

# SILVIO DA ROSA PAULA

**Telefone:** (53)984654546

**E-mail:** [silvio.rosa.paula@gmail.com](mailto:silvio.rosa.paula@gmail.com)

**Data Nascimento:** 31/01/1984

**Linkedin:** [www.linkedin.com/in/silvio-paula](http://www.linkedin.com/in/silvio-paula)

**GitHub:** <https://silviopaula.github.io/>



## SOBRE

Tenho atuado na aplicação de Economia e Ciência de Dados em projetos de mercado e pesquisas acadêmicas, com foco em Machine Learning, econometria e inferência causal para apoiar decisões orientadas por dados. Minha formação inclui graduação, mestrado e doutorado em Economia Aplicada, além de uma especialização em IA e Machine Learning.

## HISTÓRIO PROFISSIONAL

### Especialista em Regulação e Ciência de Dados ( Abr 2023 - Presente )

#### Grupo Equatorial Energia - Brasília / DF

- **Implementação e consolidação do processo de projeções internas.**

**Resultados:** desenvolvimento do fluxo de projeções internas de +150 séries mensais, por trimestre, abrangendo sete distribuidoras em diferentes estados. A iniciativa reduziu significativamente a dependência de consultorias externas, gerando economia e maior autonomia analítica, com margem de erro estabelecida e cumprida em até 2%.

**Tecnologias:** R, Python, SQL, caret, Scikit-learn, Statsmodels, Sktime, forecast, modeltime, Metrics, XGBoost, prophet, ARIMA, AutoKeras.

- **Estruturação e padronização de repositórios de dados e projeções.**

**Resultados:** Organização do acervo histórico, governança de dados e rastreabilidade dos dados utilizados em decisões estratégicas.

**Tecnologias:** R, Python, SQLite, ETL.

- **Desenvolvimento de relatórios interativos de monitoramento de desvios.**

**Resultados:** Relatórios HTML mensais para acompanhamento do realizado e dos desvios das projeções.

**Tecnologias:** R Markdown, Jupyter Notebooks, plotly, Shiny.

- **Desenvolvimento de aplicativos internos para análise, visualização e projeção.**

**Resultados:** Conjunto de ferramentas para limpeza, padronizações de dados, modelagem estatística, projeções, visualização de dashboards.

**Tecnologias:** R e Shiny, Python e Streamlit, SQL, caret, Scikit-learn, Statsmodels, Sktime, forecast, pandas, dplyr.

- **Diagnóstico e melhoria de séries com alto erro de previsão (MAPE).**

**Resultados:** Aplicação de modelos ensemble models, combinando séries temporais, machine learning e redes neurais.

**Tecnologias:** R, Python, caret, caretEnsemble, PyCaret, tuning de hiperparâmetros.

- **Estudos especiais com inferência causal e econometria aplicada**

**Resultados:** Melhor compreensão dos efeitos como pandemia, clima, renda sobre o consumo de energia

**Tecnologias:** R, Python, causalimpact, fixest.

### Pesquisador e Econometrista ( Jan 2021 - Abr 2023 )

#### Instituto de Pesquisa Econômica Aplicada - IPEA – Brasília / DF

- **Execução de rotinas de ETL e cruzamento de grandes bases de dados governamentais.**

**Resultados:** Extração e manipulação de dados restritos de múltiplas fontes em estrutura única e padronizada, viabilizando análises robustas em projetos de pesquisa aplicada.

**Tecnologias:** Big Data, Spark, R, Python, SQL, data.table, arrow, sparklyr, ff, dask, pyarrow.

- **Criação de Banco de dados abrangente sobre programas sociais relacionados a agricultura familiar.**

**Resultados:** Consolidação de dados transformada em produto analítico permanente, ampliando a capacidade de análise integrada dos programas (PAA, PNAE, entre outros) e servindo como fonte contínua para estudos internos da instituição.

**Tecnologias:** R, SQLite, arrow, tidyverse.

- **Pesquisador e econometrista em estudos de avaliação de impacto causal**

**Resultados:** além de contribuir com o texto dos artigos e ser o único responsável pela parte econométrica do estudo o qual fui contratado, fui convidado a conduzir análises de outros programas, resultando em dois artigos científicos, dois textos para discussão, capítulo de livro, apresentação em 2 congressos nacionais de economia.

