

# Baptiste Gorteau

4710 rue Fulton  
Montréal, QC, Canada  
[LinkedIn](#), [GitHub](#)

baptiste.gorteau@gmail.com  
[bgorteau.github.io](#)  
+33 7 83 47 87 18

---

## EDUCATION

**Rennes 2 University**, Rennes, France

*Master of Science*, Applied Mathematics and Statistics, September 2023

*Relevant Courseworks*: Python Programming, R Programming, Linear Models, Time Series, Deep Learning

**Rennes 2 University**, Rennes, France

*Bachelor of Science*, Mathematics and Computer Science Applied to the Humanities and Social Sciences, June 2021

*Relevant Courseworks*: Python Programming, Database Management, Algebra, Data Mining, Economy, Sociology

## EXPERIENCE

**Data Consultant**

April 2024 – April 2025

UNESCO GEM Report Team

Paris, France (Remote)

- Developed a data pipeline using R and SQL to build an education aid database from International Aid Transparency Initiative (IATI) data.
- Developed automated Excel reports using R and Python to extract insight from the database.
- Collaboration with UNICEF teams to write a report on the use of IATI data in the education aid environment.

**Developer – Data Scientist**

May 2022 – September 2023

Capgemini T.S.

Rennes, France

- Developed management rules in Python using PySpark, and created tests with Pytest to ensure their functionality.
- Contributed to the development of a Big Data ETL pipeline using HDFS, PySpark, NiFi, and cloud platforms.
- Built a system to detect anomalies in log files using various machine learning methods, leveraging Scikit-Learn and TensorFlow frameworks.
- Debugged Python, R, and SQL code.

**Data Scientist Internship**

April 2021 – July 2021

Movement, Sport and Health Laboratory

Bruz, France

- Analyzed and visualized tracking data using Python to extract insight for a research project combining soccer performance analysis and virtual reality technologies.
- Analyzed time series from tracking data to compare player trajectories and develop predictive movement models.
- Created reports and presentations using LaTeX and PowerPoint.

## PROJECTS

**Creation of a database of long-distance and middle-distance running performances** – Personal Project (April 2025)

Scraped data from PDF files using Python to build a database of 100m split times for middle- and long-distance races at international athletics competitions (2019–present).

**Stathletes 2025 Big Data Cup Finalist** – Data Competition (February 2025)

Creation of different metrics to calculate the availability of hockey players on the ice rink using their tracking data.

## SKILLS

**Languages:** French (Proficient), English (Proficient)

**Software and Languages:** Python, R(Studio), SQL, HTML, Microsoft Office — (Word, Excel, PowerPoint, Outlook),  $\text{\LaTeX}$ .

**Python packages:** Pandas, Numpy, BeautifulSoup, Selenium, Scikit Learn, TensorFlow, Matplotlib