

Né en 2000

1 à 3 ans d'expérience

Réf : 2402151137

Junior data engineer

Ma recherche

Je recherche tout type de contrat, sur toute la région Ile de France, dans le Bâtiment.

Formations

2023 : Bac +5 à École Nationale Supérieure d'Informatique et d'Analyse des Systèmes-ENSIAS sur Rabat, Maroc

Diplome d'ingénieur-Génie Logiciel

2020 : Bac +5 sur Tétouan, Maroc

Classes Préparatoires aux Grandes Écoles Tétouan Filière Technologie et Sciences Industrielles

2018 : Bac +5 à Lycée Technique Moulay Youssef sur Tanger, Maroc

Bac Sciences et Technologies Électriques

Expériences professionnelles

03/2023 - 09/2023 :

Data Engineer Intern chez Namr sur Paris

Développement d'outils de récupération des fichiers d'un datalake de 900.000 sources dans différents formats y compris des données géospatiales avec Python et PostgreSQL sur Dataiku.

- Collecter, analyser, nettoyer, intégrer et documenter des ensembles de données à partir du datalake hébergée sur Google Cloud Platform (GCP).
- Gestion des CI/CD pipelines sur GitLab CI et maintenance des scripts pour la mise à jour des flux de données.
- Collaboration avec les data analystes et les data owners pour rédiger des notebooks d'analyse ainsi que des documentations détaillées des données, dans le cadre d'un environnement Agile

2023 - 2023 :

Projets récents

Morocco Subreddit Data Pipeline

Conception de l'architecture du Data Pipeline □ Extraction des données à partir du Reddit API □

Stockage des données □ ETL □ Reporting.

Python,Apache Airflow,Google Cloud Platform (Composer, Storage, BigQuery),Power BI

Conception et mise en oeuvre d'une Data Warehouse

L'étude de cas □ Définir les tables des faits et des dimensions □ Création de la zone de transit et du

Datawarehouse □ ETL □ Reporting.

Microsoft BI: SSMS, SSIS, SSAS, Power BI

07/2022 - 08/2022 :

Data Engineer Intern chez Ocp Group sur Safi, Maroc

Conception et implémentation des pipelines ETL dans un environnement data center en optimisant les performances des transformations avec Python, PostgreSQL.

- Mettre en place les tests pour assurer la qualité des livrables et l'intégrité des données produites.

- Automatiser des tâches récurrentes telles que la sauvegarde des données, la gestion des journaux, et la suppression des données obsolètes avec Apache Airflow

07/2021 - 08/2021 :

Data Engineer Intern chez Rebbanet sur Safi

Gestion d'une MySQL Database de 100.000 enregistrements et développement du backend et du data pipeline avec Python et Flask.

- Optimisation des schémas du base de données pour améliorer la récupération et stockage des données.
- Création du Dashboard d'analyses pour fournir des informations exploitables avec Power BI.

Langues

Français (Oral : moyen / Ecrit : avancé)

Atouts et compétences

Programmation

Python • Java • C Data Engineering

Apache Airflow • BigQuery • Microsoft BI • Dataiku Cloud Computing GCP • Docker • Vmware

Bases de données PostgreSQL • PL/SQL • MongoDB • MS SQL Server (T-SQL) • Snowflake

Conception Logiciel

UML • Merise

Systèmes d'exploitation

Windows • Linux

Data Science et visualisation

Pandas • Scikit-learn • Power BI

Gestion de projet

Git • Agile (Jira) • GitLab

Permis

Permis B