



Recherche

## Stagiaire Data Scientist

Disponibilité : 2020

Filiale à 100 % du Groupe ENSAE, DATASTORM ([www.datastorm.fr](http://www.datastorm.fr)) a été créé en 2013 face à une demande grandissante d'expertise dans le domaine de la statistique, de l'économie, de la Data Science et de l'intelligence artificielle. Installés sur le campus de l'Institut Polytechnique de Paris du Plateau de Saclay, nous sommes au cœur d'un pôle majeur d'innovation et de recherche en intelligence artificielle. DATASTORM intervient sur ses projets en collaboration avec les chercheurs et enseignants des différentes écoles du plateau.

Avec le souhait de rester agile, DATASTORM privilégie une organisation très à plat avec un regroupement de l'ensemble des équipes opérationnelles au sein d'une direction unique pilotée par le directeur général et 2 directeurs adjoints. L'effectif, outre ces 3 directeurs, se compose à date de 5 Data Scientists Seniors et de 5 Data Scientists Juniors.

Notre culture de l'innovation se traduit directement dans notre organisation interne puisque nous avons choisi de laisser à chaque collaborateur de l'entreprise 20% de son temps pour travailler sur des sujets de R&D et de veille : développement de packages open-source produits, participation à des compétitions de data science, conférences et formations. Ceci permet à nos collaborateurs d'être à la pointe des outils et méthodes mis en œuvre. DATASTORM dispose d'ailleurs de son propre dépôt ouvert sur GitHub (<https://github.com/datastorm-open>), avec des contributions à l'éco-système open source (packages R et python).

DATASTORM a développé un portefeuille client très diversifié, notamment dans les secteurs de l'énergie, des médias et de la banque et assurance.



Les locaux de DATASTORM se situent au sein de l'ENSAE, sur le plateau de Saclay au 5 Avenue Le Chatelier, 91120 Palaiseau. Nous disposons également de bureaux de passage à Malakoff. De plus, une partie des études peut nécessiter des déplacements sur les sites de nos clients, la plupart en région parisienne.

## Le stage

Pour capitaliser les savoirs faire acquis pour délivrer les flux de traitement de données, les modèles de Machine Learning optimisés et les datavisualisations attendus par nos clients, nous développons un ensemble d'outils packagés internes (framework) et avons mis en place dans ce cadre une organisation agile formalisée :

- Un fonctionnement par récits utilisateurs et tickets via un board de suivi mural type Kanban ;
- Des revues de code systématiques réalisées par des pairs ;
- Des sprints de 2 semaines à l'issue desquels une revue de sprint est faite avec l'ensemble du personnel Datastorm pour partager les éléments produits et prioriser les tickets à traiter.

De ce fait, Datastorm peut vous proposer une **offre de stage originale** :

1. Ce stage sera aussi l'opportunité pour vous d'être **immergé dans notre organisation agile** :
  - Dans un premier temps, après avoir convenu avec votre maître de stage du sujet et des attendus vous participerez à sa « traduction » sous la forme d'un récit utilisateur (« story ») puis de tickets (« issues ») ;
  - Dans un deuxième temps vous traiterez les tickets qui vous auront été affecté en développant les scripts R ou Python nécessaires ; Ces développements seront réalisés dans un environnement GIT ;
  - Enfin vous exercerez votre esprit critique et votre créativité en vous impliquant dans les revues de sprint
2. Parce que votre cursus et vos attentes vous sont propres et que nous disposons de multiples dépôts de données mis à disposition par nos clients, nous vous proposons de **convenir ensemble de votre sujet de stage** ;
3. Vous travaillerez en **partenariat avec nos chercheurs et d'autres datascientists** de Datastorm.

Ce stage est indemnisé.

## Votre profil

- Master I ou II (école d'ingénieur ou université), avec une spécialisation en Statistiques, Mathématiques appliquées, Computing Science avec une forte dimension Data Science ou Data Ingénierie ;
- Familier des langages R ou Python ;
- Bonnes pratiques de développement (GIT, tests) et connaissance des environnements de calculs intensifs ;
- A l'écoute du monde open source, des nouvelles technologies et des nouveaux algorithmes ;
- Rigueur dans le travail.

## Contact

Les candidat(e)s sont invité(e)s à adresser leur candidature (CV, lettre de motivation) par mail à l'adresse suivante : [contact@datastorm.fr](mailto:contact@datastorm.fr)