.

Coursework in Computer Engineering

Federal University of Technology – Paraná (UTFPR) | Curitiba, Brazil | 2013 – 2017

Advanced data structures, algorithmic complexity, systems architecture, and scientific computing.

Associate Degree in Industrial Electronics

Federal University of Technology – Paraná (UTFPR) | Curitiba, Brazil | 2008 – 2012

Precision measurement instrumentation, control logic, and empirical testing methodologies.

LANGUAGES: Portuguese: Native • English: Professional Working Proficiency (Technical Reading & Writing)