



CATALOGUE TRIMANE



FORMATION

01. Concepts Généraux

02. Statistiques et Data Science

03. Valorisation des données

04. Technologies BI et Big Data

05. Transformation Data Driven



TRIMANE
THE DATA INTELLIGENCE
COMPANY

A PART OF THE BLOCKCHAIN GROUP

TOUR W, 102 TERRASSE BOIELDIEU,
92 800 PUTEAUX – LA DEFENSE

www.trimane.fr

SOMMAIRE

1. CONCEPTS GENERAUX	4
1.1 NIVEAU 1	4
1.1.1 <i>Initiation au décisionnel</i>	4
1.1.2 <i>Big Data : Sensibilisation et nouvelle approche</i>	5
1.1.3 <i>Big Data : Architecture et technologies</i>	5
1.2 NIVEAU 2	6
1.2.1 <i>Modélisation décisionnelle</i>	7
1.2.2 <i>Big Data : de l'expérimentation à la mise en production</i>	8
1.2.3 <i>Initiation à Spark</i>	10
1.3 NIVEAU 3	11
1.3.1 <i>Optimisation SID</i>	11
2. STATISTIQUES / DATA SCIENCE	12
2.1 NIVEAU 1	12
2.2.1 <i>Big Data : utilisation de R</i>	12
2.2.2 <i>Python / Machine Learning</i>	13
2.3 NIVEAU 2	14
2.3.1 <i>Chaîne SAS</i>	14
2.3.2 <i>Python Scientifique</i>	14
2.4 NIVEAU 3	15
2.4.1 <i>Python avancé / Machine Learning</i>	15
3. VALORISATION DES DONNEES / DATA VIZ / REPORTING	16
3.1 NIVEAU 1	16
3.1.1 <i>Power BI</i>	16
3.1.2 <i>Tableau Software Desktop Analystes</i>	17
3.1.3 <i>SAP Business Objects</i>	19
3.2 NIVEAU 2	21
3.2.1 <i>Power BI – Perfectionnement</i>	21
3.2.2 <i>Tableau Software Desktop Developpers</i>	22
3.2.3 <i>Tableau Software Serveur</i>	23
3.2.4 <i>SAP Business Objects perfectionnement</i>	24
4. TECHNOLOGIES BI ET BIG DATA	25
4.1 NIVEAU 1	25
4.1.1 <i>Développement Python</i>	25
4.1.2 <i>PL/SQL : Les fondamentaux</i>	26
4.1.3 <i>Bases No SQL & Mongo DB</i>	27
4.1.4 <i>Ecosystème Hadoop</i>	28
4.2 NIVEAU 2	29
4.2.1 <i>Developpement Python Intermédiaire</i>	29
4.2.2 <i>Talend Open studio</i>	30
4.2.3 <i>MAP R : Architecture et Administration</i>	31
4.2.4 <i>Cloudera / Hortonworks : Architecture et Admin</i>	32
4.3 NIVEAU 3	33
4.3.1 <i>Talend Avancé</i>	33
5. DATA DRIVEN	34
05.1 NIVEAU 1	34
5.1.1 <i>Economie numérique et utilisation des données</i>	34

Fondée en **Avril 2005**, TRIMANE est une ESN française, spécialisée en Data Intelligence, elle conseille, forme et collabore avec ses clients depuis plus 15 ans, dans le développement et la réussite de projets Data et intervient dans des secteurs d'activité variés : administration public, santé, industrie, finance/banque/assurance, énergie....

Notre offre Data est dimensionnée pour accompagner nos clients sur l'ensemble du cycle de vie d'un projet de la phase d'étude, de réalisation / maintenance et la formation des équipes internes.

Notre offre de formation est construite de manière à proposer des programmes adaptés à tous type de public en passant par des sessions de sensibilisation et d'acculturation des équipes métiers aux programmes des plus techniques destinées aux profils plus avertis qui souhaitent maîtriser les technologies émergentes.

En tant qu'organisme de formation agréé, nous proposons les prestations suivantes :

- 1** **Formation à la demande** basées sur des besoins spécifiques. Nous proposons de construire des programmes de formation adapté au niveau de maturité des organiser et à leur contexte. Nous sensibilisations les équipes internes à la mise en place d'une culture Data en passant par les utilisateurs finaux, le top management, les référents, les équipes techniques et les équipes supports.

Nos formations à la demande sont hybrides est peuvent être mise en place selon la définition de formation classique (théorique + pratique) en groupe ou individuel, en mode coaching ou en tutorat.

- 2** **Formation au catalogue** sur l'ensemble des thématiques suivantes :
- 01. Concepts Généraux
 - 02. Statistiques et Data Science
 - 03. Valorisation des données
 - 04. Technologiques BI et Big Data
 - 05. Transformation Data Driven

Plusieurs niveaux de formation sont proposés du Niveau 1 (Débutant) au Niveau 3 (Expert).

1. CONCEPTS GENERAUX

1.1 NIVEAU 1

1.1.1 INITIATION AU DECISIONNEL

Concepts Généraux

INITIATION AU DECISIONNEL

★ Avis : 9/10

01.1.1

1 jour (7h)

BUT

Cette formation d'initiation aux systèmes décisionnels aborde les sujets suivants :

- Les caractéristiques de la chaîne décisionnelle,
- Les bases de données,
- La modélisation décisionnelle,
- Les différentes étapes d'un projet décisionnel

COMPETENCES VISEES

Acquérir une base de connaissances en modélisation et bases de données

PRE-REQUIS :

Aucuns

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Informaticiens, consultants, chefs de projet non décisionnels, désireux de découvrir la BI ainsi que les bases de données

ANIMATEUR

Consultant BI, société TRIMANE

SOMMAIRE

Présentation du décisionnel et de ses enjeux

La chaîne du décisionnel : les différentes briques

Les sources de données

Outil d'alimentation ETL

Définition

La collecte des données

La transformation des données

Le chargement des données

La gestion des rejets

Questions/réponses – Choix d'un ETL dédié

L'entrepôt de données (datawarehouse)

Définition

La structure des données

Son rôle

Questions/réponses

Les magasins métiers (datamarts)

Définition

Les avantages et les inconvénients

Questions/réponses

Analyse multidimensionnelle : reporting et OLAP

La notion d'hypercube

Datamining

Définition

Les quatre couches de l'architecture décisionnelle

Définition des 4 couches

La collecte

L'intégration

L'organisation

La présentation

Exemple

Le projet décisionnel

Les différentes étapes

L'étude de cadrage

Spécifications et conception

La mise en œuvre

La phase de recette

Exemple

Etudes de cas

Réalisation d'un modèle en étoile
Les modèles en étoiles multiples et en flocon

1.1.2 BIG DATA : SENSIBILISATION ET NOUVELLE APPROCHE

Concepts Généraux

BIG DATA : SENSIBILISATION ET NOUVELLE APPROCHE

01.1.2

★ Avis : 8/10

2 jour (14h)

BUT

Sensibilisation au Big Data : Qu'est-ce que la Big Data ? Où se situent-elles aujourd'hui ? Démistification de cette nouvelle technologie et des apports qu'elle peut procurer.
Comprendre les différents types de modélisation, les avantages et inconvénients de chacun, les outils existant et l'utilisation faite.
Comprendre ce qu'il y a derrière cette nouvelle technologie.

COMPETENCES VISEES

Sensibilisation et initiation à l'écosystème Big Data, cette formation est une première approche destinée à l'ensemble des entités du groupe manipulant des données (techniques, métiers, Chef de projet, Responsable SI...)

PRE-REQUIS :

Bonne connaissance de SQL et modélisation de base de données.

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...
Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Développeurs, équipes métiers, top management

ANIMATEUR

Consultant BIG DATA, société TRIMANE

SOMMAIRE

SENSIBILISATION A LA BIG DATA

Qu'est-ce que la Big Data ?

Introduction : Comprendre simplement ce qu'est la Big data
Informatique connecté, objets intelligents : Pourquoi a-t-on besoin de la Big Data ?
Big Data : Historique, définition, enjeux et chiffres clés
Comment les grands acteurs utilisent la Big Data

Les limites et l'avenir de la Big Data

Infrastructure, compétences
Les barrières morales et psychologique : La vision du Big Data par l'extérieur
Juridiques
Evolution logicielle

UNE NOUVELLE APPROCHE

Les différents types de modélisation et outils utilisés

Besoin d'une nouvelle modélisation
Quand utiliser le Paradigme Clé/Valeur : Avantages, inconvénients et outils
Quand utiliser les Bases documentaires : Avantages, inconvénients et outils
Quand utiliser les Bases orientés colonnes : Avantages, inconvénients et outils
Quand utiliser le Paradigme Graph : Avantages, inconvénients et outils

Les nouvelles pratiques

Les principes de multi-nœud
Qu'est-ce que le DFS ?
Le principe et exemple de MapReduce
Le « requêtage »

Les outils de la Business Intelligence dans tout ça ?

La réponse et l'adaptation des outils de BI
Les ETL
Le Reporting

1.1.3 BIG DATA : ARCHITECTURE ET TECHNOLOGIES

Concepts Généraux

BIG DATA : ARCHITECTURE ET TECHNOLOGIES

01.1.3

2 jour (14h)

★ Avis : 7.5/10

BUT

Cette formation Bigdata : architecture et technologies vous permettra de comprendre les concepts du BigData et savoir quelles sont les technologies implémentées.

COMPETENCES VISEES

Approfondissement sur les technologies du Big Data et maîtriser son écosystème (quelles technologies à implémenter ?)

