

# Biostatisticienne

Romane LE GOFF

+33 608831762 | [romane.le-goff@laposte.net](mailto:romane.le-goff@laposte.net) | [romanelgff.github.io](https://romanelgff.github.io)

## COMPÉTENCES

**Programmation:** R (RShiny, caret, keras), Python (tensorflow, sklearn, pandas), SAS, SQL, XML  
**Algorithmes en Data Science:** Régression (Linéaire Multiple, Ridge, Lasso), Classification (PPV, SVM, Arbres de décision, Forêts aléatoires), Analyse textuelle (NLP), Clustering (K-Means, Hiérarchique), Deep Learning (CNN)  
**Logiciels:** LaTeX, Pack Office, Git

## EDUCATION

**Université Rennes 1 - ENSAI** Rennes, France  
*Master de Mathématiques Appliquées, Statistiques parcours Evaluation et Décision Publique* Sept 2020 – Sept 2022  
• European Master in Official Statistics

**University of Plymouth** Plymouth, Royaume-Uni  
*L3 Mathématiques Appliquées, Economie - Erasmus+* Sept 2019 – Mai 2020

**Université de Bretagne Occidentale** Brest, France  
*Licence Mathématiques et Informatique Appliquées aux Sciences Humaines et Sociales (L1-L2)* Sept 2017 – Mai 2019

## EXPERIENCE

**Biostatisticienne** Septembre 2022 – Présent  
*IQVIA* Courbevoie, France  
• Expérience sur le Système Nationale des Données de Santé (SNDS) et sur les bases de données IQVIA (Electronic Medical Records, délivrances en pharmacie).

**Stage data scientist** Mars 2022 – Août 2022  
*IQVIA* Courbevoie, France  
• Stage de recherche d'une durée de 6 mois dont le but était de fournir et de comparer des méthodes permettant le regroupement de trajectoires de traitements similaires, afin d'identifier des "parcours-types" de patients.  
• *Programmation: SAS, R*

**Stage en visualisation de données** Avril 2021 – Juillet 2021  
*CECLANT - Préfecture maritime de l'Atlantique* Brest, France  
• Création de tableaux de bord automatisés pour le Commandant de la Base de Défense de Brest-Lorient et pour l'Action Sociale des Armées (ASA) avec R Shiny.  
• *Modules R utilisés : RMarkdown, Shinydashboard, plotly, ggplot2, RAmCharts, leaflet*

## PROJETS

**Académiques | R** 2020-2021  
• Application Web entièrement interactive conçue pour la visualisation de données: cartographie, représentation graphique et tables statistiques informatives sur les livraisons et les stocks de vaccins de la COVID-19 au début de l'année 2021. ([Vaccins-COVID19](#)). *Réalisé avec RShiny*  
• Modélisation de la maladie de l'artère coronaire à l'aide des réseaux bayésiens, dans le but de mieux comprendre les diagnostics liés à cette maladie. ([GitHub](#)). *Réalisé avec R*  
• Analyse de la relation étroite entre les émissions de CO2 et le PIB par habitant par une modélisation économétrique. ([courbe de Kuznets](#)) *Réalisé avec R*

**Personnel | Python** 2021  
• Création de réseaux neuronaux conventionnels à partir de 16000 images de vêtements, dans un objectif de classification et d'étiquetage en 3 catégories et 17 sous-catégories. ([GitHub](#)) Ce projet a été réalisé dans le cadre d'un cours suivi en ligne ([365DataScience](#)). *Réalisé avec Python (Tensorflow)*