

# Babacar SARR

## Data Engineer | Data Analyst

📍 Paris | ✉️ [ba66bacar@gmail.com](mailto:ba66bacar@gmail.com) | ☎️ 07 75 80 67 99  
📁 [Portfolio](#) | 🔗 [GitHub](#) | 🔗 [LinkedIn](#)

### 🎯 Profil Professionnel

Actuellement alternant en Ingénierie des Données à l'EPSI, je suis passionné par la gestion, le traitement et l'analyse des données. Je recherche une **alternance (1 semaine école / 2 semaines entreprise)** pour appliquer mes compétences sur des projets innovants et contribuer à des solutions data-driven.

### 🚀 Projets Académiques

#### Pipeline de traitement en temps réel – Spark & Kafka (EPSI, Fev 2025)

- **Développé** un **Producer** pour collecter et envoyer des flux de données en temps réel vers **Apache Kafka**.
- Implémenté un **Consumer Spark Streaming**, réduisant le temps de latence et garantissant un **traitement en quasi temps réel**.
- **Configuré** Kafka pour assurer la haute disponibilité et la gestion efficace des topics avec **Conduktor**.
- **Optimisé** les performances en ajustant les partitions Kafka et les batchs Spark pour un traitement rapide.

#### Déploiement d'une Architecture Décisionnelle (EPSI, Oct 2024 - Dec 2024)

- **Automatisé** la collecte et le stockage des données VTC de New York dans un **Data Lake MinIO** via des scripts **Python**.
- **Développement d'un pipeline ETL avec Amphi.ai pour transférer un modèle en flocon vers PostgreSQL**.
- **Créé un tableau de bord interactif** sous **Power BI** pour l'analyse et la prise de décision.
- **Orchestré** la mise à jour mensuelle automatique des datasets avec **Apache Airflow**.

### 👛 Expériences Professionnelles

#### Data Analyst (Stage) – Malick Galaxie Groupe, Dakar (Nov. 2023 – Mai 2024)

- Mené une enquête d'évaluation sur le Projet de Désenclavement des Zones de Production (PDZP Dakar), **analysant plus de 500 réponses** pour identifier les principales attentes des bénéficiaires.
- Conçu des questionnaires électroniques sous kobotoolbox et Paramétré les questionnaires sous **kobocollect** pour gérer la **collecte des données** via les smartphones.
- Suivi à distance la collecte faite par les enquêteurs sur le terrain et Transfert des données vers **Microsoft Excel**.
- Apuré les données avec **Power Query** en faisant le **nettoyage et la normalisation** des données.
- Créé des graphiques (KPI's et autres) et des tableaux de bords sous **Power BI** pour une bonne visualisation et les prises de décisions.

#### Développeur Full-Stack (Freelance) – École Leeru Sangabi, Dakar (Fév. 2023 – Août 2023)

- géré les informations de plus de **1000 élèves** en implémentant une **base de données relationnelle (MySQL)** pour organiser les données relatives aux élèves, classes, matières et notes.
- Mise en place d'un **système d'authentification sécurisé** pour différents types d'utilisateurs (administrateurs, enseignants, super administrateurs, élèves).
- Géré les **permissions d'accès** aux fonctionnalités selon les rôles (consultation des notes, ajout/modification des résultats).
- Automatisé les **calculs de moyennes et la génération de bulletins** pour les élèves.

### 📚 Formation

- ♦ **MSc Ingénierie des Données** – EPSI, La Défense (Depuis 2024)
- ♦ **Master 1 Génie Logiciel** – ISI, Dakar (2023 – 2024)
- ♦ **Licence Génie Logiciel** – Unipro, Dakar (2020 – 2023)

### 🔧 Compétences

**Bases de données** : PostgreSQL, MySQL, MongoDB, SQL Server | **Big Data & Cloud** : Spark, Hadoop, Kafka, AWS (Redshift, S3) | **Programmation** : Python (pandas, NumPy, scikit-learn), SQL, R, Scala | **ETL & Orchestration** : Apache Airflow, Docker, Kubernetes, Dataiku | **Visualisation & BI** : Power BI, Excel, Amphi.ai | **Gestion de projet** : Scrum, Kanban, Git, GitHub, Jira, Trello

### 🌍 Langues & Soft Skills

- **Français** : C2 | **Anglais** : B1 (Lecture technique assurée)
- **Adaptabilité, Esprit analytique, Collaboration, Autonomie**