

## Dataiku DSS - Les fondamentaux

### Description

*Deux jours intensifs pour découvrir le logiciel DSS de Dataiku, la manipulation de données, la création de modèles jusqu'à la mise en production de modèles. Après une introduction à l'outil, vous apprendrez à maîtriser les flux de traitements disponibles dans Dataiku DSS. Basée sur de nombreux exercices pratiques, cette formation Dataiku vous permet d'être immédiatement opérationnel pour le traitement de vos données grâce à Dataiku.*

### Durée

2.00 jours (14.00 heures)

### Profils des stagiaires et pré-requis

- Analystes
- Data scientists
- Experts business intelligence
- Pas de pré-requis

### Objectifs pédagogiques

- Maîtriser l'interface de DSS Dataiku et de son utilisation
- Construire des projets et des flux de manipulation de données
- Construire un modèle de machine learning
- Compréhension de la mise en production des modèles

### Programme détaillé

- Introduction à Data Science Studio (DSS)
  - Qu'est-ce que DSS Dataiku ?
  - Quels usages et quels utilisateurs ?
  - DSS dans le paysage des plateformes data science
  - Comment installer DSS ?
- L'interface DSS
  - Découverte de l'interface
  - La notion de projet
  - La notion de flux
  - Travailler dans le Lab
  - Comment travailler en équipe avec Dataiku (Le partage, le versionnement)
- La manipulation des données avec DSS
  - Principe de la préparation des données
  - Importer des données de différents types (csv, xls, sql...)

Stat4decision

37-39 avenue Ledru Rollin Paris 75012

Tel. 01.72.25.40.82 | E-mail : [info@stat4decision.com](mailto:info@stat4decision.com) | [www.stat4decision.com](http://www.stat4decision.com)

Numéro SIRET: 81048985600015 | Numéro de déclaration d'activité: 11755352275 (auprès du préfet de région de: 75)

- Appliquer des recettes pour transformer des données (Transformation de données, gestion des types, manipulation des données textuelles, jointure de fichiers)
- Ateliers sur des données
- La data visualisation avec DSS
  - Quelques graphiques simples avec Dataluku DSS
  - Présentation de visualisations plus avancées
  - Ateliers sur des données réelles (Manipulation de données pour afficher des séries temporelles, affichage de cartographie avec DSS)
  - Le développement et le partage de tableaux de bord (Dashboard) avec DSS
- Introduction au machine learning : présentation des principes de l'apprentissage automatique
  - Construction d'un premier modèle avec Dataluku
  - Les autres modèles disponibles dans DSS
  - Atelier sur des données (préparation et construction de modèles, ajustement des hyper-paramètres)
- La mise en production d'un projet DSS
  - Les bundles
  - Les automation node et les API node
  - Les scénarios
  - Le monitoring
- Aller plus loin :
  - L'intégration de code en Python ou en R dans Dataluku DSS (les Notebooks)
  - L'utilisation d'infrastructures Big Data (Hadoop / Spark)
  - L'intégration de code en Python ou en R dans Dataluku DSS (Utilisation de Notebook et de recettes sous forme de code)
  - Le deep learning dans DSS

## Organisation de la formation

### Moyens pédagogiques et techniques

- Documents supports de formation projetés.
- Exposés théoriques
- Etude de cas concrets
- Mise à disposition en ligne de documents supports à la suite de la formation.

### Dispositif de suivi de l'exécution de d'évaluation des résultats de la formation

- Feuilles de présence.
- Mises en situation.
- Formulaire d'évaluation de la formation.