PRE-REQUIS :

Bonne culture générale sur les systèmes d'information

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Chefs de projets, architectes, et toute personne souhaitant connaître les outils et solutions pour mettre en place une architecture BigData.

ANIMATEUR

Consultant BIG DATA, société TRIMANE

SOMMAIRE

Introduction

Le besoin : volumes importants de données, traitements optimisés de flux de données au fil de l'eau, liés aux nouvelles technologies et aux nouveaux usages

Domaines concernés : recherche scientifique, médical, e-commerce, sécurité, ...

Développement des techniques sur différents aspects :

stockage, indexation/recherche, calcul

Définition ETL : Extract Transform Load

Les acteurs

Stockage

Caractéristiques NoSQL : Structure de données proches des utilisateurs, développeurs

Données structurées et non structurées, documents, images, fichiers XML, JSON, CSV, ...

Les différents modes et formats de stockage.

Stockage réparti : réplication, hachage,...

Systèmes de fichiers distribués : GFS, HDFS, ...

Les bases de données

Quelques exemples de produits et leurs caractéristiques :

cassandra, MongoDB, CouchDB, DynamoDB

Indexation et recherche

Moteurs de recherche

Principe de fonctionnement

Méthodes d'indexation

Exemple de Lucene, et mise en œuvre avec solr

Recherche dans les bases de volumes importants : exemples de produits et comparaison : dremel, drill, elasticsearch,

MapReduce

Calcul et restitution, intégration

Différentes solutions : calculs en mode batch, ou en temps réel, sur des flux de données ou des données statiques

Les produits : langage de calculs statistiques, R Statistics

Language

Outils de calcul sur des volumes importants : storm en temps réel, hadoop en mode batch

Zoom sur Hadoop : complémentarité de HDMS et MapReduce

Evolutions

Les offres Saas BigData comme Google BigQuery

Les limites. Les nouveautés annoncées

1.2 NIVEAU 2

1.2.1 MODELISATION DECISIONNELLE

Concepts Généraux

MODELISATION DECISIONNELLE

★ Avis : 9/10

01.2.1

2 jours (14h)

BUT

Cette formation Big Data a pour objectif d'appréhender les spécificités d'un Datawarehouse, son vocabulaire dédié et en quoi il répond mieux aux besoins d'analyse des données formulés par les utilisateurs qu'une base de données relationnelle « classique ».

COMPETENCES VISEES

Savoir concevoir progressivement un datawarehouse

PRE-REQUIS :

Connaître les concepts d'une base de données relationnelle

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Informaticiens, consultants, chefs de projet non décisionnels, désireux de découvrir la BI ainsi que les bases de données.

ANIMATEUR

Consultant BI, société TRIMANE

SOMMAIRE**Vocabulaire et concepts**

Le rôle d'un Datawarehouse en entreprise
Le Datamart (ou magasin de données)
Les axes de l'analyse (dimensions)
Le rôle de l'analyse (table de faits)

La démarche de modélisation en étoiles

Proposer un schéma accessible aux utilisateurs
Mettre en relation les dimensions et les faits (modèle en étoile)
Hiérarchiser les dimensions et leurs liens avec les faits (modèle en flocon)
Définition d'une constellation
Les différences fondamentales entre le modèle en étoile et les modèles traditionnels
Passage d'un modèle relationnel à un modèle décisionnel
La dénormalisation
Compatibilité entre les différents modèles
Quelles différences concrètes pour les utilisateurs des outils BI ?
Exemple de modélisation en étoile

Conception d'un Datawarehouse

Les différentes approches (top-down, bottom-up, middle-out)
Liaison avec les éléments précédents (étoile, constellation...)
Les raisons d'une mauvaise conception (dimension pas assez générique par exemple)
Les grandes lignes pour valider la conception
La documentation du Datawarehouse

Travaux pratiques

Conception pas à pas d'un Datawarehouse à travers 3 applications traitant de

- *Méthodologie de construction d'un datawarehouse, les 4 étapes*
 - *Sélectionner le processus à modéliser*
 - *Choisir le grain de données*
 - *Choisir les dimensions*
 - *Identifier les faits*
- *Le remplissage des dimensions*
- *Table de faits simple contre table de faits multiple*

1.2.2 BIG DATA : DE L'EXPERIMENTATION A LA MISE EN PRODUCTION

Concepts Généraux

BIG DATA EXPERIMENTATION

★ Avis : 7/10

01.2.2

2 jours (14h)

BUT

- Cette formation Big Data aborde les sujets suivants :
- L'origine du Big Data,
 - Vulgarisation et décryptage de l'impact du Big Data,
 - Les opportunités du Big Data,
 - Aspects juridiques : quelles données pour quels usages ?
 - L'entreprise face au défi du Big Data,
 - Présentation de l'écosystème Hadoop et son architecture
 - Data Science et analyse des données
 - Les métiers du Big Data
 - Les méthodologies projet
 - Exemples de cas d'usage
 - Présentation d'un exemple concret de mise en œuvre

COMPETENCES VISEES

Sensibilisation et initiation à l'écosystème Big Data, cette formation est une première approche destinée à l'ensemble des entités du groupe manipulant des données (techniques, métiers, Chef de projet, Responsable SI...)

PRE-REQUIS :

Bonne culture générale sur les systèmes d'information

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Développeurs, équipes métiers, top management

ANIMATEUR

Consultant BI, société TRIMANE

SOMMAIRE

L'origine du Big Data

Contexte et fondements
Le Système d'Information 2.0 (S.I. 2.0.)

Vulgarisation et décryptage de l'impact du Big Data

Quelles applications concrètes pour le Big Data
Ou se situe la Big Data ?

Les opportunités pour les directions informatiques et les métiers

La DSI au centre du jeu
Les nouveaux usages du Big Data
L'approche Big Analytics
Créer de la valeur à partir des données

Aspects juridiques et éthiques : quelles données pour quels usages ?

Quels types de données peut-on traiter (publiques, privées, personnelles, anonymes, autre...) ?
A qui appartiennent les données (à la personne concernée, l'entreprise qui les revend ou celle qui les collecte)?
Existe-t-il un règlement européen concernant l'usage de ces données et la protection du citoyen ?
Ces règles s'appliquent-elles à l'Open Data ?
Big Data ou Big Brother?
Impact global sur la vie privée
Les enjeux juridiques internationaux

L'entreprise face au défi du Big Data

Intégration de données internes et externes à l'entreprise au sein d'un gisement de données Big Data
Nature des données (structurées et non structurées)
Valorisation de ces nouvelles données
La datavisualisation
Cette nouvelle science qui permet de présenter les données de façon visuelle
Les outils de présentation des données analytiques
Les outils de datamining et de modélisation statistique
Les outils de datascience

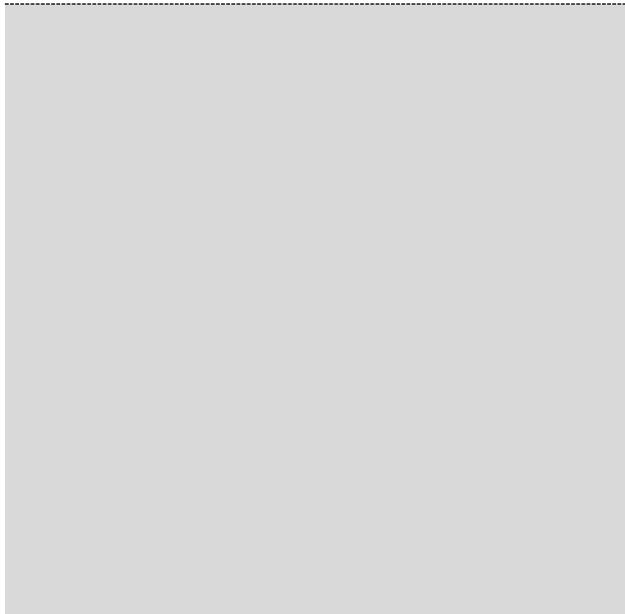
Présentation de l'écosystème Hadoop
Présentation des technologies au cœur du Big Data
Les différents acteurs du monde Hadoop
Quelle urbanisation au sein du système d'information ?

Les architectures Big Data

Comment positionner le Big Data face à l'existant ?
Quelles sont les possibilités offertes par le Big Data
Quels sont les contraintes techniques du Big Data ?
Quelles stratégies de conservation des données (chaudes, froides, "gelées") dans le temps ?
Exemples de mise en œuvre d'architectures Big Data

Data Scientist et analyse de données

Les différentes sources disponibles et leur structure



Les solutions de traitement des données structurées et non structurées
Définition de la Data Science
Les outils de Data Science et de Data discovery

Les métiers du Big Data

Administrateur
Développeur
Data Analyst
Administrateur Fonctionnel
Chief Data Officer

Quelle méthodologie projet pour le Big Data ?

Quel ROI pour les projets Big Data ?
Les étapes du projet Big Data
Quelle gouvernance pour les projets Big Data ?

Exemple concret de mise en œuvre

Définition d'une expression de besoin
Choix d'une distribution Hadoop
Mise en œuvre de la solution sur un Cloud Public
Exploitation des données

1.2.3 INITIATION A SPARK

Concepts Généraux**SPARK INTITATION**

★ Avis : 6.5/10

01.2.3**1 jour (7h)****BUT**

Cette formation a pour objectif de faire découvrir à l'apprenant Spark et son principe de fonctionnement. Ce dernier apprend en outre à procéder à la manipulation de données et à la programmation avec les RDD (Resilient Distributed Datasets).

