

Arthur DANJOU

Data Science & Applied AI | Available for a final-year internship starting April 2026.

✉ arthur.danjou@dauphine.eu ☎ +33 7 82 57 64 85 📍 Paris, France 🌐 arthurdanjou.fr 🐙 [arthurdanjou](https://github.com/arthurdanjou) 🌱 [arthurdanjou](https://www.linkedin.com/in/arthurdanjou)

Profile

Master's student in Applied Mathematics – Statistical and Financial Engineering (Data Science track) at Paris-Dauphine University (PSL), I am looking for a **final-year internship** in the field of **Applied Artificial Intelligence**. Passionate about **statistical** modeling, machine learning, and their real-world implementation, I aim to contribute to innovative projects as a **Data Scientist** or **Applied AI Researcher**. I am fully committed to initiatives that combine **mathematical modeling** and **practical applications**. Rigorous, curious, and motivated by **applied research**, I aspire to deepen both my theoretical and practical skills, with the long-term goal of pursuing a **PhD in Artificial Intelligence**.

Work Experience

ArtDanjProduction

AUTO-ENTREPRENEUR – REMOTE

FEB 2022 – TODAY

- Administration of a personal **homelab** (servers, databases, storage, backups) used as a platform for machine learning projects and MLOps experimentation.
- Design, development, and maintenance of web, data, and cloud projects using Python, TypeScript, Vue.js, Nuxt 3, and R.
- Maintaining and updating live production projects.
- Creating technical documentation and blog content, and publishing on my personal website (arthurdanjou.fr).
- Actively learning and applying **new technologies** (LLMs, MLOps, NLP, Agents IA) through personal projects within a DevOps environment.

Sevetys

DATA ENGINEER INTERN – PARIS, FRANCE

JUN – JULY 2025

- Designing and implementing data cleaning protocols for client and patient data.
- Implemented daily data cleaning pipelines on Azure with PySpark, enhancing data quality by **20–30%** depending on the clinic.
- Developed an **automated monthly data quality report** tracking completeness, consistency, and usability metrics by clinic.
- Worked closely with business teams to validate and optimize data quality rules.

Education

MSc in Statistical and Financial Engineering, Data Science track (2nd year)

PARIS DAUPHINE-PSL UNIVERSITY – PARIS, FRANCE

Key courses: Machine Learning (theory and practice), Deep Learning, NLP, Reinforcement Learning, Clustering, Non-parametric Estimation, Learning Theory, Generative AI, Data Quality, Data Visualization (R), SQL.

Capstone project: Data Project – End-to-end data science project in the format of a "Hackathon 3.0".

MSc in Mathematics and Applications, Statistics track (1st year)

PARIS DAUPHINE-PSL UNIVERSITY – PARIS, FRANCE

Key courses: Supervised Statistical Learning, Generalized Linear Models (GLMs), Monte Carlo Methods, Data Analysis, Non-parametric Statistics, Time Series, Numerical Optimization.

Projects completed:

- Monte Carlo Simulation** (R & LaTeX): Analyzed and compared different estimators through Monte Carlo methods and algorithmic simulations.
- Generalised Linear Models** (R & LaTeX): Developed a linear regression model to predict bike-sharing demand, assessing both predictive performance and model adequacy.
- Statistical Learning** (Python & LaTeX): Implemented and compared supervised learning algorithms (k-NN, neural networks, logistic regression, Naive Bayes) for breast cancer detection and patient classification.

BSc in Mathematics (Licence – French undergraduate degree)

PARIS-SACLAY UNIVERSITY – ORSAY, FRANCE

- Schelling Segregation Model** (Python & LaTeX): Statistical and probabilistic analysis of Schelling's segregation model, with simulation and results visualization using Python.

Skills

- Programation:** Python, R, TypeScript, Java, Git et LaTeX.
- Librairies & Frameworks:** NumPy, Pandas, Scikit-learn, PyTorch, Matplotlib, Seaborn
- Databases:** SQL, Redis
- Cloud & DevOps:** Proxmox, Docker, Azure, Linux, SysAdmin
- Multivariate Data Analysis:** Principal Component Analysis (PCA), Correspondence Analysis (CA), clustering techniques, and correlation analysis
- Supervised Learning:** k-NN, linear and logistic regression, CNN, Naive Bayes
- Unsupervised Learning:** clustering, dimensionality reduction, k-means, CNN
- IA & Modern Models:** Natural Language Processing, Transformers, Large Language Models, AI agents, embeddings, and fine-tuning

Soft Skills

Strong Analytical Skills – Scientific Curiosity – Technical Independence – Effective Communication – Collaborative Agile Teamwork

Languages

- Native French
- English (B2/C1 proficiency)
- Spanish (A2 proficiency)

Interests & Hobbies

AI and related concepts, Applied Mathematics, Personal cloud projects (homelab), Team sports

Portfolio

All my academic and personal projects are available on my website:

arthurdanjou.fr.

You can also find my blog here:

<https://go.arthurdanjou.fr/writings>