

EMMANUEL PAGUIEL BOUENDO



emmanuelpaguel@gmail.com



06 95 23 93 47



Portfolio & Projets | GitHub



TOURS, FRANCE

Étudiant en Master 2 Économiste d'Entreprise (MÉcEn), je me forme en data science, machine learning et économétrie. Je recherche un stage de fin d'études (6 mois) à partir d'avril 2026 dans les domaines de la data science, de la business intelligence ou de l'analyse économique.

FORMATION

Master Économie d'Entreprise et des Marchés (MÉcEn)

Université de Tours / 2024 - 2026 (en cours)

Spécialisation : Data Science, Machine Learning, Modélisation Économétrique.

- Modules principaux : Analyse de la concurrence & Théorie des jeux, Pricing & Segmentation, Méthodes micro-économétriques, Marketing Analytics, Entrepreneuriat, Web scraping & Machine Learning, Système de gestion de base de données, Développement logiciel.

Licence Économie - Parcours Économiste de l'Entreprise

Université de Tours / 2023 - 2024

- Théorie micro/macroéconomique, Mathématiques appliquées à l'économie, Statistiques Fréquentiste & bayésiennes, Économétrie.

EXPERIENCES PROFESSIONNELLES

Conseiller Clientèle - Société Générale Congo

Brazzaville, Congo / Avril 2019 - Août 2022

- Analyse des besoins clients et recommandation de produits financiers à partir de critères de risque et de profil client.
- Instruction de dossiers de crédit (analyse de risque, conformité)
- TraITEMENT back-office et gestion de portefeuille client.
- Compétences : relation client, analyse financière, gestion du risque, CRM

COMPÉTENCES TECHNIQUES

Machine Learning	XGBoost, Random Forest, SVM, Régression (linéaire, logistique), Classification, Clustering (K-means), Feature Engineering, Validation croisée, Sélection de modèles, SHAP.
Langages	Python: Pandas, NumPy, scikit-learn, Matplotlib, Plotly, Seaborn, Streamlit. R: Shiny, tidyverse, ggplot2, dplyr, tidymodels
Outils et Plateformes	Git/GitHub, SQL, MySQL, Power BI, Tableau, Excel VBA, Jupyter, VS Code, Spyder.
Analyse et statistiques	Statistiques descriptives & inférentielles, Tests d'hypothèses, Modélisation économétrique, Data Mining, Web Scraping (BeautifulSoup), Data Cleaning.

🌐 LANGUES : Français (natif) / Anglais courant

↳ Centres d'intérêt : Dessin (croquis, illustration) • Développement de projets open source en data science

PROJETS

Prédicteur de Salaires – Métiers de la Data

Janvier 2026 • projetmachinelearning.streamlit.app

Application exploratoire des déterminants salariaux dans les métiers de la Data.

À partir de 5 868 offres d'emploi HelloWork (janvier 2026), ce projet analyse les facteurs associés aux niveaux de rémunération (expérience, compétences, localisation, secteur).

Un modèle XGBoost ($R^2 = 0.337$) a été retenu après comparaison. L'interface propose une estimation personnalisée, une interprétation SHAP, et une feuille de route carrière basée sur des profils similaires.

Modélisation Prédictive du Marketing

Bancaire | Octobre 2025

Analyse des facteurs associés à la souscription à un dépôt à terme.

Ce projet explore les déterminants de la conversion client dans une campagne télémarketing bancaire (Portugal, 2008 - 2010, 41 000 clients).

Après comparaison d'algorithmes, le Random Forest (AUC = 0.945) a été retenu pour estimer la propension à souscrire. Une segmentation client (K-means) et un prototype de scoring sont intégrés à une application Shiny.

Évaluation Comparative des Modèles de Prédiction du Bonheur | Mars 2025

Analyse des liens entre bien-être subjectif et variables socio-économiques à partir de données d'enquête.

Ce projet compare la capacité de 7 algorithmes (Random Forest, SVM, XGBoost, etc.) à discriminer les niveaux de bien-être subjectif à partir de données d'enquête. L'évaluation repose sur des métriques complémentaires (AUC, F1-score, précision, rappel) et une validation croisée.

Analyse des Accidents Routiers en France (2015–2024) | Janvier 2025

paguystream.shinyapps.io/CrimStream2

Exploration des tendances criminelles locales à partir de données officielles.

À partir de la base BAAC (ministère de l'Intérieur, 2015–2024), ce projet permet d'observer les variations géographiques, temporelles et contextuelles des accidents corporels.

L'application Shiny offre des indicateurs agrégés par territoire, des visualisations cartographiques, et une analyse des facteurs associés à la gravité des accidents.