COMPETENCES VISEES

Comprendre l'environnement Apache Spark
Savoir utiliser le package PySpark pour communiquer avec Spark

PRE-REQUIS :

Bonne connaissance en Big Data (notamment Hadoop, MapReduce...) et en langage Python

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...
Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Développeurs, équipes métiers, top management

ANIMATEUR

Formateur Data Science, société TRIMANE

SOMMAIRE**Introduction à Spark**

Présentation de Spark
L'environnement Spark : RDD, DataFrame, DataSet
Installation de la plateforme Spark (plateforme distribuée, en local et en cloud)

Spark pour la manipulation des données

Utilisation de SparkSQL et des DataFrames pour manipuler des données
Charger des données depuis Hadoop, depuis des fichiers csv...
Transformer des données (création de DataFrames, ajout de colonnes, filtres...)

Programmation avec les RDD (Résilient Distributed Datasets)

Rdds élémentaires
Rdds clé/valeur

1.3 NIVEAU 3

1.3.1 OPTIMISATION SID

Concepts Généraux**OPTIMISATION D'UN SYSTEME D'INFORMATION DECISIONNEL (SID)**

★ Avis : 9/10

01.3.1**1 jour (7h)****BUT**

Cette formation a pour de répondre aux questions suivantes :

- Comment organiser son architecture technique pour contenir les charges ?
- Comment organiser son Datawarehouse pour optimiser la volumétrie ?
- Comment améliorer les temps de réponse du reporting ?

COMPETENCES VISEES

Savoir optimiser un datawarehouse et de manière plus vaste un SID

PRE-REQUIS :

Savoir modéliser un datawarehouse, connaissances SQL nécessaires

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Informaticiens, consultants, chefs de projet non décisionnels, désireux de d'approfondir la BI

ANIMATEUR

Formateur BI, société TRIMANE

SOMMAIRE**Organisation de l'architecture physique**

Composantes d'un système d'information décisionnel
Dimensionnement des serveurs
Les avantages et pièges de la parallélisation
La gestion de la volumétrie
Prendre en compte les flux de données et goulots d'étranglement
Le Clustering
Mise en pratique des principes par résolution de problématiques courantes

Optimisation d'un Datawarehouse existant

Table d'agrégation
Utilisation des magasins de données
Le partitionnement : Oracle 10g et plus

- Les différents types de partitionnement (par intervalle, hachage, liste)
- Leurs cas d'utilisation

Les index

- Les différents types d'index
- « Best practice » sur les index

Utilisation du plan d'exécution pour organiser son optimisation

2. STATISTIQUES / DATA SCIENCE

2.1 NIVEAU 1

2.2.1 BIG DATA : UTILISATION DE R

Statistiques / Data Science

MACHINE LEARNING SOUS R

★ Avis : 6/10

001

2 Jours (14 h)

BUT

La formation a pour objectif de permettre à l'apprenant d'avoir une vision transverse des principaux algorithmes supervisés et non supervisés et de savoir mettre en œuvre les algorithmes dans R sur un problème de classification ou de régression (apprentissage supervisé d'une cible binaire ou continue).

COMPETENCES VISEES

Maîtrise des algorithmes de Machine Learning,
Mise en œuvre des process d'apprentissage sur des cas réels sous R.

PRE-REQUIS

Connaissance des bases en statistique et mathématique.
Connaissance d'un des langages suivants : Python, Matlab, R, Java

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...
Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Développeurs, Ingénieurs en traitement de données, Data Analyst.

ANIMATEUR

Formateur Data Science, société TRIMANE

SOMMAIRE

1. Présentation générale

Généralités.
Typologie d'apprentissage : apprentissage supervisé, apprentissage non supervisé, Apprentissage profond.
Cycle de vie d'un projet de Machine Learning.

1. L'apprentissage automatique

Définitions, vocabulaire, modèle vs algorithme.
Typologie : classification, regression, clustering, rules, anomaly, recommandation.
Libraries R, comparaison avec Python.
Framework caret.
Techniques de contrôle de l'apprentissage : split, grid-search / Cross Validation, métriques.
Techniques de pre-processing : imputation, normalisation.
Apprentissage sous le capot : Fonction de perte, descente de gradient.
Introduction au Deep-learning.

2. Apprentissage supervisé

Algorithmes supervisés : régression pénalisée, arbres, randomforest, boosting.
Applications et démonstrations.
Travaux pratiques : analyse prédictive.

3. Apprentissage non supervisé

Analyse factorielle et réduction des dimensions.
Application sur l'image.
Techniques de clustering.
Applications : clustering d'individus, réaffectation avec KNN.

Cas pratique : résoudre un problématique de Machine Learning avec R.

Statistiques / Data Science

PYTHON / MACHINE LEARNING

★ Avis : 7/10

02.2.2**2 Jours (14 h)****BUT**

Cette formation a pour objectif de rendre l'apprenant autonome sur Python, de lui permettre de comprendre, modifier et écrire le code pour les besoins les plus courants (manipulations de données, statistiques descriptives), mais aussi pour des requêtes avancées (calculs complexes, formules conditionnelles).

COMPETENCES VISEES

Comprendre les enjeux et les fondamentaux du Machine Learning.
Concevoir et entrainer des modèles prédictifs adaptés à nos données d'apprentissage.
Évaluer, optimiser et améliorer les performances de modèles d'apprentissage automatique.

PRE-REQUIS :

Connaissance de base du langage Python (boucles for, conditions, types de variable, etc.). Public ayant déjà acquis de bonnes bases du langage Python (niveau intermédiaire) et débutant en Machine Learning

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...
Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Développeurs, ingénieurs en traitement de données, Data Scientist.

ANIMATEUR

Consultant Expert Python, société TRIMANE

SOMMAIRE**1. Présentation générale**

Généralités

Typologie d'apprentissage : apprentissage supervisé, apprentissage non supervisé, Apprentissage profond
Cycle de vie d'un projet de Machine Learning

2. La manipulation des données et l'exploration des données avec Python

Principales bibliothèques Python en sciences des données (Numpy, SciPy, Matplotlib, Pandas)

Maîtrise du calcul matriciel et de la manipulation de données
Analyse et visualisation des données

3. La bibliothèque Python (Scikit-Learn)

Comment utiliser la documentation ?

Intégration de Scikit-Learn avec d'autres bibliothèques (Pandas, Numpy, SciPy, Matplotlib, etc.)

Représentation des données par des tableaux (Numpy, Scipy, Pandas, Python)

Représentation d'une prédiction par une classe (prédicteur, classifieur, estimateur)

Comment choisir le bon algorithme d'apprentissage automatique ?

4. Les algorithmes de Machine Learning

Apprentissage supervisé VS non supervisé

Régression linéaire.

Classification.

Clustering.

Systèmes de recommandation.

Évaluation de la performance d'un modèle.

Optimisation paramétrique

Cas pratique : Résolution d'un problème de Machine Learning avec Python. Utilisation des implémentations mises à disposition par la bibliothèque *Scikit-Learn*.

1. Pipeline de préparation et nettoyage des données.

2. Apprentissage automatique supervisé.

3. Visualisation, analyse et interprétation des résultats.

2.3 NIVEAU 2

2.3.1 CHAINE SAS

Cette formation n'est actuellement pas disponible.

2.3.2 PYTHON SCIENTIFIQUE

Statistiques / Data Science

PYTHON SCIENTIFIQUE
 Avis : 7/10
02.3.2**5 Jours (35 h)****BUT**

- Utiliser les principales bibliothèques de calcul numérique dont Numpy, SciPy, Pandas et Matplotlib
- Paralléliser ses traitements sur des architectures modernes permettant le calcul distribué

COMPETENCES VISEES

- Pouvoir réaliser des calculs complexes au travers l'utilisation de bibliothèques spécifiques et le machine learning
- Pouvoir déployer des algorithmes de Machine Learning

PRE-REQUIS :

Niveau Technicien informatique
 Cette formation nécessite des compétences intermédiaires en programmation et calcul statistique

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...
 Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Informaticiens et ingénieurs travaillant dans les domaines des statistiques, la géomatique...

ANIMATEUR

Consultant Expert Python, société TRIMANE

SOMMAIRE**JOUR 1
Initiation**

- Introduction
- Pourquoi utiliser Python pour le calcul scientifique ?
- Les distributions Python pour les « data scientists »
- Les outils de développement : (Syntaxe de base, Structures de contrôles, Structures de données, Organisation du code...)
- Présentation et exercices d'application sur la bibliothèque Numpy et Pandas (calcul scientifique, manipulation des vecteurs et des matrices, manipulation de données).

**JOUR 2
Exploitation de données et visualisation**

- Présentation et exercices d'application sur la bibliothèque Matplotlib (visualisation des données).
- Présentation et exercices d'application sur Scipy
- Présentation et exercices d'application sur Seaborn
- Présentation et exercices d'application sur Scikit

**JOUR 3
Machine Learning**

- Initiation au Machine Learning
- Apprentissage supervisé
- Apprentissage non-supervisé

**JOUR 4
Application et étude de cas**

- Exercices d'application avec des algorithmes de Machine Learning
- Étude de cas (Machine Learning avec Python)

**JOUR 5
Solution d'optimisation**

- Étude de cas (suite)
- Parallélisme (Multi-procesus, Multithreading)
- Bilan et questions Apprentissage non-supervisé

2.4 NIVEAU 3

2.4.1 PYTHON AVANCE / MACHINE LEARNING

Statistiques / Data Science**PYTHON AVANCE - MACHINE LEARNING**

★ Avis : 7/10

02.4.1**2 Jours (14 h)****BUT**

Pouvoir réaliser des calculs complexes au travers l'utilisation de bibliothèques spécifiques et le machine learning.

