

Vanessa Schemes Martins Pinto

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

Portugal / Remoto | +351 922 004 270 | vanessa.smartinsp@gmail.com | linkedin.com/in/vanessasmp | Portfólio: vsmp.pages.dev

RESUMO

Estatística com sólida atuação como Senior Quantitative Researcher e Data Scientist, especializada em experimentação, inferência causal e modelagem econométrica aplicada a decisões de negócio e produto com foco em modelos de alta explicabilidade (glass-box).

Combina Python, R e SQL avançado com fundamentação probabilística rigorosa para traduzir problemas complexos de receita e comportamento em hipóteses testáveis, métricas de impacto e simulações contrafactuais.

Experiência no desenho de experimentos robustos (D-Optimal, testes laboratoriais e quase-experimentais), modelagem de escolha e priorização de atributos, com ênfase na comunicação executiva de resultados ancorados em margens de incerteza, limites metodológicos e critérios acionáveis de decisão.

EXPERIÊNCIA PROFISSIONAL

Quantitative Researcher & Data Scientist | Chameleon Tecnologia | Remoto | fev. 2023 – presente

- **Liderança Técnica e Autonomia End-to-End:** Responsável pela definição metodológica, desenho amostral e validação estatística de ponta a ponta em todos os estudos quantitativos complexos da empresa. Liderança técnica de equipe de análise, conduzindo revisão, padronização de amostragem e capacitação técnica de analistas e estagiários.
- **Diagnósticos Estratégicos e Key Drivers:** Condução de estudos avançados de inteligência quantitativa utilizando GLMs, testes de hipótese, Key Driver Analysis, Decomposição de Shapley, análise perceptual (MCA) e matrizes de impacto assimétrico (Kano / IPA), isolando os atributos reais que governam a retenção e a percepção de valor.
- **Modelagem de Escolha e Simuladores de Demanda:** Desenvolvimento de motores analíticos e simuladores de prateleira baseados em Logit Multinomial com Classes Latentes (LC-MNL) e otimização numérica, permitindo a clientes simular cenários contrafactuais de precificação, lançamentos de produtos e elasticidade cruzada de portfólio.
- **Desenho Experimental Ótimo e Amostragem:** Estruturação de desenhos experimentais eficientes (D-Optimal Designs e blocos balanceados) e arquiteturas de testes A/B/quase-experimentos; aplicação de métodos de redução de variância (como CUPED e pós-estratificação) e dimensionamento de poder estatístico, maximizando a sensibilidade dos testes e reduzindo custos e tempo de exposição.
- **Rigor Metodológico e Diagnósticos Econométricos:** Tratamento e remediação de violações severas de premissas em dados observacionais e experimentais — incluindo diagnóstico e correção de multicolinearidade (decomposição, regularização), heterocedasticidade (erros-padrão robustos/sanduiche, transformações) e tratamento de dados censurados e truncados para neutralizar vieses de seleção e garantir estimativas consistentes.
- **Engenharia Analítica e Aplicações Interativas:** Desenvolvimento de pipelines de dados modulares em Python, R e SQL avançado, além de interfaces analíticas interativas (Streamlit), traduzindo modelos probabilísticos e incerteza analítica em simuladores intuitivos para tomadores de decisão.

Estagiária de Análise Quantitativa e Dados | Chameleon Tecnologia | Remoto | jun. 2022 – fev. 2023

- **Pipelines de processamento:** Limpeza e auditoria de consistência estatística de microdados em Python e R, reduzindo tempo de preparação de bases complexas.
- **Confecção de relatórios:** Auxílio na análise e compilação para entrega ao cliente final nos formatos de apresentação de slide e documento escrito.

Analista de Suporte e Infraestrutura de TI | ISATEC | Curitiba, Brasil | mar. 2013 – fev. 2018

- Suporte a usuários, gestão de ambientes de infraestrutura de TI, redes locais, servidores e resolução de incidentes.

