

## **Dataviz : solutions opensource**

**Durée:** 1 jour

640 €

15 mars

21 juin

11 octobre

13 décembre

### **Public:**

Chefs de projet, architectes, développeurs, data-scientists souhaitant mettre en oeuvre des solutions d'analyse et de visualisation des données

### **Objectifs:**

Savoir quelles sont les solutions et outils opensource d'analyse et de restitution de données, connaître les spécificités, contraintes et apports des différentes solutions.

### **Connaissances préalables nécessaires:**

Connaissances générales sur le Bigdata, le data-mining, l'analyse de données.

### **Programme:**

#### **Introduction**

: Le besoin : types de données, exemples de démarches et d'analyse  
Définitions : apprentissage supervisé, apprentissage automatique  
Arbres de décision, de régression, régression automatique  
Classifieurs. Scoring  
Fonctionnalités des outils de dataviz : analyses statistiques, classifications, rapprochements, production de recommandations, représentations graphiques,  
Présentation de quelques outils : Mahout, Giraph, Agile, spagobi

#### **Mahout**

: Présentation Mahout.  
Positionnement dans l'offre BigData et Machine Learning : Hadoop, Spark,..  
Fonctionnalités. Mode autonome ou mode distribué  
Exemples d'algorithmes fournis avec Mahout.

#### **Giraph**

: Principe du projet Giraph. Algorithmes de graphes. Infrastructure Hadoop. Cas d'utilisation.

#### **spagoBI**

: Historique de spagoBI, positionnement, fonctionnalités.  
Installation et démonstrations depuis l'interface web. Exemples d'analyse avec Cockpit.