COMPETENCES VISEES

Approfondir ses connaissances sur le Machine Learning avec Python

Être capable de :

Développer de puissants modèles prédictifs en Python

PRE-REQUIS :

Bonne connaissance du langage Python (niveau intermédiaire)
Compétences avancées en programmation et calcul statistiques/scientifique (exemple : mettre en œuvre les algorithmes dans Scikit-Learn sur un problème de classification ou de régression).

PROFILS DES STAGIAIRES

Développeurs, ingénieurs en statistiques, IA, traitement de données

Personnes souhaitant maîtriser et se perfectionner en Python sur le volet Machine Learning : Modéliser, explorer et classer ses données.

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

ANIMATEUR

Consultant Expert Python, société TRIMANE

SOMMAIRE**1. Choisir les algorithmes de machine learning**

Comprendre les enjeux

Tour d'horizon des principaux algorithmes.

Classification: k-Nearest Neighbors (k-NN).

Arbre de décision, Random Forest, XGBoost.

Régression : Régression logistique.

Clustering : K-Means, DBScan.

Les différentes méthodes de Scoring.

2. Concepts avancés

Validation croisée

Ensemble Machine Learning : cumuler les algorithmes pour une meilleure précision.

Automatiser les manipulations de données avec un pipeline.

3. Procédure d'entraînement et d'évaluation des algorithmes

Séparation du jeu de données : entraînement, test et validation.

Techniques de bootstrap (bagging).

Exemple de la validation croisée.

Définition d'une métrique de performance.

Descente de gradient stochastique (minimisation de la métrique).

Courbes ROC et de lift pour évaluer et comparer les algorithmes.

Matrice de confusion : faux positifs et faux négatifs.

4. Traiter les données en parallèle

Pourquoi paralléliser les traitements ?

Adapter les algorithmes.

Une complexité complémentaire.

Les frameworks de distribution à disposition : Spark et Dask.

5. Déployer en production

Intégrer un pipeline à une chaîne de déploiement automatisée (continuous delivery).

Packager un modèle : Predictive Model Markup Language.

Créer un endpoint REST avec python Flask.

Déployer dans le cloud.

Cas pratique : Evaluer la performance et automatiser un modèle de Machine Learning

3. VALORISATION DES DONNEES / DATA VIZ / REPORTING

3.1 NIVEAU 1

3.1.1 POWER BI

Valorisation des données / DataViz

POWER BI

★ Avis : 8/10

03.1.1

2 jours (14h)

BUT

Cette formation a pour objectif de fournir à l'utilisateur une interface simple et intuitive permettant de réaliser rapidement des tableaux de bord, dans le but de faciliter la compréhension des données par des représentations visuelles. Elle permet donc de rendre les utilisateurs autonomes dans la production des tableaux de bord, des analyses, dans l'exploration des informations, et dans le partage des données et des récits.

COMPETENCES VISEES

- Comprendre le fonctionnement des outils de datavisualisation
- Réaliser des tableaux de bord en quelques clics
- Produire des analyses de manière autonome
- Explorer et travailler sur des volumes de données conséquents
- Retrouver de manière simple des données cachées
- Partager ses explorations

PRE-REQUIS :

Connaissance tableur / Excel appréciées

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Développeurs, consultants informatiques

ANIMATEUR

Consultant Tableau, société TRIMANE

SOMMAIRE

Introduction:

- Présentation de l'offre BI Microsoft.
- La Suite Power BI: Power BI Desktop, le service Power BI et les applications Power BI Mobile

Acquisition et préparation des données :

- Extraire, transformer et charger des données dans Power BI.
- Utiliser des fichiers plats, Excel, bases de données relationnelles, sources SSAS, Web.
- Choisir les colonnes, choisir les lignes, filtrer, trier et supprimer les doublons

Enrichissement et préparation des données :

- Nettoyer et compléter les jeux de données.
- Fractionner les colonnes, les formater et définir le type de données, assembler et fusionner des tables.
- Créer des colonnes calculées. Empiler et fusionner des requêtes.

Définir le modèle de données :

- Utiliser la vue Diagramme pour définir des relations entre les tables.
- Concevoir des mesures avec les fonctions statistiques.
- Fonctions DAX pour concevoir des mesures élaborées : CALCULATE, ALL, ALLEXCEPT, FILTER, RANKX etc.
- Utiliser les fonctions DAX Time Intelligence

Conception de rapport Power BI:

- Création de rapports, pages, tableaux de bord.
- Afficher des données : Table, Matrice, Carte.
- Exploiter les paramètres de données et de formatage.
- Insérer des éléments visuels. Exploiter les visualisations cartographiques.



- Importer des éléments visuels (.pbviz) depuis la galerie : jauges, images etc.
- Ajouter des outils de filtrage, des segments, des KPI. Mise en page pour écran ou smartphone

Présentation de l'offre PowerBI Online:

- Présentation des fonctionnalités du service Power BI, stockage, présentation, partage.
- Épingler les visualisations dans un tableau de bord.
- Présentation de conception de rapport avec le service Power BI.
- Quels jeux de données dans Power BI?

3.1.2 TABLEAU SOFTWARE DESKTOP ANALYSTES

Valorisation des données / DataViz

TABLEAU SOFTWARE DESKTOP ANALYSTES

03.1.2

★ Avis : 7/10

2 jours (14 h)

BUT

Tableau Desktop est l'outil de développement de tableaux de bord de la solution Tableau Software. La partie analyste couvre les concepts importants de Tableau Desktop, ainsi que les techniques permettant de créer des visualisations et tableaux de bord interactifs

À l'issue de cette formation, chaque participant comprendra le positionnement de l'outil dans la suite Tableau Software, il aura une vision des différentes fonctionnalités proposées et saura identifier les enjeux de la BI agile.

COMPETENCES VISEES

Pouvoir construire des visualisations avancées sur des tableaux de bord et les déployer.

PRE-REQUIS :

Niveau ingénieur/Master informatique

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Chefs de projet, responsable de service, analystes

SOMMAIRE

Présentation de la Self-Service BI

La self BI selon Tableau Software : Rupture avec la BI conventionnelle

Architecture de Tableau Desktop

Fichiers Tableaux et structure d'un classeur : Classeurs, sources de données (directe, extraits...), feuilles de calcul, tableaux de bord, histoires
Structure des dossiers
Panneau de connexion aux données
Création du premier Tableau de bord

Se connecter aux données

Les différentes sources : fichiers plats, cubes, bases de données, cloud...
Jointure gauche, droite, externe...

Manipulation des données

Le Data Blending ou la fusion des données
Les groupes
Les agrégations : A la volée, par défaut, agrégation des totaux
Les champs calculés
Les paramètres
Les dates
Les filtres
Les types de données et fonctions usuelles (ZN...)
Les calculs de tables : Calcul de table rapide, édition des calculs de tables, et champ calculés créant des calculs de tables (RUNNING_SUM, INDEX(), LAST()...)
Les calculs de tables avancés : le partitionnement et l'adressage

ANIMATEUR

Consultant Tableau, société TRIMANE

Visualisations avancées

Cartographie de base
Cartes combinées
Ensembles, prévisions, courbes de tendances et lignes de projection
Pareto
Analyse des comportements d'achats
WaterFall ou diagrammes en cascade

Les tableaux de bord

Les composés de tuiles, flottants
Les actions : Activation du filtre par défaut, filtres personnalisés
Les actions de filtre, de sélection et d'URL

Les histoires (Story telling)

Histoire sur les trois clients ayant les plus grosses baisses de ventes
Les annotations

Partager son travail

Partage en local ou sur le portail Tableau Server
Exports : Image, données, tableaux à double entrées (tableaux croisés), PDF
Enregistrement de vues personnalisées
Partage des classeurs via scripts ou URL
Partage en local : format twb et twbx

3.1.3 SAP BUSINESS OBJECTS

Valorisation des données / DataViz

SAP BUSINESS OBJECTS WEBINTELLIGENCE 4.2 (NIVEAU I)

★ Avis : 8/10

03.1.3

2 jours (14h)

BUT

La nouvelle suite décisionnelle SAP Business Objects BI 4. permet de répondre à différents besoins :

- Comment transformer les données en rapports formatés - Comment visualiser les indicateurs clés de performance de l'entreprise
- Comment créer facilement ses rapports et les partager
- Comment analyser ses données OLAP de façon très détaillée - Comment piloter son activité et répondre à des questions métier

A l'issue de cette formation, chaque participant aura une expertise sur les principaux concepts de la solution, les fonctions d'interrogation, d'analyse, de reporting de la suite BI 4.

