

## Data Scientist

**Stane**, société à mission, nous révolutionnons la recherche clinique en médecine de ville pour la rendre plus rapide, efficace et accessible. Notre ambition ? Donner un coup d'accélérateur à la recherche clinique tout en allégeant la charge des professionnels de santé et en améliorant l'expérience des patients.

Tes missions, en collaboration avec le directeur Data et l'ensemble des équipes:

### 1- Développement et maintenance des pipelines d'intégration de données :

- Créer et optimiser des pipelines de traitement et d'intégration de données provenant des logiciels de gestion de cabinets médicaux.
- Mettre en œuvre des solutions robustes pour l'extraction, la transformation et le chargement (ETL/ELT) des données dans notre entrepôt de données.
- Automatiser les workflows pour assurer une collecte continue et fiable des données.
- Collaborer avec les équipes produit et les professionnels de santé pour comprendre les spécificités des sources de données.

### 2- Construction et optimisation des requêtes dans l'entrepôt de données

- Écrire des requêtes performantes pour extraire des données pertinentes et répondre aux besoins des équipes produit, data science et recherche. (SQL, Opensearch/Elastic Search)
- Concevoir des vues et des modèles de données adaptés aux cas d'usage (analyses cliniques, tableaux de bord, études épidémiologiques, etc.).
- Contribuer à l'amélioration continue de la structure et de la performance de l'entrepôt de données.
- Optimiser les pipelines existants pour automatiser les requêtes entre APIs et l'entrepôt de données.

### 3- Enrichissement et structuration de données

- Concevoir et développer des pipelines d'extraction d'information dans les textes (NLP)
- Concevoir et développer la standardisation et l'enrichissement des données extraites à l'aide de terminologies médicales et de bases de connaissances.

### 4- Qualité et sécurité des données

- Participer à la mise en place de standards de qualité pour garantir l'intégrité, la complétude et la cohérence des données.
- Appliquer et respecter les normes RGPD et les réglementations en matière de traitement des données de santé.

### Profil recherché :

#### Hard skills :

- ETL/ELT (dbt, Dagster, Perfect)
- SQL et PostgreSQL
- CSV, JSON, HL7, FHIR (format d'interopérabilité spécifiques à la santé)
- Data Warehousing et architectures cloud
- Python
- Machine learning, Deep Learning
- LLM, LangChain, LangGraph

#### Soft skills :

- Esprit analytique et rigueur dans le traitement des données.
- Curiosité et esprit d'innovation.
- Capacité à travailler dans un environnement évolutif.
- Capacité à collaborer avec des équipes multidisciplinaires (développeurs, data scientists, médecins).
- Sensibilité aux enjeux de confidentialité et d'éthique des données de santé.

#### Expérience :

- Diplôme en informatique, ingénierie, ou domaine lié aux données de santé, niveau bac+5 minimum (un PhD est un atout).
- 2 à 5 ans d'expérience sur des projets similaires.
- Une expérience dans le secteur de la santé ou des données médicales.

#### Conditions de travail :

- Type de contrat : CDI
- Localisation : Lyon
- Rémunération : selon notre grille salariale + de nombreux avantages

Merci d'envoyer vos candidatures à l'adresse email [recrutement@stane-groupe.fr](mailto:recrutement@stane-groupe.fr) avec l'intitulé du poste en objet : **Data Scientist**

Nous sommes une entreprise qui valorise la diversité et l'inclusion. Nous encourageons les candidatures de personnes talentueuses, quel que soit leur handicap, et nous nous engageons à créer un environnement de travail accessible et inclusif pour tous.