Nicolo Foppa Pedretti

+1(917)-294-6771 | nicolo.fp@gmail.com | GitHub: nicolofp | LinkedIn

Quantitative Data Scientist with 8 years spanning research, energy, and finance sectors .
Specialized in statistical modeling, machine learning, and Bayesian methods with proven publication track record.

Experience

Critical Path Institute   Sep 2024 – Current
Quantitative Medicine Scientist (Data Scientist) – Pharmacometrics   *Remote, US*

- Developed disease progression models using linear/non-linear mixed effects frameworks in R and Stan, applying frequentist and Bayesian approaches to analyze longitudinal Alzheimer's data and identify predictive biomarkers
- Engineered automated ETL pipelines in R to transform clinical trial data from SDTM/ADaM standards into analysis-ready datamarts with comprehensive data quality checks and validation procedures

Icahn School of Medicine at Mount Sinai   Feb 2019 – Apr 2024
Biostatistician II (Data Scientist) – Environmental Medicine and Climate Science   *New York, NY*

- Analyzed multi-omics datasets (metabolomics, epigenetics) using machine learning, dimensionality reduction (PCA, t-SNE, UMAP), and linear mixed-effects models for biomarker identification and pathway validation in R
- Developed Bayesian mixture regression R package using Stan and applied advanced time series models (Distributed Lag Models, ARIMA, Bayesian splines, Gaussian State Space Models) for environmental health analysis using Stan, JAGS, and R
- Implemented machine learning pipelines (XGBoost, tree-based models, polynomial regression) for biomarker prediction and validation on large-scale genomic data using R and Julia
- Built automated ETL pipelines and reproducible workflows using SQL, R, and DuckDB, generating publication-ready reports with RMarkdown/Quarto, ggplot2, and plotly for peer-reviewed publications

InchCapital   Apr 2017 – Dec 2018
Quantitative Analyst – R&D   *Milan, Italy*

- Executed time series analysis using ARIMA and Hidden Markov Models with harmonic regression to detect trading patterns and anomalies, designing quantitative strategies across multi-asset portfolios with rigorous back-testing and Monte Carlo simulations using R, Python, and Julia
- Optimized investment portfolios using Markowitz theory and advanced peer-reviewed methodologies to maximize risk-adjusted returns, establishing key performance metrics and automated executive reporting systems using R/RMarkdown for management and client deliverables

Gala S.p.A.   Jun 2016 – Nov 2018
Data Scientist – Market Intelligence and Risk Control   *Milan, Italy*

- Built mathematical and statistical models for portfolio performance monitoring and commodity price forecasting, applying machine learning and data mining techniques to energy time series data to identify profitable trading opportunities using R
- Validated and optimized quantitative trading models while managing ETL processes across PostgreSQL databases, deploying parallel processing solutions on Linux servers and automating workflows through custom RShiny applications using R and SQL

Education

Universitá degli Studi Milano-Bicocca   Sep 2013 – Jan 2016
Master of Science in Statistics   *Milan, Italy*

Universitá degli Studi di Milano   Sep 2009 – Dec 2013
Bachelor of Science in Applied Mathematics   *Milan, Italy*

Skills

Languages: Italian (native), English (professional/advanced)
Programming Languages: R, Stan, Julia, Python, SQL, Bash
Development Frameworks: RShiny, RMarkdown, Quarto, DuckDB