

Mehdi Somrani

Data Engineer

mehdisomrani.netlify.app | GitHub: somranimehdi | LinkedIn: Mehdi SOMRANI

Email: MehdiSomrani@proton.me | Téléphone: +216 XX XXX XXX

PROFIL

Data Engineer avec 3+ ans d'expérience dans l'optimisation de plateformes analytiques à grande échelle et de solutions BI. Expert en Python, SQL et PySpark, axé sur l'amélioration de la fiabilité des données et la collaboration d'équipe pour maximiser l'impact business. Capacité éprouvée à concevoir des pipelines de données de bout en bout à travers des projets personnels avancés.

EXPÉRIENCE PROFESSIONNELLE

Ingénieur Data/BI | Orange Tunisie 08/2023 – Présent

Équipe BI Marketing : Analytique, Reporting & Pipelines

- Segmenté 10M d'enregistrements/jour via **SQL** et **PySpark** pour 100+ campagnes.
- Créé des tableaux de bord **Power BI** et scénarios CVM pour 3M+ de clients.
- Automatisé le reporting et optimisé **SQL** (gain de 7h/semaine, runtime -40%).
- Développé des modèles de prévision ML en **Python** (précision 80-85%).

Stagiaire Big Data | Orange Tunisie 02/2023 – 05/2023

Équipe Réseau & Services : Infrastructure & Analytique Temps Réel

- Déployé un cluster **Hadoop** (3 nœuds) et pipelines **Kafka** (100k+ rec/h).
- Automatisé les flux (**Apache NiFi**, **Shell**) et le suivi (**Elastic**, **Kibana**).

PROJETS

Pipeline de Streaming de Données en Temps Réel : Conception d'un pipeline complet conteneurisé pour le traitement en streaming de centaines d'événements en direct par minute.

Technologies : Python, Apache Kafka, Apache Spark, Elasticsearch, Kibana, Docker

Pipeline ETL & ML : Ingestion d'API, transformation JSON et modélisation (79% de précision).

Technologies : Python, Airflow, Pandas, scikit-learn, Docker

Data-Guard : Pipeline Medallion Résilient : Développement d'un pipeline tolérant aux pannes avec Spark, Kafka et Docker pour valider, mettre en quarantaine et fournir des données de ventes au détail de haute qualité.

Technologies : PySpark, Docker, DuckDB, Superset, Power BI

FORMATION

Certification en Data Engineering, Mines Paris – PSL 2024–2025

Master en Business Analytics & Data Science, Université Virtuelle de Tunis 2023–2025

Licence en Informatique (Big Data & Analyse de Données), ISAMM 2020–2023

COMPÉTENCES

Data Engineering : Python, SQL, PySpark, Hadoop, Spark, Airflow, Kafka, Machine Learning, ETL, Data Warehousing, SQL Server, MongoDB

Méthodologies : Agile, Scrum, Kanban

BI & Outils : Power BI, Kibana, Excel, Docker, Git, FastAPI, Linux

Langues : Anglais (EFSET C2), Français (Intermédiaire), Arabe (Langue maternelle)

VIE ASSOCIATIVE & PROJETS COLLABORATIFS

Projet Apoo | Initiative Personnelle (2021 - En pause)

Création et gestion des personnages, conception visuelle et collaboration avec l'OMS et le Croissant-Rouge.

Organisateur | GDG Sfax (09/2023 - 02/2024)

Organisation d'ateliers et de meetups ; accompagnement technique des participants.

Designer Graphique | Data Science Club TBS (10/2021 - 07/2023)

Conception des supports graphiques d'événements et soutien à l'animation des ateliers.

CERTIFICATIONS

Associate Data Engineer : DataCamp (2024) | MS PowerBI (PL300 - Équipe)