

Krystian Korzec

Full-Stack Data Scientist | Causal Inference & Bayesian Modeling

Contact Information

- **Location:** Warsaw, Poland
 - **Phone:** +48 730 555 510
 - **LinkedIn:** [linkedin.com/in/krystian-korzec](https://www.linkedin.com/in/krystian-korzec)
 - **Email:** k.korzecds@gmail.com
 - **GitHub:** github.com/krystiankorzec
-

Professional Summary

Quantitative Full-Stack Data Scientist with over 8 years of experience bridging rigorous statistical methodology with production-grade software engineering. Specialized in Causal Inference, Bayesian Statistics, and mathematical optimization. Proven track record of developing end-to-end analytical solutions—from estimating heterogeneous treatment effects (CATE) to architecting containerized, production-ready web applications (PyShiny) and scalable data pipelines. Adept at managing the full lifecycle of data products using a modern cloud and database stack.

Key Competencies & Tech Stack

- **Causal Inference & Experimentation:** Heterogeneous Treatment Effect (CATE) estimation, Double Machine Learning, A/B Testing, Synthetic Controls, Geo-experimentation, Uplift Modeling.
 - **Statistical & Optimization Methods:** Hierarchical Bayesian Modeling, MCMC simulation, Mixed-Integer Linear Programming, Stochastic Optimization (SAA).
 - **Programming & Libraries:** Python (Pandas, Polars, Scikit-learn, PyMC, Pyomo, PyShiny), R, Advanced SQL, PySpark.
 - **Cloud, Infrastructure & Databases:** GCP (BigQuery), AWS, Snowflake, MongoDB, Redis, Docker, CI/CD, Git, Airflow.
-

Employment History

Data Scientist | Allegro

2024 Jul – Present | Warsaw, Poland

- Formulated and deployed hierarchical Bayesian predictive models and Sample Average Approximation workflows (Project Charon) for real-time contact allocation under operational uncertainty.
- Co-architected the methodology for geographic experimentation frameworks (geo-testing/synthetic controls) to robustly measure the causal impact of pricing interventions.
- Developed causal pricing models to estimate price elasticity (CATE) and optimize product selections, moving beyond pure prediction to direct counterfactual policy optimization.
- **Tech Stack:** Python, PyMC, Pyomo, GCP, BigQuery, GitHub, Airflow.

Contract ML Engineer (Part-Time / B2B) | Roche (via Seargin)

2024 Aug – 2026 Apr | Remote

- Designed and deployed a full-stack, AI-driven platform (PyShiny web app) to automate the generation of complex non-clinical statistical reports, significantly reducing manual research overhead.
- Architected the application infrastructure using containerized deployments (Docker) integrated with robust CI/CD pipelines.
- Implemented state management using Redis and established comprehensive testing to ensure the high performance and reliability of the analytical tools.
- **Tech Stack:** Python, R, PyShiny, Docker, GitLab, Redis, Gemini API.

Data Scientist | Allianz

2023 Jul – 2024 Jul | Warsaw, Poland

- Developed and maintained data-intensive statistical applications and analytical pipelines using Python (PySpark, Pandas).
- Utilized Amazon EMR Studio to perform large-scale data processing tasks for actuarial and analytical workflows.
- **Tech Stack:** Python, AWS, GitHub, SQL.

Data Science Analyst | Accenture

2022 Nov – 2023 Jul | Warsaw, Poland

- Delivered analytical solutions solving complex allocation problems using Integer Programming (Pyomo) and Supply Chain Optimization techniques.
- Spearheaded the refactoring of legacy statistical R code into scalable PySpark/R solutions in Azure Databricks, drastically reducing execution time and enforcing engineering rigor.
- **Tech Stack:** Python, GCP, Databricks, GitHub, PySpark.

Data Scientist | Nowa Era

2021 Jan – 2022 Sep | Warsaw, Poland

- Designed and deployed causal uplift models (Heterogeneous Treatment Effects) for targeted retention campaigns, successfully isolating the true incremental impact of marketing interventions from organic behavior.
- Built predictive models for early detection of at-risk users, transitioning the analytical focus towards proactive intervention strategies.
- Architected analytical infrastructure and end-to-end data pipelines leveraging AWS and Snowflake.
- **Tech Stack:** Python, R, Snowflake, GitHub, Docker, AWS.

Data Scientist | PKO BP

2018 Jul – 2021 Jan | Warsaw, Poland (*Promoted to Data Scientist following a 6-month Data Science Internship*)

- Utilized spatial analytics and statistical segmentation to build Automated Valuation Models (AVM) for real estate pricing prediction.
 - Engineered automated monitoring pipelines to track model stability and detect geographic feature drift.
 - **Tech Stack:** Python, Hadoop, SQL.
-

Education

Warsaw School of Economics (SGH)

Master's Degree in Quantitative Methods in Economics and Information Systems (MIESI) - Master Thesis:
Spatio-Temporal Bayesian Model of PM2.5 and PM10 Air Pollution

University of Warsaw (MIMUW)

Completed 2 Years of B.Sc. Coursework in Pure Mathematics - Built a rigorous theoretical foundation in probability theory, real analysis, and linear algebra supporting advanced statistical modeling.