Philipp Kleer

he/him | Recife, PE | +55 81 99406-3729 | philipp.kleer@posteo.com | kleer | philkleer

ABOUT ME

Ph.D.-trained Data Scientist with **8+ years of experience** in **quantitative research**, statistical modeling, and **data-driven decision support**. Strong background in building **reproducible analytical workflows**, developing **production-grade data applications**, and translating complex analyses into **actionable insights for diverse stakeholders**. Experienced in **international and interdisciplinary environments**.

WORK EXPERIENCE

Data Scientist
July 2025 — Present

Center for Studies and Research in Network Technology and Operations - Measurements (CEPTRO), NIC.br — *São Paulo, Brazil*

- Built and deployed **production-grade data applications** in **R/Shiny**, refactoring large monolithic codebases into a modular architecture, reducing LOC and improving performance and maintainability.
- Designed and maintained **data pipelines, databases, and APIs** (PostgreSQL), improving **data quality, structure, and reproducibility** across multiple analytical portals.
- Developed **interactive dashboards and data visualizations** (ggplot2, plotnine, ggiraph, plotly) with standardized design and narrative reporting for technical and policy audiences.
- Implemented **CI/CD pipelines** (GitLab) with automated testing (testthat), linting, pre-commit hooks, and **reproducible environments** (renv, uv) for internal R packages and applications.
- **Containerized and deployed applications** using **Docker** and **Kubernetes**, supporting automated builds, releases, and production environments.
- Applied **statistical modeling and Bayesian regression** to analyze network quality and other large-scale datasets, supporting exploratory and decision-oriented analyses.
- Contributed to **machine learning and data innovation initiatives**, including early work on **semantic segmentation** and experimentation with **LLMs** in development and testing workflows.

Lecturer
April 2025 — July 2025

Justus-Liebig-University Giessen — *Giessen, Germany*

- Authored **3 publications/chapters**, including **applied material on teaching R** to social science students.
- **Designed and taught 4 courses** on political socialization and teachers' role (~ 200 students)

Data Science & MLOps (Training)
October 2024 — March 2025

DataScientest (certified by Panthéon-Sorbonne, in total 505 hours) — *Remote*

- Completed a **505-hour certified Data Science & MLOps program** (Panthéon-Sorbonne).
- Built **machine learning models** for classification, regression, and clustering using scikit-learn, TensorFlow, and PyTorch.
- **Developed** and **tracked ML pipelines** with **MLflow, DVC, Dagshub, Airflow**, and **ZenML**.
- Deployed models using BentoML and gained **hands-on experience with Docker**, **Kubernetes**, **Prometheus**, and **Grafana**.
- Created **analytical visualizations** and **interactive apps** with matplotlib, seaborn, plotly, and Streamlit.

Postdoctoral fellow
February 2023 — September 2024

Justus-Liebig-University Giessen — *Giessen, Germany*

- Led data-driven research projects using **R** and **Python**, **delivering statistically robust insights** for academic and policy audiences.
- Applied advanced statistical and machine-learning methods (**post-double selection**, EFA/CFA, SEM, multilevel models, regression) with **strong emphasis on data quality and validation**.
- Produced **8 publications** and technical reports, **translating complex analyses into actionable insights**.
- Collaborated in **international research teams** and conducted research stays in Brazil and Spain.

Doctoral fellow
October 2015 — September 2022

Justus-Liebig-University Giessen — *Giessen, Germany*

- Conducted **large-scale quantitative research** using **R**, **Git**, and **reproducible workflows**.
- Applied **advanced statistical modeling techniques**, including multilevel models, SEM, Bayesian regression, and quantitative text analysis.
- Published **6 peer-reviewed papers/reports** and presented findings at **9 international conferences**.
- Designed and **taught research methods courses** and supervised 11 thesis projects.
- Received **advanced training in machine learning with R**, **Bayesian modeling**, **multilevel regression**, and **quantitative text analysis**.

Project assistant (workshop management)
May 2019 — May 2022

Justus-Liebig University Giessen — *Giessen, Germany*

- Designed and maintained a **SQL-based database** to **automate reporting** and **certification workflows**, significantly reducing manual effort.
- Coordinated large academic events and supervised student assistants, ensuring on-time delivery and quality standards.

Skills

Data Science & Analytics: Statistical Modeling | Bayesian Regression | Machine Learning | Exploratory Data Analysis | Data Quality & Validation | Quantitative Research | Feature Engineering |

Programming & Data: R (`tidyverse`, `Shiny`, `ggplot2`) | Python (`scikit-learn`, `PyTorch`, `TensorFlow`) | SQL (`PostgreSQL`) |

Data Visualization & Reporting: Shiny Dashboards | Interactive Visualization | `ggplot2` | `plotnine` | `ggiraph` | `plotly` | Quarto |

MLOps & Engineering: CI/CD (GitLab) | Docker | Kubernetes | MLflow | Airflow | `renv` | `uv` |

Version Control & Collaboration: Git | GitHub | GitLab | Code Review | Reproducible Workflows |

Languages: English (fluent, C1) | Portuguese (fluent, B2/C1) | German (native, C2)

Projects

- **Shiny application Observatório de Inteligência Artificial (OBIA):** Refactored a large production Shiny application into a modular, testable architecture, reducing code size by 41%. Implemented CI/CD, automated testing, linting, and reproducible environments (`renv`) to support reliable multi-developer collaboration.
- **Shiny application Autodiagnóstico do Sistema de Administração dos Recursos de Tecnologia da Informação:** Developed and deployed a production-grade Shiny dashboard assessing IT governance among national entities, including data integration, standardized visual design, and CI/CD-based deployment for public-facing analytics.
- **Network Technology Analysis & Presentation:** Conducted statistical and exploratory analyses of network technologies, translating results into high-impact presentation for technical and policy stakeholders at IX Forum 2025.
- **PhD project:** Applied Bayesian hierarchical models (`brms`, `RStan`) to analyze large-scale regional data, producing reproducible statistical analyses and peer-reviewed publications.
- **ML Ops Pipeline Project:** Implemented an end-to-end MLOps workflow for model training, versioning, and monitoring using MLflow and DVC, simulating production-like model lifecycle management.

Education

Doctorate (Dr. rer. soc.)   May 2016 — December 2022
Justus-Liebig University Giessen   *Giessen, Germany*
- degree in Social Scienes with specialization on quantitative methods; doctorate was accomplished at the Chair of Empirical Research Methods

Master of Arts   October 2011 — May 2014
University of Vienna   *Vienna, Austria*
- degree in Political Science with a focus on quantitative methods and Central Eastern European societies

Bachelor of Science   October 2008 — August 2011
University of Cologne   *Cologne, Germany*
- degree in Social Sciences with a strong emphasis on empirical research methods