

# Xavier Rousseau

Data Engineer & Analyst | De l'architecture à l'analyse



 rousseau.pro@outlook.fr

 [linkedin](#)

 [github](#)

 +336 95 29 05 05

 Landerneau, France

## PROFIL

Je travaille depuis plus de 10 ans dans la donnée, avec une approche simple : rendre les données plus fiables, plus compréhensibles et vraiment utiles. J'ai souvent été le lien entre les équipes métiers, la DSI et les data engineers, pour structurer les référentiels, automatiser les traitements et fiabiliser les flux.

Depuis quelques années, je m'oriente progressivement vers des missions plus techniques autour des pipelines, de l'architecture et de l'orchestration. Je complète cette évolution par une formation en data engineering, pour gagner en maîtrise sur les outils comme Airflow, Docker, Spark ou les bases cloud.

Ce que je cherche : des projets concrets, bien pensés, où je peux apporter à la fois ma vision fonctionnelle et mes compétences techniques.

## COMPETENCES CLES

### Architecture & orchestration des données

- Mise en place de pipelines avec Airflow, Kestra et Airbyte
- Traitement des données avec Python (nettoyage, jointures, agrégations simples)
- Participation à la création d'un data lake basé sur Hadoop et Hive chez Financo
- Conteneurisation avec Docker
- Stockage et ingestion de fichiers avec MinIO, premières expériences sur GCP

### Bases & technologies de données

- Utilisation régulière de PostgreSQL dans différents projets
- Expérimenté aussi avec MongoDB, MySQL et SQLite

### Analyse & visualisation

- Analyses exploratoires, détection de problèmes, création d'indicateurs
- Automatisation de traitements avec Python, SQL, SAS ou R selon les contextes
- Conception de dashboards avec Power BI, Tableau ou Metabase

### Gouvernance & qualité des données

- Fiabilisation des données et suivi des corrections
- Documentation des règles de gestion et suivi des corrections
- Mise en place de contrôles pour limiter les erreurs dans les flux et les bases

### Collaboration & accompagnement

- Recueil des besoins, rédaction claire, tests et documentation
- Interface naturelle entre les équipes métiers, IT et data
- Capacité à faire le lien entre les enjeux du terrain et les contraintes techniques

## FORMATION

### **Formation Data Engineer – OpenClassrooms (en cours) :**

Formation professionnalisante centrée sur des projets concrets, avec une approche moderne de l'ingénierie des données. J'y ai appris à concevoir des pipelines robustes, à manipuler différentes bases de données (relationnelles et NoSQL), à orchestrer des traitements, et à comprendre les fondements d'une architecture orientée services. J'ai aussi découvert des outils que je continue à approfondir, comme Airbyte, Kestra, ClickHouse ou encore Flower pour le monitoring.

 **Master 2 – Statistiques & Économétrie** – Université de Rennes I – 2010 à 2011

 **Licence – Mathématiques** – Université de Bretagne Occidentale – 2008 à 2009

## EXPERIENCE

### 📌 Consultant Data Analyst Senior | Jack Data

Sept. 2023 – Oct. 2024 | Paris | À distance

Mission chez BNP Paribas autour d'un projet stratégique de finance durable : refonte d'un modèle d'impact environnemental.

- Diagnostic de la chaîne analytique existante, cartographie des risques, identification d'anomalies
- Refonte complète du modèle : nettoyage, transformation, indicateurs consolidés, score final
- Documentation des règles de gestion, traçabilité complète du traitement pour les audits
- Collaboration étroite avec la direction data, les métiers et la DSI
- Automatisation des traitements (Python, SQL), restitution sous Power BI

### 📌 Data Analyst | Qwant

Juin 2022 – Fév. 2023 | Nice | À distance

Amélioration de la qualité des données et des outils décisionnels dans un contexte agile.

- Refonte des dashboards (architecture, validation, suivi)
- Structuration des bases analytiques (logs, clics, segments utilisateurs)
- Automatisation d'analyses récurrentes avec Python et SQL
- Visualisation des données avec Kibana et découverte de la suite Elastic
- Premiers usages d'Apache Zeppelin pour l'analyse exploratoire
- Découverte de Prometheus et Grafana pour le suivi des indicateurs techniques
- POC et benchmark dataviz (Tableau, Superset, Metabase) pour mieux répondre aux besoins utilisateurs

### 📌 Data Analyst | Financo (Groupe Arkéa)

Mars 2014 – Mai 2022 | Guipavas

8 ans sur des projets structurants autour des référentiels, de la qualité des données et de la modernisation des traitements.

- Référent métier pour la mise en place du data lake marketing / commerce, basé sur Hadoop et Hive
- Co-construction avec la MOE des premières briques : tables Oracle, vues matérialisées, logique datamart
- Conception et fiabilisation d'un référentiel client unifié : cadrage, règles de gestion, documentation
- Développement de traitements automatisés (SQL, SAS) pour les analyses et les ciblage
- Migration des requêtes SAS Desktop vers SAS Serveur pour améliorer la performance et la maintenabilité
- Modélisation des portefeuilles clients et contrats pour le pilotage commercial et budgétaire
- Mise en place de contrôles qualité et prise en compte des exigences RGPD dans les traitements
- Création de tableaux de bord interactifs pour les métiers sous Tableau
- Participation aux comités SI et aux projets de transformation du système décisionnel

### 📌 Consultant Décisionnel & AMOA | Anaya

Oct. 2011 – Déc. 2013 | Paris

Mission longue au Crédit Agricole Consumer Finance (20 mois), puis au Ministère de l'Économie (8 mois), sur des projets de scoring, refonte décisionnelle et pilotage.

**Crédit Agricole Consumer Finance :**

- Tests fonctionnels du moteur de règles crédit (ONE HORUS) avec outils automatisés (SAS, DSV Toolkit)
- Simulation d'un score unifié pour Sofinco et Finaref, étude de solvabilité client
- Reporting marketing automatisé (SAS Add-in), analyse de résultats d'enquêtes
- Recette fonctionnelle des bases de ciblage utilisées par le CRM

**Ministère de l'Économie :**

- AMOA pour l'outil de pilotage du recouvrement : recueil des besoins, suivi MOE/MOA
- Vérification fonctionnelle, requêtes SAS, accompagnement des utilisateurs

## PROJETS PERSONNELS

🔗 **Datalyzer** - Application web d'analyse exploratoire de données. Développée sous Streamlit, cette app propose une analyse interactive : nettoyage (valeurs manquantes, doublons, outliers), scoring de qualité, visualisations, snapshots, export. Architecture modulaire avec composants indépendants. Évolution prévue vers une version API avec FastAPI.

🔗 **Stack data open-source complète** - Conception d'une stack unifiée pour orchestrer, transformer, valider et interroger des données (structurées et semi-structurées). Intégration prévue du RAG (Retrieval-Augmented Generation) pour enrichir les analyses avec des modules LLM.