COMPETENCES VISEES

Acquérir un perfectionnement en matière de construction de tableaux de bord et de reporting sur SAP BO BI 4

PRE-REQUIS :

Posséder un bon niveau de connaissance XI3 « Utilisateur »

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Utilisateurs finaux, tout acteur impliqué dans le pilotage d'une organisation

ANIMATEUR

Consultant SAP BI, société TRIMANE

SOMMAIRE

Introduction à SAP BO WEBI 4.2

- Découvrir l'environnement
- Se connecter et se déconnecter du portail d'informations InfoView
- Sélectionner l'éditeur de documents WebIntelligence *Application* : *navigation dans l'InfoView*

- Création et modification d'un document
- Création d'une requête
- Exécuter et sauvegarder une requête
- Modifier une requête
- Visualiser et modifier les propriétés d'une requête
- Visualiser les propriétés d'un document "Drag and Drop" d'objets
- Dupliquer un tableau
- Rajouter un rapport à un document
- Rajouter une ligne ou une colonne à un tableau
- Sauvegarder un document sous plusieurs formats

Nouveau - Possibilité d'ouvrir et traiter plusieurs documents en même temps

Application : comment créer un premier document avec un requête simpl

Conceptions de rapports

- Manipulation de rapports et de tableaux
- Les différents modèles de tableaux WebIntelligence
- Créer un tableau dans un rapport
- Modifier les propriétés d'un tableau
- Cas particulier sur les tableaux croisés

Application : création, transformation et mise en forme de tableaux

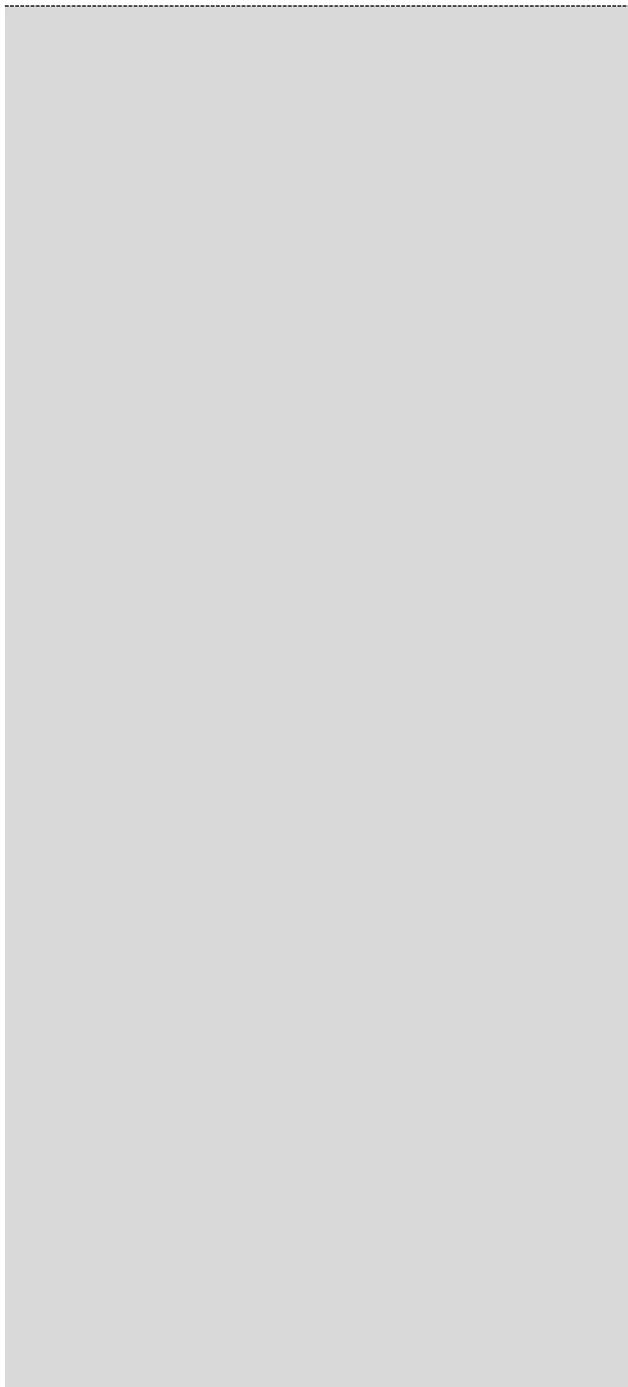
Manipulation de diagrammes

- Les différents types de diagrammes
- Créer un diagramme
- Transformer un tableau en diagramme

Application : création, transformation et mise en forme de diagrammes

Analyse des données

Les filtres :



Utiliser un filtre (prédéfini, personnalisé) Supprimer un filtre
Créer une invite pour sélectionner un sous-groupe de données
Sélectionner un sous-groupe de données basé sur des filtres multiples
Définir la répartition des filtres multiples

La notion de rupture (calcul et tris) :

Créer une rupture Calculer des totaux et des sous-totaux pour les données
Trier les données dans le rapport
Modifier l'ordre et le format des ruptures
Application : création d'un rapport intégrant un tableau avec rupture, calcul et tri

Filtres de rapports :

Sélectionner un filtre de données dans un rapport

Classement :

Ajouter un classement aux données
Application : création d'un filtre de rapport avec un hit parade

Les sections :

Diviser un rapport en sections
Calculer des totaux et sous-totaux par section
Poser un filtre de section
Tier les sections dans un rapport
Application : création d'un rapport avec sections, calculs et mise en page

Les formules :

Créer une formule pour afficher les résultats d'une invite
Créer une formule pour calculer une valeur
Créer une variable à partir d'une formule
Application : création de calculs, de formules et utilisation des variables

Nouveau - Appliquer un tri simple ou complexe dans un bloc

Formater les données d'un document

Rajouter un texte à un document
Modifier le format texte ou numérique dans un document
Nouveau - Suivi des données modifiées
Mise en forme conditionnelle des données (alerteurs)
Rajout d'un fond d'écran dans un document
Rajout d'une pagination dans un document
Changer le format d'un tableau, d'une cellule, d'une section

Nouveau - Modifier le formatage d'un diagramme

3.2 NIVEAU 2

3.2.1 POWER BI – PERFECTIONNEMENT

Valorisation des données / DataViz**POWER BI PERFECTIONNEMENT**

★ Avis : 8/10

03.2.1**2 jours (14h)****BUT**

Power BI, la solution de Data Visualisation de Microsoft, permet de créer des tableaux de bord riches et interactifs grâce aux multiples éléments graphiques qu'elle propose.

Ce stage vous permettra d'optimiser vos connaissances, d'aller plus loin dans la mise en forme et de maîtriser les fonctions de calcul les plus puissantes.

COMPETENCES VISEES

Plus concrètement, cette formation Power BI Avancé vous apportera les connaissances et compétences nécessaires pour:

Interroger et associer des sources de données hétérogènes

- Concevoir des mesures, des colonnes et des tables avec le langage DAX (Data Analysis Expressions)
- Utiliser des éléments visuels pour concevoir des tableaux de bord.

PRE-REQUIS :

Bonne connaissance de PowerBI

Pour suivre cette formation Power BI Avancé, il est important d'avoir suivi la formation Power BI - Initiation ou d'en posséder les connaissances équivalentes, et avoir pratiqué le développement sous Power BI de façon assez poussée

Niveau : Avancé

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Développeurs, consultants informatiques

ANIMATEUR**SOMMAIRE****Introduction:**

- Rappels sur le workflow de traitement Power BI Desktop - Power BI Online
- Présentation des nouveautés majeures des dernières versions de Power BI

Acquisition et préparation des données :

- Utilisation des sources de données avancées : listes SharePoint, bases SQL Server, fichiers PDF, sites web, jeux de données existants...
- Focus sur le mode « Direct Query » pour une base SQL Server. Avantages et limitations.

Fonctions avancées en M

- Filtrage de lignes
- Création de colonnes : à partir d'exemples, personnalisées, index, conditionnelles ...

Création de nouvelles tables en DAX (Data Analysis Expressions):

- Passage d'une table volumineuse à une table agrégée : Summarize / Calculatetable
- Création d'une table à partir d'une jointure: NaturalInnerJoin / NaturalLeftOuterJoin / Union

Focus sur les jointures entre tables dans le modèle de données

- Bien comprendre la cardinalité d'une jointure. Point d'attention sur la relation de « plusieurs à plusieurs »
- Explication sur le sens des jointures

Visuels avancés avec Power BI

- Indicateur de performances clés
- Influenceurs clés
- Construction d'un forecast
- Carte de forme (carte personnalisée)
- Carte de flux (trajets)
- Nuage de mots

Mise en forme avancée

- Créer un thème, un modèle de rapport
- Créer une home page et des menus de navigation en utilisant les signets

Consultant Tableau, société TRIMANE

- Utiliser des images pour améliorer la lisibilité du rapport
- Créer des rapports "Maitre-Détail" en utilisant les filtres d'extractions
- Navigation entre rapports
- Choisir la mesure à afficher avec un segment
- Afficher sur le rapport les filtres positionnés

Publication du rapport créé avec Power BI

- Retour sur les bases de la publication
- Publication sur Microsoft Teams
- Consommation des rapports sur l'application mobile
- Statistiques d'utilisation des rapports

Mettre à disposition son rapport, sécuriser les accès

- Interaction avec d'autres outils Office 365
- Déclencher une alerte à partir d'un indicateur avec Microsoft Flow
- Intégrer une application Power Apps dans un rapport Power BI

3.2.2 TABLEAU SOFTWARE DESKTOP DEVELOPPERS

Valorisation des données / DataViz

TABLEAU SOFTWARE DESKTOP DEVELOPPERS

★ Avis : 9/10

03.2.2

1 jour (7h)

BUT

Elle est conçue pour les professionnels qui ont acquis une solide expérience en utilisant Tableau et qui souhaitent passer au niveau supérieur. Vous maîtrisez sans doute la création des feuilles de calcul et des tableaux de bord, mais les aspects plus complexes vous posent encore un problème.

COMPETENCES VISEES

Les connaissances et compétences acquises permettent aux participants de créer rapidement une plate-forme Tableau robuste et évolutive. Remarque : Ce cours est dispensé dans un environnement de serveur Windows.

