

Data scientist pour la R&D 2.0 !

15/09/2021 – 15/09/2022, DUREE 12 MOIS

ORGANISME ET EQUIPE D'ACCUEIL

La Mass Transit Academy (MTA) promeut la culture Mass Transit au sein de SNCF Transilien. Le Mass Transit, c'est l'exploitation des trains en zone dense (beaucoup de voyageurs : 3.4 millions/jour et beaucoup de trains : plus de 6000 trains/jour). Le Lab' MTA est le laboratoire d'innovation interne de la MTA qui a pour mission de valoriser les données de comptage grâce à des projets innovants. Ces 4 dernières années, le Lab' a bâti son expertise sur l'exploitation conjointe des données de flux voyageurs et de flux de trains.

MISSIONS CONFIEES

Vous serez responsable des projets en cours du Pôle Data du Lab' MTA, et vous assurerez le portage de nouveaux projets ainsi que la valorisation des travaux effectués. Votre mission consistera à participer à la conception et au pilotage de la stratégie de l'équipe data. Plus concrètement, vous serez amené-e à :

1. Manipuler et fouiller les données complexes issues des nouvelles rames compteuses afin d'en tirer le meilleur parti (manipulation, visualisation, modélisation, optimisation, ...);
2. Développer dans le cadre d'une nouvelle PoC, une application Power BI alimentée par des données de comptages issues des capteurs au niveau des portes des trains en Île-de-France. Le travail consistera à développer de nouvelles fonctionnalités en accord avec les utilisateurs et accompagner l'industrialisation de ce dashboard aussi bien en termes d'architecture que de développement de graphiques et/ou d'indicateurs ;
3. Assurer le suivi du PoC [Hector](#), un outil Web développé en Django, en monitorant le fonctionnement de l'outil et en développant de nouvelles fonctionnalités éventuelles (ex : prévision de la charge à bord le long d'une course).

La recherche et les projets from scratch sont l'ADN du Lab' Mass Transit Academy, vous serez ainsi au cœur d'une mission de production de nouvelles idées/solutions pour le futur de Transilien. Il s'agira également de répondre à des sollicitations rapides d'autres entités pour analyser leurs données.

DESCRIPTION DU PROFIL RECHERCHE

Vous souhaitez manipuler des petits et gros volumes de données concernant les voyageurs et les trains d'Île-de-France. Vous êtes issu·e d'un cursus d'ingénieur / universitaire avec une dominante en statistiques et développement informatique, et vous avez plus de 2 ans d'expérience dans le domaine. Vous aimez proposer de nouvelles idées et porter des projets autour des données. Vous êtes aussi capables de développer des outils de A à Z en mode PoC, de la conception de l'architecture à la visualisation d'indicateurs en passant par la connexion et le traitement de bases de données. Pour cela :

- Vous maîtrisez les langages de programmation **R, Python** ;
- Vous avez déjà mené un projet **Django** et **PowerBI** ;
- Vous n'avez pas peur de construire des diagrammes UML, et n'êtes pas rebuté·e par un code CSS ou HTML ;
- Vous connaissez Azure Databricks et savez manipuler des grandes bases de données ;
- Connaître les fondamentaux de l'apprentissage statistique (régression linéaire, forêts aléatoires, ...) est un plus.

Vous êtes autonome dans votre travail, mais savez être pédagogue et curieux·e ce qui vous permet d'interagir facilement avec des clients internes pour construire avec eux un outil ou un projet.

Les données de Transilien sont tellement riches que nous avons besoin d'un bricoleur ou d'une bricoleuse des données hors pair !

CONTRAT : mission d'intérim

GRATIFICATION : 30 – 45k brut

LIEU : 10 rue Camille Moke Saint Denis 93220 + télétravail selon la situation sanitaire

CONTACT

Merci de faire parvenir CV et lettre de motivation par courriel : marc.deruelle@sncf.fr