

Dataviz : solutions opensource

Durée: 2 jours

1320 €

15 au 16 mars
4 au 5 juin

25 au 26 octobre
13 au 14 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 : Fonctionnalités des outils de dataviz : analyses statistiques, classifications, rapprochements, production de recommandations, représentations graphiques, Présentation de quelques outils : Mahout, Giraph, Agile, spagobi, Neo4j, Spark GraphX, Gephi
Positionnement et fonctionnalités.

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.

Dataviz : solutions opensource

- Neo4j** : Présentation du produit. Requêtage avec le Cypher Query Language
Travaux pratiques : création d'un graphe,
Requêtes de recherche, navigation dans le graphe
- Spark GraphX** : Fourniture d'algorithmes, d'opérateurs simples pour des calculs statistiques sur les graphes
Travaux pratiques :
exemples d'opérations sur les graphes.
- Gephi** : Mise en oeuvre. Gestion des graphes. Positionnement des objets.
Import de données.
Travaux pratiques :
Mise en oeuvre avec Spark/GraphX