PRE-REQUIS :

Niveau ingénieur/Master informatique / Avoir suivie la formation Tableau Desktop niveau 1

MODALITES D'EVALUATION :

SOMMAIRE

Les fondamentaux

1. Rappels
 - Connexion aux données
 - Champs auto-générés
 - Champs continus et discrets
2. Approfondissements
 - Connexion aux données avancée

Manipulations des données

1. Rappels
 - Les hiérarchies
 - Les groupes
 - Les tris
 - Les calculs de tables simples
 - Les ensembles
2. Approfondissements
 - Les tris imbriqués
 - Les ensembles
 - Les calculs de tables avancées
 - Les filtres
3. Les paramètres

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...
Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Développeurs, consultants informatiques

ANIMATEUR

Consultant Tableau, société TRIMANE

4. Les LOD : Les expressions LOD sont utilisées pour exécuter des requêtes complexes à des niveaux de détails différents.

Cartographies avancées

Visualisations avancées

1. Boite à moustache
2. Diagramme à puce
3. Pareto
4. Waterfall
5. Gantt
6. Sparlines
7. Bump
8. Prévisions

Présenter ses analyses

1. Les Tableaux de bord complexes

Publier sur Tableau public

Fonctionnalités collaboratives Tableau Server

Administration : fonctionnalités Tableau

Public et administration

3.2.3 TABLEAU SOFTWARE SERVEUR

Valorisation des données / DataViz

TABLEAU SOFTWARE SERVER

★ Avis : 8.5/10

03.2.3

1 jour (7h)

BUT

Cette formation permettra aux stagiaires de pouvoir administrer et configurer le portail Tableau Server

COMPETENCES VISEES

À l'issue de cette formation, chaque participant saura utiliser Tableau de manière avancé

PRE-REQUIS :

Niveau ingénieur/Master informatique

PROFILS DES STAGIAIRES

Développeurs, consultants informatiques

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

SOMMAIRE

Architecture et configuration système de Tableau Server

Architecture de Tableau Server
Outil Tableau Server configuration : Onglet général, Data Connections, Servers, Alerts and Subscriptions, SSL, SAML, Kerberos, SAP Hana, Open ID

Maintenance du serveur et de la base de données

Sauvegarde des données Tableau Server
Restauration des données Tableau Server
Suppression des fichiers superflues
Scripts de sauvegarde et maintenance du serveur

Commandes usuelles

Commandes tabadmin
Commandes tabcmd

Architecture et configuration du portail Tableau Server

Les sites du portail
La gestion des utilisateurs
La gestion des groupes
La gestion des autorisations et droits d'accès
Les programmations et planifications

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

ANIMATEUR

Consultant Tableau, société TRIMANE

Les tâches d'extraits et d'abonnements
L'état du serveur et des sites
Les paramètres du serveur et des sites
Les paramètres de compte
Sécurité d'accès aux données et aux classeurs
Déploiement des extraits

Tableaux de bords d'administration

Tableaux de bord sur le portail
Accès à la base de données embarquées POSTGRESQL
Analyse des performances d'un tableau de bord sur le portail

3.2.4 SAP BUSINESS OBJECTS PERFECTIONNEMENT

Valorisation des données / DataViz

SAP BUSINESS OBJECTS WEBINTELLIGENCE 4.2 (NIVEAU II)

03.2.4

★ Avis : 9/10

1 jour (7h)

BUT

La nouvelle suite décisionnelle SAP Business Objects BI 4. permet de répondre à différents besoins :

- Comment transformer les données en rapports formatés
- Comment visualiser les indicateurs clés de performance de l'entreprise
- Comment créer facilement ses rapports et les partager
- Comment analyser ses données OLAP de façon très détaillée - Comment piloter son activité et répondre à des questions métier

A l'issue de cette formation, chaque participant aura une expertise sur les principaux concepts de la solution, les fonctions d'interrogation, d'analyse, de reporting de la suite BI 4.

COMPETENCES VISEES

Acquérir un perfectionnement en matière de construction de tableaux de bord et de reporting sur SAP BO BI 4

PRE-REQUIS :

Avoir suivi le stage « BO WEBI 4.2 niveau I » ou posséder un niveau équivalent.

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...
Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Utilisateurs finaux, tout acteur impliqué dans le pilotage d'une organisation

ANIMATEUR

Consultant SAP BI, société TRIMANE

SOMMAIRE

Rappel des principales fonctionnalités de SAP BO WEBI 4

Les univers BO et leur choix
La structure d'un document
L'éditeur de requêtes
Les classes et les objets La gestion des rapports

Quizz de validation des acquis

Création de formules
Formules de calcul basées sur des fonctions
Fonctions de type chaîne de caractères
Fonctions de type booléen
Fonctions de type numérique
Fonctions de type date
Opérations conditionnelles dans des formules
Création et utilisation de variables Applications

Création de requêtes complexes

Combinaison de requêtes
Création de sous-requêtes
Classement de requêtes
Maîtrise des opérateurs et des contextes de calculs
Opérateurs "Dans", "PourTout" et "PourChaque" Applications

Conceptions de rapports

Conception de rapports multiblocs
Conception de rapports multisources
Gestion des blocs et des cellules
Synchronisation des données provenant d'univers différents
Rassembler des sources de données multiples dans un même bloc
Nouveau – modifier le style par défaut
Applications

4. TECHNOLOGIES BI ET BIG DATA

4.1 NIVEAU 1

4.1.1 DEVELOPPEMENT PYTHON

Technologies BI et Big Data

DEVELOPPEMENT PYTHON – INITIATION

★ Avis : 9/10

04.1.1

2 Jours (14 h)

BUT

La formation vise à découvrir le langage Python, et rendre l'utilisateur autonome pour accéder, manipuler et synthétiser les informations disponibles dans les bases de données avec Python.

COMPETENCES VISEES

Être capable de :

- Lire et comprendre un code Python, langage de référence en data science.
- Manipuler et gérer les tableaux de données.
- Interroger, manipuler, ordonner et modifier un jeu de données avec Python.

PRE-REQUIS :

Appétence en mathématique

Cette formation nécessite des compétences intermédiaires en programmation orientée objet et calcul statistique.

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Tous public
Développeurs, Ingénieurs en traitement de données, MOA, Managers

ANIMATEUR

Formateur Data Science, société TRIMANE

SOMMAIRE

1. Présentation du langage Python et de son environnement

-Généralités.
-Distributions existantes avec focus sur ANACONDA, présentation des environnements de travail et des modules courants).

2. Les types de données

-Choisir les structures de données.
-Structures avancées de Python.

3. Fondamentaux de Python

-Variables et types.
-Opérateurs et tests.
-Boucles.
-Fonctions.
-Classes et modules.

4. Data

-Chargement et exploration d'un jeu de données.
-Data cleaning.
-Data processing.
-Pandas.
-Exportation des données.
-Fonctions de base pour la visualisation de données (bibliothèques : Matplotlib, Seaborn, ou Plotly).

5. Cas pratique : Pandas

-Créer des tables de données.
-Manipuler des tables de données.
-Trafic séquentiel de gros fichiers.

Technologies BI et Big Data

PL/SQL : LES FONDAMENTAUX

★ Avis : 8.5/10

04.1.2**5 Jours (35 h)****BUT**

Maîtriser le langage PL/SQL pour être capable de développer des applications classiques à partir et pour une base Oracle

COMPETENCES VISEES

Savoir récupérer des enregistrements et des colonnes de tables, utiliser des instructions, déclarer des variables PL/SQL, contrôler des flux de code, décrire des procédures stockées

PRE-REQUIS :

Connaissances SQL nécessaires

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...
Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Informaticiens, consultants

ANIMATEUR

Consultant Data, société TRIMANE

SOMMAIRE**Ecrire des scripts PL/SQL**

Développer votre premier script en PL/SQL
Syntaxe de base du langage
Types de données et conversion de types
Le cas particulier des dates (to_date, to_char, nls_date_format)
Déclarer, utiliser des variables et des constantes
Variables de type composite : tables PL/SQL, records, LOB, Bind variables
Portée d'une variable

Utilisation des curseurs

Déclarer et utiliser des curseurs
Le cycle de vie d'un curseur (open, fetch, close)
Utilisation des curseurs pour simplifier la programmation
Mises à jour sur les données traitées par un curseur
Propriétés des curseurs (%FOUND, %NOTFOUND, %ISOPEN, %ROWCOUNT)
Passer des paramètres à un curseur
Spécificités des curseurs implicites
Curseur en mise à jour (FOR UPDATE)

Procédures stockées et packages en PL/SQL

Utilisation des packages
Création de procédures et de fonctions stockées
Les packages prédéfinis fournis par Oracle (dbms..)
Création, compilation et stockage des programmes avec SQL PLUS
Définition et appel d'une procédure stockée
Les droits d'accès sur les procédures stockées
Gestion des dépendances

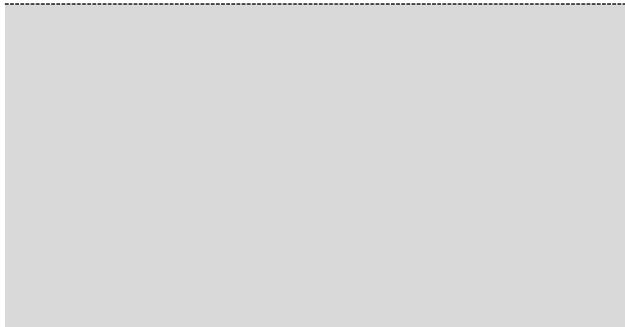
Gestion des erreurs

Philosophie des exceptions
Les types d'exception (prédéfinie ou utilisateur)
Code erreur et message associé (SQLCODE, SQLERRM)
Capture des exceptions
Gestion et propagation des exceptions
Personnaliser le traitement des erreurs

Les triggers

Utilisation des triggers (déclencheur)
Différents types de triggers
Valeur actuelle et nouvelle valeur (:NEW, :OLD)
Utilisation de la clause FOR EACH ROW
Activation et désactivation des triggers (vues DBA TRIGGERS)
Les triggers sur les vues et les événements systèmes
Implémenter des contraintes d'intégrité à l'aide des triggers
Triggers et transaction

Les transactions



Philosophie d'une transaction de bases de données
 Les instructions COMMIT, ROLLBACK et SAVEPOINT
 La problématique des accès concurrents (verrou)
 Exemples de blocages en PL/SQL
 La gestion standard des verrous par Oracle
 Transactions en lecture seule, niveaux d'isolation, verrouillages implicites et explicites
 Exemple de transaction trop volumineuse (rollback segment)

Recommandations et best practices

4.1.3 BASES NO SQL & MONGO DB

Technologies BI et Big Data

BASES NO SQL & MONGO DB

★ Avis : 7/10

04.1.3

2 jours (14h)

BUT

Power BI, la solution de Data Visualisation de Microsoft, permet de créer des tableaux de bord riches et interactifs grâce aux multiples éléments graphiques qu'elle propose.

