

Mehdi Somrani

Data Engineer

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

Email: MehdiSomrani@proton.me | Phone: +216 58 XXX XXX | Tunis

SUMMARY

Data Engineer with 2+ years of experience optimizing high-scale analytics platforms and BI solutions. Expert in Python, SQL, and PySpark, with a focus on enhancing data reliability and team communication to drive business impact. Proven ability to architect end-to-end data pipelines through advanced personal projects.

EXPERIENCE

Data/BI Engineer | Orange Tunisie 08/2023 – Present

BI Marketing team: analytics, Reporting, data pipelines, and decision support

- Performed data segmentation and analysis with SQL and PySpark on 10M daily records from data lakes.
- Built CVM scenarios for 3M customers, supporting 100+ B2C/B2B campaigns.
- Automated reporting workflows and optimized SQL queries, saving 7+ hours/week and improving execution time by 40%.
- Developed forecasting models achieving 70–85% accuracy.

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

Network & Services team : infrastructure and real-time analytics

- Deployed a 3-node Hadoop cluster on Ubuntu.
- Implemented Kafka ingestion pipelines handling 100k+ hourly records.
- Automated data flows with Apache NiFi and shell scripting; monitored using the Elastic and Kibana.

PROJECTS

Real-Time Data Streaming Pipeline : Built an end-to-end data pipeline within containers to handle the streaming of 100s of live events/minute.

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

Airflow ETL & ML Pipeline : API ingestion, JSON transformation, regression models (79%).

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

Data-Guard: Resilient Medallion Pipeline : Built a fault-tolerant data pipeline using Spark, Kafka, and Docker to validate, quarantine, and serve high-quality retail sales data.

Tools: Pyspark, Docker, DuckDB, Superset, PowerBI

EDUCATION

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

MSc Business Analytics & Data Science, Virtual University of Tunis 2023–2025

BSc Computer Science (Big Data & Data Analysis), ISAMM 2020–2023

SKILLS

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

Methodology Agile, Scrum, Kanabn

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

Languages: English (EFSET C2), French (Intermediate), Arabic (Native)

COMMUNITY LIFE

Apoo Project | Personal Initiative (2021 - Paused)

Created characters, managed social media, designed visuals, and collaborated with WHO and Red Crescent.

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

Planned workshops and meetups; guided participants technically.

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

Designed event graphics and supported workshops.

CERTIFICATIONS

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