



BENJAMIN ROCHETEAU

Irvine, California | +1 (949) 527-8240 | benjaminphrocheteau@gmail.com

Work Eligibility: French Citizen & US Green Card

PROFESSIONAL

Banque of France – Analytic Modeling Team, France | Data Science Intern 07/2024 – 12/2024

Causal Inference (Regression, Fixed Effects), Machine Learning (Random Forest, XGBoost), R, Python, LaTeX, Web Scraping

- Developed a forecasting model of bank response to short term interest rate shocks. Compared accuracy across both econometric (linear regression, fixed effect) and machine learning (XGBoost, Random Forest) models across 4 different metrics to inform policymaker about financial stability.
- Collected and cleaned thousands of data points through web scraping using an API for financial data in R. Used the data to feed into a modeling pipeline for forecasting and projections.
- Summarized work into reports and presentations to communicate methods and results to team and stakeholders across technical and non-technical background.

Moody's Analytics – Predictive Analytics Team, Czech Republic | Analytics Intern 06/2023 – 08/2023

Causal Inference (Regression), Machine Learning (GAM, Random Forest, XGBoost), R, AWS EC2 and VMs, GitHub, Web Scraping, ETL

- Updated the team's modeling dataset for an annual model refresh using ETL and aggregating across multiple internal sources, scraping external sources, and validating the collected data.
- Transitioned team codes to GitHub and AWS EC2 from a shared drive in order to streamline the workflow. Wrote SOPs and hosted training sessions in order to train the team on the use of the new code-sharing and data-storage solutions.
- Wrote clear project documentation to streamline the onboarding of future hires. Created vignettes for each function, dependency charts, and how-to guides for setting up R, Github, AWS, and model deployment.

GeoSys – Data Science Team, France | Statistical Consultant 09/2023 – 03/2024

Machine Learning (XGBoost), Panel Data, Time Series, Functional Data Analysis, Big Data, Sample Design, Python, Google Collab

- Developed a crop-identification model to showcase functional data analysis (FDA) techniques to the team. Inspired by agricultural literature, we calculated the normalized distribution vegetation index (NDVI) using red light and near-infrared light band data and transformed time into growing degree days, which in conjunction with the FDA techniques improved model accuracy by over 20%.
- Put in place a big data workflow to be able to train models on the full dataset, which was over 10 times bigger than our computers' RAM. Implemented iterators to be able to train our models, and used splines for dimensionality reduction in order to avoid issues like overfitting.
- Implemented James Bethel's stratification optimization algorithm in order to calculate the optimal strata sizes for model performance, and combat the issue of drastically unbalanced class distributions. The algorithm choses strata sizes as a function of class and feature variance.

Toulouse School of Economics – Macroeconomics 3, France | Teaching Assistant 09/2023 – 12/2023

AD-AS, ISLM, Mundell-Fleming, Communication, Teaching

- Led weekly TA sessions to present the solutions to the problem sets and graded student exams

University of California, Irvine – History of the Fed, USA | Teaching Assistant 06/2020 – 08/2020

Federal Reserve System, Communication, Teaching

- Wrote lecture and reading summaries, and created weakly revision quizzes for the students

EDUCATION

Toulouse School of Economics, France

- M.S in Data Science (2024)
- B.S. in Economics and Mathematics (2022)