

Mehdi Somrani

Ingénieur Data

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

E-mail: MehdiSomrani@proton.me | Tél: +216 58 XXX XXX | Tunis

PROFIL

Ingénieur Data avec plus de 2 ans d'expérience dans l'optimisation de plateformes analytiques à grande échelle et de solutions BI. Expert en Python, SQL et PySpark, avec une forte orientation vers la fiabilité des données et la communication pour générer un impact métier. Capacité avérée 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 Marketing BI : analytique, reporting, pipelines de données et aide à la décision

- Segmentation et analyse de données avec SQL et PySpark sur 10M records quotidiens issus de data lakes.
- Construction de scénarios CVM pour 3 millions de clients, soutenant plus de 100 campagnes B2C/B2B.
- Automatisation des flux de reporting et optimisation des requêtes SQL, permettant un gain de 7+ heures/semaine et une amélioration du temps d'exécution d'environ 40%.
- Développement de modèles de prévision atteignant une précision de 70–85%.

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

Équipe Réseau & Services : infrastructure et analytique temps réel

- Déploiement d'un cluster Hadoop à 3 nœuds sur Ubuntu.
- Mise en place de pipelines d'ingestion Kafka traitant plus de 100 000 enregistrements par heure.
- Automatisation des flux de données avec Apache NiFi et scripts shell ; supervision via Elastic et Kibana.

PROJETS

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

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

Pipeline ETL & ML avec Airflow : Ingestion via API, transformation JSON, modèles de ML (79%).

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

Data-Guard : Pipeline Médaille Résilient : Construction d'un pipeline de données tolérant aux panne, pour valider, mettre en quarantaine et servir des données de ventes retail de haute qualité.

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

FORMATION

Certification en Ingénierie des Données, 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

Ingénierie des Données : Python, SQL, PySpark, Hadoop, Spark, Airflow, Kafka, Machine Learning, ETL, Entrepôt de Données, SQL Server, MongoDB

Méthodologie : Agile, Scrum, Kanban

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

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

VIE ASSOCIATIVE

Projet Apoo | Initiative Personnelle (2021 - En pause)

Création de personnages, gestion des réseaux sociaux, conception visuelle et collaboration avec l'OMS et la Croix-Rouge.

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

Conception de visuels pour les événements et participation aux ateliers.

CERTIFICATIONS

Associate Data Engineer : DataCamp (2024) | MS Power BI (PL300 – The Team)