KONSTANTIN HOFFIE

MSc Statistics

khoffie myemail@earthmail.com Europe, Earth

ABOUT ME

I am a passionate researcher and data scientist specializing in mathematical modeling and Bayesian methods to solve real-world problems. Eager to learn and grow, I plan to finish my PhD thesis in February. I seek to transition from research to industry to make a tangible impact on society.

SKILLS

Skill Set	Skills
Programming languages	R (advanced), Julia, Python
Coding	Modular and clean code, familiar with unit testing, collaborated via GitHub
Data Science	Advanced skills in data wrangling, visualization, and modeling; analyzed large datasets (17M rows), primarily in R
Machine Learning & AI	Conceptual knowledge of deep learning, especially Convolutional Neural Networks
Statistics & Causal Inference	(Generalized) linear models, survey and register data analysis, experience in selecting data; long-time reader and occasional commenter on Statistics blogs
Research & Collaboration	Clear writing, comfortable engaging with non-technical audiences

WORK EXPERIENCE

Research Assistant (fulltime) — **Earth University**, Earth

</> R | Julia | Turing.jl | Parallelization | spatial data — Sept 2021 – Dez 2024

- I implemented and extended the code for a mathematical model to analyze migration patterns between German districts. While the conceptual framework was developed collaboratively with scientist, I managed data preparation, integrated the prototype into a flexible pipeline, created visualizations for clear result communication and conceived the mathematical notation.
- The nonlinear Bayesian model, implemented in `Julia` with `Turing.jl`, overcomes key limitations of standard approaches, providing unbiased predictions of migration flows and revealing latent preferences for urban and rural areas. The code is available here: .

Internship Data Scientist — **Earth GmbH**, Europe

</> R | python | keras | shiny — April 2019 - March 2020

- I built an image classification tool using `Python` and `Keras` to find out the race of customer's dogs. Insurance premium was chosen based on the predicted race. This tool is still in use.
- I built a dashbord using `R` and `Shiny` for management and employees to see how the company is doing.

EDUCATION

PhD Data Science (forthcoming) — **University of Earth**, Earth

</> R | Julia | Turing.jl | Parallelization | spatial data — Sept 2021 - March 2025

- Thesis: A mathematical model of internal migration (migrations on earth). For details see section "Work Experience" above. I hand in my thesis in February.
- Publication: Forthcoming, I hand in the paper in March.

Msc Statistics (1.6) — **Earth University**, Earth

</> python | tensorflow — Okt 2016 - July 2020

- Thesis: Deep Learning for Image Manipulation Detection (In cooperation with Earth GmbH)
  I replicated a deep learning paper which applies the Faster R-CNN architecture to detect image manipulation. I successfully ran their model in `python` and `tensorflow`. This work included debugging their code, augmenting data, fine-tuning a pretrained network and exploring the architecture of CNNs.

BA Statistics (1.6) — **University of Earth**, Earth

</> R | stan — Okt 2012 - Sept 2016

- Thesis: An empirical evaluation of the randomized response technique. I developed a simple model in `R` and `stan` to analyze response rates to sensitive survey questions.

SIDE PROJECTS

Make migration model fully reproducible — **Reproducible Research**,

</> R | Julia | nix | git — khoffie/earth-matters

- Reproducibility matters, to have confidence in one's results and so that others can use and extent the model.