Ce stage vous permettra d'optimiser vos connaissances, d'aller plus loin dans la mise en forme et de maîtriser les fonctions de calcul les plus puissantes.

COMPETENCES VISEES

Plus concrètement, cette formation Power BI Avancé vous apportera les connaissances et compétences nécessaires pour:

Interroger et associer des sources de données hétérogènes

- Concevoir des mesures, des colonnes et des tables avec le langage DAX (Data Analysis Expressions)
- Utiliser des éléments visuels pour concevoir des tableaux de bord.

PRE-REQUIS :

Niveau : Avancé

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Développeurs, consultants informatiques

SOMMAIRE

Introduction aux Big Data

BDR : Base de données relationnelle
 Base de données décisionnelles
 Quelques chiffres
 Big Data : Vs
 Un peu d'histoire
 Théorème CAP
 Présentation des technologies au cœur du Big Data
 SGBDR : propriétés ACID
 Limites des SGBDRs

Les bases de données NoSQL

Les bases NoSQL
 Mouvement NoSQL

La solution MongoDB

Popularité des SGBD NoSQL
 Avantages et fonctionnalités clés de MongoDB

Manipulation des données avec MongoDB (CRUD)

JSON vs BSON
 Stockage des données sur MongoDB : JSON vs BSON
 Importation et exportation des données
 Shell MongoDB et interrogation de la BD
 Insertion de document(s)
 Insertion de document(s)
 Suppression de collections et documents
 CRUD avancés : Opérateurs de comparaison
 Méthodes : sort() / limit() / skip()
 Indexes
 Upsert ()

Agrégation des données

Map-Reduce (MR)
 Framework d'agrégation (FA)
 Map-Reduce VS Framework d'agrégation

ANIMATEUR
 Consultant Big Data, société TRIMANE

Jointures et références
 Référencement Manuel
 BDRRefs
 Etape : \$lookup (jointure entre collections)

Recherche d'information
 Définition
 Recherche textuelle | Opérateur : \$text
 Recherche / Navigation à facettes

Modélisation des données Big Data
 Les données semi-structurées
 Modélisation d'une collection JSON
 Traduction du schéma E/A au niveau logique
 Relations entre les données
 L'importance du schema
 Les avantages et inconvénients des bases de données NoSQL
 Que choisir : Relationnel ou NoSQL

4.1.4 ECOSYSTEME HADOOP

Technologies BI et Big Data

ECOSYSTEME HADOOP

★ Avis : 7/10

04.1.4

1 jour optionnel + 2 jours (7h +14h)

BUT
 Installer hadoop, comprendre l'écosystème et savoir l'utiliser.

COMPETENCES VISEES
 Installation et manipulation d'un environnement Hadoop

PRE-REQUIS :
 Avoir suivi « Big Data : de l'expérimentation à la mise en production »

MODALITES D'EVALUATION :
 Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...
 Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES
 Tous profils souhaitant mettre en œuvre la big data, plus précisément : Hadoop

ANIMATEUR
 Consultant Big Data, société TRIMANE

SOMMAIRE

Installer Hadoop sur une station de travail (1 jour optionnel)
 Les différents modes de fonctionnement de Hadoop : Local, pseudo distribué, totalement distribué, clusters virtualisés
 Installation en mode pseudo-distribué
 Installation pas-à-pas sur une station de travail

Les Nouvelles pratiques pour Hadoop
 Le HDFS
 Le Map Reduce dans Hadoop

L'écosystème de Hadoop
 Pig,hive, API ? Passage en revue et mise en pratique
 Outils orientés bases de données
 Outils d'exploitations
 Intégration de Hadoop dans un système d'information

Déveollper des programmes Hadoop
 Le Wordcount
 Le joinTables

Quand utiliser Hadoop
 Exemple de problématique
 Exemple réel d'utilisation

4.2 NIVEAU 2

4.2.1 DEVELOPPEMENT PYTHON INTERMEDIAIRE

Technologies BI et Big Data**DEVELOPPEMENT PYTHON INTERMEDIAIRE****04.2.1****2 Jours (14 h)**

★ Avis : 8/10

BUT

Cette formation a pour objectif de rendre l'apprenant autonome sur Python, de lui permettre de comprendre, modifier et écrire le code pour les besoins les plus courants (manipulations de données, statistiques descriptives), mais aussi pour des requêtes avancées (calculs complexes, formules conditionnelles).

COMPETENCES VISEES

Gagner en autonomie sur Python, comprendre, modifier et écrire le code pour les besoins les plus courants (manipulations de données, statistiques descriptives), mais aussi pour des requêtes avancées (calculs complexes, formules conditionnelles) et automatiser une analyse des données.

Acquérir les bases pour :

- Appliquer les bonnes pratiques.
- Maîtriser les designs patterns.
- Apprendre à améliorer les performances.
- Garantir la qualité.

PRE-REQUIS :

Connaissance de base du langage Python. Cette formation nécessite des compétences intermédiaires en programmation et calcul statistique.
(un débutant ou un ancien utilisateur ayant peu de pratique s'orientera sera orienté vers la formation Python Initiation)

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...
Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Développeurs, ingénieurs en traitement de données

ANIMATEUR

Consultant Expert Python, société TRIMANE

SOMMAIRE**1. Présentation générale**

-Le calcul scientifique et utilisation de Python
-Les distributions Python pour les DataScientists (WinPython, Python Anywhere, Python (x,y), EnthoughtCanopy, Anaconda)
-Exemples d'outils de développement

2. Bases du Python scientifique

-La Scipy Stack : présentation
-Clacul numérique avec Numpy: la bibliothèque Python de calcul numérique (algèbre linéaire) permettant de manipuler des vecteurs et matrices à plusieurs dimensions.
-Analyse des structures de données complexes avec Pandas.
-Visualisation des données avec Matplotlib (Affichage de plusieurs types de représentations interactives, 2D comme 3D).
-Gestion des erreurs numériques : comprendre les problèmes d'erreurs numériques dans le calcul scientifique.

3. Usages avancés

Maîtrise de Numpy et Pandas, ainsi que les bibliothèques Scipy et Scikit Image.

4. Visualisation de données

Les principales librairies de visualisation disponibles pour Python
Seaborn.
Altair.
Visualisation de données cartographiques.
Applications Open Sources pour la visualisation de larges jeux de données.

5. Parallélisme

Parallélisation.
Multi-Threading.
Multi-processing.

Cas pratique :

Exercices d'application avec des algorithmes de Machine Learning
Étude de cas (Machine Learning avec Python)
Solution d'optimisation

- Parallélisme (Multi-procesing, Multithreading)
- Bilan et questions Apprentissage non-supervisé

4.2.2 TALEND OPEN STUDIO

Technologies BI et Big Data – Acquisition des données**ETL : TALEND**

★ Avis : 9/10

04.2.2**3 jour (21h)****BUT**

Etre capable de développer des jobs d'ETL de manière autonome

COMPETENCES VISEES

Base/ Perfectionnement

PRE-REQUIS :

Bonnes connaissances en SQL et programmation Java
 Disposer de l'ensemble des connaissances nécessaires pour la réalisation de premiers jobs et leur mise en production. Avoir un aperçu complet sur le Business Modeler, le gestionnaire des méta-données, le Job Designer, l'exécution, le débogage et le déploiement des jobs.

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...
 Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Développeurs, consultants, Chefs de projet

ANIMATEUR

Consultant Talend, société TRIMANE

SOMMAIRE**Introduction**

Le marché de l'intégration de données
 Les solutions du marché
 Vue d'ensemble de TOS

Modéliser et documenter

Présentation du Business Modeler
 Documentation
 Association documentation / jobs

Premiers pas avec TOS

Vue d'ensemble de TOS
 Génération de données de tests Composants de communication avec les bases de données
 Composants de communication avec les fichiers Gestionnaire de métadonnées Grands Principes (Gestion des contextes)
 Centralisation des connexions Centralisation des structures de fichiers, de bases de données Transformations
 Composants de transformations Composants de « mapping »
 Gestion des filtres et des rejets
Gestionnaire de métadonnées
 Grands Principes (Gestion des contextes)
 Centralisation des connexions
 Centralisation des structures de fichiers, de bases de données

Transformations

Composants de transformations
 Composants de « mapping »
 Gestion des filtres et des rejets

Manipulation de données avec des composants avancés

Composants liés à Internet

Fonctionnalités de développement (Java)

Intégrer du code Java dans ses jobs
 Gestion des routines (créer sa propre librairie d'opération Java)

Déploiement et Débogage

Exporter son job
 Optimiser et déboguer ses jobs

Technologies BI et Big Data – Acquisition des données**MAP R / ARCHITECTURE ET ADMINISTRATION****04.2.3****3 jour (21h)**

★ Avis : 6/10

BUT

Cette formation MapR pour architectes et administrateurs nécessite de savoir naviguer et modifier des fichiers dans **Linux**. Une connaissance préliminaire d'**Hadoop** n'est pas exigée.