ÁREAS DE DOMÍNIO

- **Experimentação & Decisão Causal:** Desenho experimental ótimo (D-Optimal), testes A/B/n, testes sequenciais, redução de variância (CUPEd), Inferência Bayesiana (ROPE/HDI), Controle Sintético Generalizado, Diferenças em Diferenças (DiD), Propensity Score Matching (PSM).
- **Rigor Econométrico & Controle de Viés:** Diagnósticos de premissas e erros não-esféricos, matrizes de covariância sanduíche robustas à heterocedasticidade (HC0–HC3, Newey-West), diagnóstico de multicolinearidade (VIF, colapso via bootstrap), tratamento de dados censurados e truncados (Modelos Tobit, Heckman Two-Step), reponderação amostral (*Raking* / IPF)
- **Produto, Demanda & Pricing:** Análise Conjunta (CBC), MaxDiff, Logit Multinomial com Classes Latentes (LC-MNL), decomposição de utilidade (*Part-Worths*), matrizes de elasticidade cruzada e otimização numérica restrita.
- **Marketing Science & Atribuição:** Marketing Mix Modeling (MMM), decomposição de Shapley, transformações de Adstock (Weibull/Geométrico), curvas de saturação e rendimentos decrescentes, diagnóstico de multicolinearidade (VIF/BKW) e Análise de Correspondência Múltipla (MCA).
- **Experiência do Cliente & LTV:** Importância Derivada vs. Declarada, Importance-Performance Analysis (IPA), Penalty-Reward Analysis, modelagem longitudinal, análise de sobrevivência (Kaplan-Meier, Cox, Modelos AFT) e mitigação de churn.
- **Séries Temporais & Previsão:** Modelos em Espaço de Estados (SSM), Filtro de Kalman (RTS), Componentes Não Observados (UCM), tendência amortecida, sazonalidade trigonométrica, família ARIMA/SARIMAX, modelos com troca de regime (TAR, STAR) e processos estocásticos (Markov, filas).
- **Machine Learning & Explicabilidade:** Algoritmos baseados em árvores (XGBoost, Random Forest), calibração de probabilidades (Platt Scaling, Brier Score), threshold tuning para custos assimétricos e Explainable AI (SHAP, PDP).
- **Processamento de Linguagem Natural (PLN):** Mineração de texto, embeddings vetoriais, busca semântica, Named Entity Recognition (NER), modelagem de tópicos e extração estruturada de dados.

STACK ANALÍTICO & FERRAMENTAS

Linguagens & Computação Científica: Python (Pandas, NumPy, SciPy, Scikit-learn, Statsmodels, PyMC, Streamlit, Plotly, spaCy), R (Tidyverse, ggplot2, lme4, survival, forecast, fable).

Engenharia de Dados & Nuvem: SQL avançado (CTEs, Window Functions, otimização de queries), Snowflake, Google BigQuery, bancos NoSQL (JSON) e Git.

Desenvolvimento & Produtividade com IA: VS Code, Cursor, GitHub Copilot, esteiras de desenvolvimento assistidas por LLMs (aceleração de código, refatoração e testes automatizados).

Visualização & Comunicação Executiva: Power BI (modelagem Star Schema, DAX avançado), Tableau, Storytelling Quantitativo para liderança técnica e de negócio.

FORMAÇÃO ACADÊMICA

Pós-graduação em Processamento de Linguagem Natural

Universidade Federal de Goiás (UFG) | 2025 – 2026

Ênfase em modelagem de dados não estruturados, arquiteturas neurais e fluxos de recuperação de informação.

Bacharelado em Estatística

Universidade Federal do Paraná (UFPR) | 2018 – 2023

Formação em uma das principais universidades públicas de pesquisa do Brasil. Bacharelado de 5 anos.

Currículo denso focado em Fundamentos de Probabilidade, Matemática, Teoria da Medida, Inferência Clássica e Bayesiana e Desenho Experimental Avançado. Pesquisa aplicada em séries temporais não-lineares com troca de regime, combinando modelagem econométrica e métodos computacionais.

Engenharia de Computação (Créditos Concluídos)

Universidade Tecnológica Federal do Paraná (UTFPR) | 2013 – 2017

Estruturas de dados avançadas, complexidade algorítmica, arquitetura de sistemas e computação científica.

Técnico em Eletrônica Industrial

Universidade Tecnológica Federal do Paraná (UTFPR) | 2008 – 2012

Instrumentação de medição de precisão, lógica de controle e ensaios empíricos.