



Mediametrie

***Vous êtes passionné par l'univers des médias ? Rejoignez-nous!***

Leader des études médias, Médiamétrie observe, mesure et analyse les comportements du public et les tendances du marché. Face au développement des technologies numériques et des nouveaux moyens de communication, elle doit adapter en permanence ses dispositifs de mesure d'audience.

Dans ce contexte, le **Data Lab** de Médiamétrie est une équipe pluridisciplinaire dédiée à la R&D sur les méthodes et outils d'analyse des données massives. Au sein de cette équipe, nous recherchons, en stage pour 6 mois, sur notre site de Sophia-Antipolis un :

**Assistant Data scientist H/F**

Vous travaillerez sur **l'analyse et l'apprentissage statistique des flux de données.**

Confronté à des volumes de données de plus en plus importants ainsi qu'à la nécessité de réaliser des traitements avec des latences de plus en plus faibles, Médiamétrie veut poursuivre ses efforts et accroître ses compétences dans le traitement des flux de données, comme par exemple des données produites par la mesure du streaming vidéo en France. En effet, le traitement des flux de données pose questions, à la fois pour ce qui est des aspects technologiques mais également en terme méthodologique et algorithmique.

Parmi les technologies en vue, [Apache Flink](#)<sup>1</sup> semble apporter des réponses intéressantes en offrant la possibilité de traiter à la fois des flux et des lots de données avec un unique moteur d'exécution. Ce stage propose d'étudier cette technologie, de l'appliquer dans une contexte métier et de la comparer aux solutions déjà existantes à Médiamétrie.

Au-delà, des problématiques technologiques, les flux de données amènent également des contraintes qui rendent inefficaces voire impossibles les méthodologies d'analyse et d'apprentissage classiques. Ce stage a également pour objectif d'appréhender les difficultés inhérentes à la nature des flux de données et d'envisager des solutions techniques et algorithmiques pour les surmonter.

**Vos missions seront les suivantes :**

- Prendre en main de la technologie [Apache Flink](#)<sup>1</sup>
- Se familiariser avec l'écosystème de services Amazon [AWS](#)<sup>2</sup>, e.g. EMR et Kinesis
- Réaliser une implémentation CEP et une implémentation batch d'un traitement métier déjà connu chez Médiamétrie, comme par exemple le calcul du top-k des vidéos les plus regardées
- Rédiger un état de l'art sur le traitement et l'apprentissage statistique des flux de données
- Envisager des solutions techniques et algorithmiques aux problèmes posés par les flux de données (e.g. complexité en temps et en espace, flux non-stationnaires)

1- <http://flink.apache.org/>



Pour mener à bien ces missions, **vous êtes en formation informatique/statistique/data science BAC+4/+5**, et vous avez des connaissances en programmation informatique et en apprentissage statistique. Vous vous intéressez aux technologies et aux méthodes de traitement des données massives. Enthousiasme et curiosité sont vos atouts pour ce stage. **Merci d'envoyer votre candidature via l'adresse suivante : [mdefoin-platel@mediametrie.fr](mailto:mdefoin-platel@mediametrie.fr)** / Mission basée à Sophia-Antipolis (06).