



# Data Scientist

BALLOGOU Essi Carole Claudia

ballogouessicarole@gmail.com

+33 6 60 67 98 98

67000 Strasbourg

github.com/ballogou-essi

linkedin.com/in/essi-ballogou

## FORMATIONS

### Master Économie - Data science - Université de Strasbourg | Strasbourg, France

Cours : Deep Learning, Machine Learning, Reinforcement Learning, Econométrie, Statistiques

Sept. 2024 (En Cours)

### Licence Économie et Gestion- Université Clermont Auvergne | Clermont-Ferrand, France

Cours : Mathématiques, Statistiques, Économétrie, Analyse financière, Macroéconomie

Sept. 2022 - Mai 2024

## EXPÉRIENCES

### Data Analyst – Axa France | Strasbourg, France

- Python, PySpark, SQL et Power Platform

Sept. 2025 (en cours)

### Animatrice - UFCV (Vacances adaptés - Handicap) | Clermont-Ferrand, France

- Respect des protocoles de sécurité, planification et organisation des activités adaptées

Dec.2022 - Janv. 2023

### Tutrice en Economie et en Mathématiques – Soutien scolaire | Lomé, TOGO

- Explication claire et adaptée aux besoins des élèves

Sept.2020 - Dec. 2021

## COMPÉTENCES TECHNIQUES

- **Data Science** : PySpark, Pandas, NumPy, Scikit-learn, PyTorch
- **Visualisation** : Power BI
- **Bases de données** : Power Query, Excel, SQL
- **Languages de programmation** : Python, R, SQL, SAS

## PROJETS

### Migration de programmes SAS vers PySpark

- Migration de traitements SAS vers PySpark dans le cadre de la transition vers un Datalake Azure Databricks.
- Réécriture et optimisation des pipelines de transformation de données

### Outil de pilotage Power BI avec Power App intégré - Gestion des dérogations

- Conception d'un tableau de bord Power BI centralisant les données liées aux dérogations
- Développement de mesures **DAX** (KPI dynamiques, calculs conditionnels, suivi des statuts)
- Automatisation des flux via **Power Automate** pour fiabiliser et structurer le processus.

### Land-Use Classification from Satellite Images - Deep Learning (EuroSAT)

- Développement d'un **modèle de classification d'images** satellite pour identifier les types d'occupation des sols
- Mise en place d'un pipeline complet : prétraitement des images, data augmentation, entraînement et évaluation du modèle
- Analyse des performances via accuracy, confusion matrix et métriques de classification multi-classes.

### Système de recommandation de films - MovieLens 25M

- Manipulation d'un dataset à grande échelle (25M d'interactions) et gestion d'une matrice user-item creuse
- Comparaison des approches en termes de précision et de qualité des recommandations.

## LANGUES

- **Français** : Langue maternelle
- **Anglais** : Niveau courant (B2)

## CENTRES D'INTÉRÊT

- **Dessin** : Je cultive ma créativité et ma capacité à représenter des concepts de manière visuelle et structurée
- Veille technologique et scientifique via **ScienceDirect, Medium, Academia**
- **Chant en choeur (MezzoSoprano)** : Je développe l'écoute active, la coordination et l'esprit d'équipe