COMPETENCES VISEES

Cette formation Hadoop pour architectes et administrateurs vous permettra sur MapR de savoir:

- Comment élaborer une architecture Hadoop
- Comment Hadoop Distributed File System (HDFS) et Yarn/MapReduce fonctionnent
- Comment exploiter et gérer un cluster Hadoop
- Comment configurer Hadoop et ses outils
- Comment mettre en œuvre un cluster Hadoop
- Comment gérer la maintenance des nœuds dans un cluster
- Comment déplacer et gérer les données dans un cluster
- Comment administrer un cluster Hadoop

PRE-REQUIS :

Cette formation MapR pour architectes et administrateurs nécessite de savoir naviguer et modifier des fichiers dans **Linux**. Une connaissance préliminaire d'**Hadoop** n'est pas exigée.

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Cette formation MapR pour architectes et administrateurs est destinée aux administrateurs systèmes et aux personnes qui gèrent et maintiennent les clusters Apache Hadoop en production ou qui évoluent dans un environnement de développement.

ANIMATEUR

Consultant BI et Big Data, société TRIMANE

SOMMAIRE**Introduction**

Présentation générale d'Hadoop
Exemples d'utilisation dans différents secteurs
Historique et chiffres clés : Quand parle-t-on de Big Data ?
Les possibilités d'implantation de l'infrastructure et les coûts associés

L'écosystème d'Hadoop

Le système de fichier HDFS
Le paradigme MapReduce et l'utilisation à travers YARN
Le garant des informations : Zookeeper
Le système de stockage de la donnée basé sur le socle : HBase
La présentation de l'architecture globale : où installer les services ?
La définition du dimensionnement pour un cluster Hadoop
Les particularités de MapR

Installation du socle d'Hadoop

Installation et configuration manuelle pas-à-pas sur plusieurs machines de :
HDFS, YARN, Zookeeper, HBase

Manipulation des données dans un cluster Hadoop

Architecture, Objectif et fonctionnement de : Hue, Hive, Pig, Oozie, Sqoop, Flume, Spark
Mise en pratique sur les différents outils

Exploitation d'un cluster Hadoop

Architecture, objectif et fonctionnement du Manager : MapR Control System
Mise en pratique

Concepts d'architecture complexe

Architecture en rack : les avantages et fonctionnement et paramétrage du rack awareness
Scalabilité et performances : le Namenode Federation
High Availability des serveurs maîtres
Utilisation des snapshots : sauvegarde et restauration

Concepts d'architectures complexes

Architecture en rack : les avantages et fonctionnement et paramétrage du rack awareness
Scalabilité et performances : le Namenode Federation
High Availability des serveurs maîtres
Utilisation des snapshots : Sauvegarde et restauration

Technologies BI et Big Data – Acquisition des données**CLOUDERA / HORTONWORKS : ARCHITECTURE ET ADMIN****04.2.4**

★ Avis : 5.5/10

3 jour (21h)**BUT**

Cette formation Cloudera/Hortonworks pour architectes et administrateurs est destinée aux administrateurs systèmes et aux personnes qui gèrent et maintiennent les clusters Apache Hadoop en production ou qui évoluent dans un environnement de développement.

COMPETENCES VISEES

Cette formation Hadoop pour architectes et administrateurs vous permettra sur Hortonworks de savoir:

- Comment élaborer une architecture Hadoop
- Comment Hadoop Distributed File System (HDFS) et YARN/MapReduce fonctionnent
- Comment exploiter et gérer un cluster Hadoop
- Comment configurer Hadoop et ses outils
- Comment mettre en œuvre un cluster Hadoop
- Comment gérer la maintenance des nœuds dans un cluster
- Comment déplacer et gérer les données dans un cluster
- Comment administrer un cluster Hadoop

PRE-REQUIS :

Cette formation Hortonworks pour architectes et administrateurs nécessite de savoir naviguer et modifier des fichiers dans Linux. Une connaissance préliminaire d'Hadoop n'est pas exigée.

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Cette formation MapR pour architectes et administrateurs est destinée aux administrateurs systèmes et aux personnes qui gèrent et maintiennent les clusters Apache Hadoop en production ou qui évoluent dans un environnement de développement.

ANIMATEUR

Consultant Talend, société TRIMANE

SOMMAIRE**Introduction**

Présentation générale d'Hadoop
Exemples d'utilisation dans différents secteurs
Historique et chiffres clés : Quand parle-t-on de Big Data ?
Les possibilités d'implantation de l'infrastructure et les coûts associés

L'écosystème d'Hadoop

Le système de fichier HDFS
Le paradigme MapReduce et l'utilisation à travers YARN
Le garant des informations : Zookeeper
Le système de stockage de la donnée basé sur le socle : HBase
La présentation de l'architecture globale : où installer les services ?
La définition du dimensionnement pour un cluster Hadoop
Les particularités de MapR

Installation du socle d'Hadoop

Installation et configuration manuelle pas-à-pas sur plusieurs machines de :
HDFS, YARN, Zookeeper, HBase

Manipulation des données dans un cluster Hadoop

Architecture, Objectif et fonctionnement de : Hue, Hive, Pig, Oozie, Sqoop, Flume, Spark
Mise en pratique sur les différents outils

Exploitation d'un cluster Hadoop

Architecture, objectif et fonctionnement du Manager : MapR Control System
Mise en pratique

Concepts d'architecture complexe

Architecture en rack : les avantages et fonctionnement et paramétrage du rack awareness
Scalabilité et performances : le Namenode Federation
High Availability des serveurs maîtres
Utilisation des snapshots : sauvegarde et restauration

Concepts d'architectures complexes

Architecture en rack : les avantages et fonctionnement et paramétrage du rack awareness
Scalabilité et performances : le Namenode Federation
High Availability des serveurs maîtres
Utilisation des snapshots : Sauvegarde et restauration

4.3 NIVEAU 3

4.3.1 TALEND AVANCE

Technologies BI et Big Data – Acquisition des données**ETL : TALEND AVANCE**

★ Avis : 8/10

04.3.1**2 jours (14h)****BUT**

Cette formation permet de maîtriser les composants et fonctionnalités avancés de Talend Open Studio et de créer des composants spécifiques afin d'étendre les fonctionnalités du logiciel.

COMPETENCES VISEES

Perfectionnement

- Maîtriser les composants avancés de Talend Open Studio ;
- Fédérer son code pour gagner du temps ;
- Debugger et déployer ses jobs ;
- Étendre Talend Open Studio en créant ses propres composants

PRE-REQUIS :

Avoir suivi la formation Talend - Niveau 1, ou posséder les connaissances et compétences équivalentes.

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Concepteurs-Développeurs initiés à l'utilisation de Talend

ANIMATEUR

Consultant Expert Talend, société TRIMANE

SOMMAIRE**Introduction****Gestionnaire de métadonnées**

Centralisation des connexions et structures
Gestion des contextes

Maitrise des composants évolués

Composants java
Composants xml

Les erreurs

Gestion des rejets et logs

Réutilisabilité

Gestion des routines

Le mode DEBUG

Les différents types de debugging
Focus sur le debug Java

Optimisation

Estimer la performance d'un job
Points d'attentions

Cycle de vie des développements

Export des éléments
Exécution en production, paramètre de la command line

Création de composants

Le Template ?
La notion de Begin/main/end

La communauté de Talend

5. DATA DRIVEN

05.1 NIVEAU 1

5.1.1 ECONOMIE NUMERIQUE ET UTILISATION DES DONNEES

Data Driven Enterprise

ECONOMIE NUMERIQUE ET UTILISATION DES DONNEES (STRATEGIE D'ENTREPRISE ET BIG DATA)

05.1.1

★ Avis : 9/10

1 jour (7 h)

BUT

Cette formation a pour objectif de permettre à des dirigeants d'entreprise de voir comment s'intègre un projet big data/data science dans une stratégie d'entreprise.

COMPETENCES VISEES

Base/ Perfectionnement

PRE-REQUIS :

Basique

MODALITES D'EVALUATION :

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

PROFILS DES STAGIAIRES

Ce stage est destiné à :

- des managers souhaitant appréhender concrètement les apports du big data dans leur domaine.,
- des managers souhaitant aborder ce virage.

ANIMATEUR

L.RIGAUD (SUPAERO 1994) – Président fondateur du groupe TRIMANE

SOMMAIRE

Le Big data/data science
Le Phénomène Big Data
Le Phénomène Datascience
Des Problèmes d'utilisation
Les Problèmes d'interprétations

Stratégie d'entreprise
Définir son objectif,
Définir ses indicateurs,
Définir les modalités de suivi de ses indicateurs
Définir l'organisation pour y aller : entreprise datadriven ♣ La datagouvernance

Promouvoir un projet big data
A garder en tête sur le big data,
La démarche de promotion

Approche pilotage projet big data
Lancer un projet Big Data,
Les causes de flops