

# Helge Liebert

DATA SCIENTIST · QUANTITATIVE ANALYST

Bristenstrasse 12, 8048 Zürich, Switzerland

☎ +41 78 606 27 63 | ✉ helge.liebert@gmail.com | 🌐 hliebert.github.io  
🏠 hliebert | 📺 hliebert | 📅 23.01.1986 | 🇨🇭 Swiss, German



*“12 years of experience with statistical modeling of real world data. Experience developing data-driven applications end-to-end in python—from PoC to user-facing application. I’m lost without vim keybindings.”*

## Experience

---

### Senior Data Scientist

Zurich, CH

SWISS NATIONAL BANK

Since June 2022

- Definition of tasks for the team in agile work environment. Code review/mentoring of junior team members.
- Responsible for short-term inflation forecast, reporting to senior management.
- Introduced a new forecasting model for headline inflation (xgboost applied to time series, mlflow).
- User-facing applications: Inflation monitor; news monitor and media sentiment dashboard. (ETL pipeline using Apache Spark/HDFS on Cloudera, dash, automated reports using quarto).
- Tooling for recurring computations, packaged and published on internal PyPI mirror. (20k+ LoC, testing using pytest with 50% coverage, linting, type checks).
- Modular Gitlab CI/CD pipeline and project templates used by multiple teams (pre-commit, black, ruff, pytest).

### Senior Researcher

Zurich, CH

UNIVERSITY OF ZURICH

Nov. 2020 – May 2022

- Uplift modeling for optimal insurance assignment (generalized random forests, integer programming).
- Project lead on a worldwide survey (130k participants, budget 1.3M CHF, led three full-time RAs).
- Implemented ETL pipeline on UZH Science Cloud.
- Teaching: Methods for Unstructured Data, Computational Data Analytics for Economists, Text Analysis.
- Lab sessions served with JupyterHub/Repo2Docker using K8s on Google Cloud.

### Data Scientist

Boston, USA

HARVARD MEDICAL SCHOOL

Mar. 2019 – Nov. 2020

- Project lead “Predicting Work Capacity in the Social Security Disability Insurance Program” (led team of three).
- Mined work history event data for 3.8M disability insurance applicants.
- Predicted residual earnings capacity based on work and medical history.
- PoC deployed on Harvard on-prem infrastructure, A/B testing of project success.

### Data Scientist

Zurich, CH

SANITAS, ADVANCED ANALYTICS

May 2018 – Feb. 2019

- Analyzed reimbursement outside of the provisions of statutory health insurance.
- Built a model to detect wrongful reimbursement.

### Doctoral Researcher

St. Gallen, CH

UNIVERSITY OF ST. GALLEN

Feb. 2012 – Feb. 2018

- Data-driven research projects (non-parametric methods, causal ML). Multiple A-ranked publications.
- Teaching: Econometrics. Administration of internal IT infrastructure.

# Education

---

## Ph.D. in Econometrics

Feb. 2012 - Sep. 2017

UNIVERSITY OF ST. GALLEN

St. Gallen, CH

- Specialization track: Econometrics. Thesis grade: 6.0/6.0. Best dissertation award 2017.
- Selected coursework: Computational Statistics, Database Systems, Econometrics of Big Data, Empirical Strategies for Causal Inference, Machine Learning and Causal Inference.
- Research visit: University of Uppsala, Department of Statistics.

# Skills

---

Programming	Python, R, SQL, Bash
ML toolkit	functime, lightbgm, matplotlib, mlflow, numpy, pandas, pyspark, pytorch, scikit-learn, sktime, xgboost
SW dev. in Python	dash, packaging, project templating, pytest, pyCharm
Other tools	Apache Spark, Cloudera, Docker, GIS Git, Gitlab, Shell, Vim, VSCode
Operating systems	Linux, MacOS, Windows
Soft skills	Project management, working in multi-project and multi-culture environments
Languages	German (Native), English (fluent), French, Spanish (basics)

# Interests

---

Climbing, Bouldering	Mostly outside on weekends, structured training during the week. Level: 7b os/8a rp. I enjoy the combination of physical and mental challenges in climbing.
Cycling	Cycle travel with friends in Europe, my bike is my daily companion when commuting. I enjoy tinkering with and repairing older bicycles.
Open source	I am a long-time Linux user with contributions to open source libraries (e.g. sktime, a scikit-learn-compliant library for time series analysis and forecasting).

OCTOBER 10, 2024 · ZURICH

  
HELGE LIEBERT