

## EDUCATION

Université de Montpellier   <i>DU Big Data, Data Science et Analyse des risques</i>	10/2023 – 09/2024
Université de Caen   <i>Master 2 Statistiques Appliquées et Analyse Décisionnelle</i>	09/2021 – 09/2023
Université de Caen   <i>Licence 3 Mathématiques Appliquées et Licence 3 Informatique</i>	09/2019 – 09/2021

## COMPETENCES

- Data Engineering:** Apache Kafka, Hadoop, Spark, Airflow, AWS Glue, Databricks
- ETL & Data Pipeline:** Python (Pandas, PySpark), SQL, Airflow, Talend
- Cloud & DevOps:** AWS (S3, EC2, Lambda), GCP, Docker, Kubernetes, Terraform, Ansible
- Big Data Tools:** Apache Hadoop, Hive, Spark, HDFS
- Gestion de bases de données :** SQL/NoSQL, PostgreSQL, MongoDB, Cassandra
- Data Analytics & Modeling:** Dimensional Modeling, OLAP, Data Warehousing
- Version Control & CI/CD :** Git, Jenkins, GitLab CI
- Gestion de projet :** Agile/Scrum (Jira, Trello, Slack)

## EXPERIENCES

<b>TrouveTout</b> - (Start Up)   Data Engineer & Developpeur Backend ( <i>WebApp</i> )	07/2023 – 05/2024
<ul style="list-style-type: none"><li><b>Data Pipeline Development :</b> Conception et mise en place de processus ETL en Python et Airflow pour le traitement des données de la marketplace.</li><li><b>Développement Backend :</b> Création d'un backend évolutif avec Django pour la gestion de l'ingestion, la transformation et le stockage des données.</li><li><b>CI/CD Pipelines :</b> Mise en place de pipelines CI/CD automatisés avec Docker, Jenkins et GitLab, réduisant le temps de déploiement de 30 %.</li><li><b>Data Quality Monitoring:</b> Implémentation de contrôles de validation des données et détection d'anomalies pour assurer l'intégrité des données.</li></ul>	
<b>ORANGE</b> - (Telecom Sector)   Stage <i>Data Science/Engineer</i>	11/2022 – 05/2023
<ul style="list-style-type: none"><li><b>Data Infrastructure:</b> Developed a data lake architecture using Hadoop and Spark to handle over 200 TB of telecom data.</li><li><b>Real-time Data Processing:</b> Deployed Kafka for real-time streaming of customer data, enabling better churn analysis.</li><li><b>Cloud Migration:</b> Assisted in migrating on-premises data infrastructure to AWS, utilizing S3, Glue, and Redshift.</li><li><b>Predictive Modeling:</b> Supported the data science team by providing scalable data pipelines for churn prediction using Apache Spark.</li></ul>	
<b>Laboratoire de Mathématiques Nicolas ORESME</b>   Stage <i>Data Analyst</i>	03/2021 – 07/2021
<ul style="list-style-type: none"><li><b>Data Transformation :</b> Optimisation des pipelines de données pour un projet de recherche, réduisant le temps de traitement de 20 %.</li><li>Développement de dashboards interactifs avec SAS Visual Analytics, permettant une meilleure compréhension des relations entre les variables.</li></ul>	

## CERTIFICATIONS & PROJETS PERSONNELS

**AWS Cloud Practitioner | Microsoft Certified: Azure Fundamentals | Databricks Certified Associate Dev**

- Customer Segmentation (Orange Telecom) :** Conception de pipelines de données évolutifs pour la segmentation des clients avec Apache Spark, améliorant l'efficacité des campagnes marketing de 20 %.
- Projet Data Warehousing :** Création d'un data Warehouse sur AWS Redshift pour une plateforme e-commerce, augmentant les performances des requêtes de 40 %.
- Real-time Pipeline Data Processing :** Développement d'un pipeline temps réel avec Kafka et Spark pour le traitement des données de capteurs IoT.