**Tecnologias:** R, Stata, did, fixest, fixest::sunab, MatchIt, Matching, WeightIt, diff, csdid, psmatch2.

**Pesquisador assistente ( Jun 2021 – fev 2023 )**  
**Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa (IDP)**

- **Produção técnica no LAIPP (Laboratório de Avaliação e Inovação em Políticas Públicas)**  
*Atividades:* Elaboração de textos técnicos baseados em estudos científicos de impacto internacional, traduzindo evidências acadêmicas, especialmente conceitos de estatística e econometria, em linguagem acessível para gestores públicos.  
*Link:* [LAIPP - IDP](#)
- **Monitoria em Lógica de Programação com Python**  
*Atividades:* Apoio a discentes de graduação na disciplina de introdução à programação aplicada à análise de dados.  
*Tecnologias:* Python.
- **Pesquisador assistente em projetos de pesquisa de pós-graduação**  
*Atividades:* Apoio a docentes e discentes de pós-graduação na redação de artigos científicos e coorientação de dissertações.  
*Tecnologias:* R, Stata.

**Consultor Técnico em Dados e Economia Aplicada ( Set 2021 - Jan 2023 )**  
**Deltainfra - Brasília / DF**

- **Consultoria técnica sob demanda em análise de dados e modelagem econométrica.**  
*Resultados:* Apoio em avaliação de RFPs e projetos no setor de energia, com foco em dados e viabilidade quantitativa.  
*Tecnologias:* R, Excel

**FORMAÇÃO ACADÊMICA**

- **Doutorado em Economia Aplicada ( Jan 2019 - Dez 2022 )**  
*Universidade Federal de Pelotas - Pelotas - RS*  
*Status:* **Concluído**
- **Especialização em Machine Learning e Inteligência Artificial ( Jul 2021 - Mar 2022 )**  
*Anhanguera Educacional - EAD*  
*Status:* **Concluído**
- **Mestrado em Economia Aplicada ( Jan 2017 - Dez 2019 )**  
*Universidade Federal de Pelotas - Pelotas, RS*  
*Status:* **Concluído**
- **Bacharelado em Ciências Econômicas ( Jan 2011 - Dez 2016 )**  
*Universidade Federal de Pelotas - Pelotas, RS*  
*Status:* **Concluído**

**FORMAÇÃO COMPLEMENTAR**

Cursos	Conhecimentos em Estatística
<ul style="list-style-type: none"><li>• SQL</li><li>• Power Bi</li><li>• Git e GitHub</li><li>• Redes Neurais em R</li><li>• Formação Cientista de Dados Python e R</li><li>• LGPD</li><li>• Scala and Spark for Big Data</li><li>• AutoML</li><li>• Series temporais com redes neurais</li><li>• Cloud Computing - AWS, Google Cloud e Azure - Dados na nuvem</li><li>• Linguagem Julia</li></ul>	<ul style="list-style-type: none"><li>• <b>Econometria:</b> Regressão Linear e Não Linear; Séries Temporais, Pannel De Dados, Econometria Espacial, Inferência Causal.</li><li>• <b>Data Science:</b> Machine Learning e Deep Learning, Redes Neurais, Classificação, Clusterização e NLP.</li><li>• <b>Inferência Causal:</b> Instrumental Variables (IV), Difference-in-Differences (DiD), Regression Discontinuity (RD), Propensity Score Matching (PSM), Synthetic Control, Causal ML e A/B Testing.</li><li>• <b>Geoestatística:</b> Interpolação Espacial IDW, Krigagem, Shapefiles, QGIS, GEODA.</li><li>• <b>Estatística:</b> Amostragem, Análise Descritiva, Modelagem Preditiva, Intervalos de Confiança, Probabilidade, Inferência Bayesiana, testes não Paramétricos, Inferência Estatística e Testes de Hipóteses.</li></ul>
<ul style="list-style-type: none"><li>• <b>Idiomas:</b> Português: Nativo; Inglês: Leitura e escrita; Espanhol: Leitura e escrita</li><li>• <b>Áreas de conhecimento:</b> Setor Elétrico, Seguro Rural, Crédito Rural, Avaliação de Programas Sociais, Inovação, Educação.</li><li>• <b>Dados governamentais:</b> Manipulação de Microdados Governamentais como PNADc, Censo, Caged, Rais etc.</li><li>• <b>Utilização de LLMs:</b> ChatGPT, Gemini, Claude, Grok, Deepseek, Qwen, Loveble, LLMs locais, cursor Ai, Windsurf.</li></ul>	