



Data Scientist

- 📍 Geneva, Switzerland
- ✉ david.munoztord@mailbox.org
- 🏠 david-munoztord.com
- 🔗 munozt0
- 📱 david-munoz-tord

Skills

R	9 yrs.
Python	5 yrs.
JavaScript	3 yrs.
Bash	7 yr.
SQL	4 yrs.
Data management	5 yr.
Data Analysis	9 years
Data Engineering	5 yrs.
Data Visualization	7 years

Libraries

- ▶ echarts, ggplot2, plotly, seaborn
- ▶ R Shiny, Dash, Streamlit
- ▶ sf, raster, rgdal
- ▶ TensorFlow, Keras, PyTorch, Scikit-learn
- ▶ Rstan, brms, JAGS
- ▶ H2O, XGBoost, LightGBM
- ▶ tidymodels, caret, mlr

Tools/Frameworks

- ▶ Git, Terminal
- ▶ Docker
- ▶ CI/CD
- ▶ AWS
- ▶ ML OPS
- ▶ ETL, NIFI

Summary

Experienced Data Scientist proficient in data analysis, software development, and data management. Skilled in R, Python, and SQL, with expertise in utilizing advanced data science tools and frameworks. Recognized for contributions to publications and active involvement in the data science community. Ready to contribute expertise in enhancing data-driven decision-making and information management in a challenging environment.

Experience

Data Scientist

10/2022 - Present

Freelancer

- Led the development and launch of a scientific project in collaboration with EPFL, transitioning initial scripts into an accessible software to clients via the cloud, emphasizing innovation and continuous improvement.
- Provided technical consulting to a biotech start-up, translating academic research into actionable solutions and advising on machine learning implementations, showcasing communication and training skills.
- Collaborated with cross-functional teams within an NGO to develop and implement data reporting infrastructure, emphasizing team management and capacity building initiatives, while implementing process automation best practices to enhance efficiency.

Full Stack Data Scientist

04/2022 - 01/2024

FlowBank

- Contributed as a key member of the data migration to AWS cloud, ensuring seamless transfer, scalability, and multitasking abilities.
- Conducted in-depth analysis to classify customers using Bayesian inference, revealing distinct segments and patterns for data-driven decision-making and targeted strategies.
- Engineered and deployed a sophisticated fraud detection system, leveraging neural networks and support vector machines to accurately identify fraudulent activities.
- Developed internal tools for data analysts in R packages, providing training and support for their use, indicating capacity building skills.
- Designed and implemented a datawarehouse solution for efficient data processing and orchestration, contributing to streamlined data engineering workflows.
- Led the development of a comprehensive ML-Ops infrastructure, leveraging AWS cloud infrastructure for seamless deployment and scalability.
- Designed and maintained a data visualization dashboard in R Shiny, facilitating communication of key performance indicators, demonstrating strong reporting skills, and showcasing expertise in data visualization.

Lecturer in Biotatistics

09/2020 - 01/2022

University of Geneva

- Conducted lectures on statistics for the graduate program, mentored students in research projects, and provided guidance on statistical analysis and data visualization, demonstrating training and mentoring skills.
- Created and shared freely available data workshops in R and Python, emphasizing capacity building in statistical analysis and data visualization, showcasing commitment to continuous learning and knowledge sharing.

PhD Candidate in Computational Neuroscience

03/2020 - 01/2022

Campus Biotech

- Processed terabytes of data for a large scale of longitudinal clinical trials (EEG/fMRI signal and blood samples).
- Used various statistical methods to analyze large scale datasets: Hierarchical Bayesian modeling, Multivariate pattern analysis, and model-free reinforcement learning.
- Demonstrated commitment to medical data management protocols compliance and proficiency in grant acquisition.
- Collaborated with Caltech data engineers to implement ICA-based denoiseification for better identification of activation and artifact components in fMRI.

Education

Ph.D. Candidacy in Computational Neuroscience

03/2020 - 01/2022

Campus Biotech - Switzerland

Topic: Computational modeling of addiction

MSc in Neuroscience

09/2018 - 03/2020

University of Geneva - Switzerland

Thesis: Denoising of high resolution fMRI signal using tICA

Diploma in American Studies

09/2017 - 06/2018

Smith College - USA

Capstone: Psychological framing in political discourse on twitter during the U.S. elections

BSc in Cognitive Science

09/2014 - 06/2017

University of Geneva - Switzerland

Thesis: Early detection of ASD event related potentials

Languages

French	Native
Spanish	Native
English	Fluent
Italian	Proficient

Technical Skills

- ▶ Geographic information systems
- ▶ Sampling methodologies
- ▶ Machine Learning
- ▶ Linux System Administration
- ▶ Visualization for Scientific Data
- ▶ Spatial Analysis
- ▶ Bayesian Statistics

Extracurricular

- 🔗 Top 2% on [StackOverflow](#) (2024)
- 🐙 Top 25% on [GitHub](#) (2023)
- 🏆 Global Fellow scholarship (2017)
- Co-founder of [Go-Fast](#), a socially responsible Bike Messenger cooperative in Geneva.

Internship in Computational Neuroscience

03/2019 - 03/2020

Geneva Neurocenter

- Developed and implemented [computational models](#) to study neural system mechanisms and dynamics.

Projects

Volunteer

actif-traffic / Pro-Velo

R, SQL, JavaScript

- Helping to evaluate the implications of reverting to a 30 km/h speed limit through the analysis of public data from the Swiss Federal Statistical Office. The findings aim to offer valuable insights for potential policy considerations and decision-making processes in a measured and informative manner.

Co-President

We Data

R, Python, Rust, JavaScript

- [We Data](#) is an organization that shares knowledge on code in data science (Blog, YouTube channel), give workshops about statistics, and do coding demonstrations. We also organize the [R-Lunches](#).

Maintainer

echarts4r

R, JavaScript

- Collaborated on the development of [echarts4r](#), a R package that expands the capabilities of R-based interactive visualizations with a powerful rendering engine, while enhancing flexibility and ease of use.

Contributor

hBayesDM

STAN, R, Python

- Collaborated on the Hierarchical Bayesian modeling of Decision-Making tasks library ([hBayesDM](#)). Built Q-learning algorithm for probabilistic selection task.

Maintainer

Firebase

R, JavaScript

- Contributed extensively to the [firebase](#) integration package for R, empowering the Shiny user community by enabling user authentication and secure file storage through Firebase Storage.

Creator

DbVieweR

R, SQL

- Created a shiny app ([DbVieweR](#)) simulating a database management system featuring functions like login authentication, save/create/delete tables, add/rename columns, using either a PostgreSQL or SQLite back-end.

Publications

See my [publications pages](#) for a comprehensive list my scientific publications.

References

[Pierre Saouter](#) ✉, Head of Data Science at FlowBank

[John Coene](#) ✉, Founder and Lead developer at Opifex

[Ben Meuleman](#) ✉, Statistician at University of Geneva

[Meirav Banon Skornik](#) ✉, Senior Analyst at CICAD

[Jose Nunes](#) ✉, Professor of Biostatistics at the University of Geneva

[Eva R. Pool](#) ✉, Professor at Campus Biotech