

Student in Data Science

Romane LE GOFF

+33 608831762 | romane.le-goff@laposte.net | romanelgff.github.io

TECHNICAL SKILLS

Programming: R (RShiny, keras), Python (Matplotlib, Bokeh, Tensorflow, sklearn), SAS, SQL, XML
Data Science Algorithms: Regression (Linear, Multiple Linear, Ridge, Lasso), Classification (Naïve Bayes, KNN, SVM, Decision Tree, Random Forest), Text mining (NLP), Clustering (K-Means, Hierarchical), Deep Learning
Public policy: Policy analysis (Experimental economics, Micro-econometrics), Decision making (Cost-Benefits Analysis), Game theory, Content writing
Software: LaTeX, Office package, Elasticsearch, SPSS

EDUCATION

Université de Rennes 1 - ENSAI Rennes, France
Master in Applied Statistics : Public Evaluation & Decision Making *Sept 2020 – Expected to graduate in 2022*
• European Master in Official Statistics

University of Plymouth Plymouth, United-Kingdom
Mathematics and Economics BSc., Year 3 (Erasmus) *Sept 2019 – May 2020*

Université de Bretagne Occidentale Brest, France
Mathematics and Computer Sciences Applied to Human and Social Sciences BSc., Year 1 & 2 *Sept 2017 – May 2019*
• Mention "Bien" - *Equivalent to a 2:1 honours degree*

EXPERIENCES

Data scientist internship 28/02/21 – now
IQVIA *Courbevoie, France*

- Six-month research internship with the purpose of finding machine learning methods to cluster care pathways for diabetic patients using pharmacy deliveries of anti-diabetic treatments. *Programming languages used: SAS, R*

Data analyst internship 26/04/21 – 23/07/21
Atlantic Maritime Prefecture *Brest, France*

- Created automated dashboards for the Commander of the Defence Base and for the Social Action of the Armies (ASA) with R Shiny and Elasticsearch, which became to be a great asset in their day-to-day work.
- Carried out a study using ASA data that determined the best location for the construction of a military nursery on the defence base, while highlighting the main familial trends in the area.

PROJECTS

Academics | R-Python 2020-2021

- Predicting diagnosis of malignant pleural mesothelioma with patient health records ([article](#)) *Made with Python*
- Evaluation of the closed relationship between CO2 emissions and GDP per capita using econometric modelling methods. ([Kuznets curve](#)) *Made with R and LaTeX*
- Web application designed for fully interactive data visualisation, mapping, graphing and statistical tables on deliveries and stocks of COVID-19 vaccines at the beginning of 2021. *Made with RShiny*

Personal | Python 2021

- Creating CNNs that classify and label images from 16,000 coloured images in order to split them into 3 categories (17 subclasses). ([GitHub](#)) Conducted in the frame of an online course ([365DataScience](#)). *Made with Python*

ABOUT ME

Languages: French (native language), English (C1), Spanish (B1), Italian (A2)
Unique experiences: 2020 BUCS Cross Country Championships in Edinburgh ; 2021 Saint-Pol - Morlaix half-marathon ; 2022 Paris Marathon