

SOMMAIRE

VIII– BI , Big data et data science

Oracle BI 12c Build Repositories	3
Oracle BI 12c Create Analyses and Dashboards Ed 1	3
Implementing Data Models and Reports with SQL Server 2014 [20466DC]	4
Designing Business Intelligence Solutions with Microsoft SQL Server 2014[20467DC].....	4
Les fondamentaux du big data : Enjeux et Perspectives du Big Data	5
bases de données NoSQL	5
Big Data, méthodes et solutions pratiques pour l'analyse des données.....	6
Analyzing Big Data with Microsoft R [20773AC].....	6
Perform Cloud Data Science with Azure Machine Learning [20774AC].....	7

Oracle BI 12c: Build Repositories

Objectifs :

build and verify the three layers of an Oracle Business Intelligence (BI) repository . use the Oracle BI Administration Tool to construct a simple repository.

Prérequis

basic skills in administering oracle 12C, and relational database.

Participants

. Développeurs d'ap-

Programme

5 jours

Repository Basics	Table Sources	Variables
Building the Physical Layer of a Repository	Adding Calculations to a Fact	Modeling Time Series Data
Building the Business Model and Mapping Layer of a Repository	Working with Logical Dimensions	Modeling Many-to-Many Relationships Setting an Implicit Fact Column
Building the Presentation Layer of a Repository	Enabling Usage Tracking Using Model Checker and Aggregates	Importing Metadata from Multidimensional Data Sources
Testing and Validating a Repository	Using Partitions and Fragments	Security
Managing Logical	Using Repository Va-	Cache Management

Oracle BI 12c: Create Analyses and Dashboards Ed 1

Objectifs :

creating Oracle BI analyses and dashboards in business intelligence applications . building basic analyses to include in dashboards, with more complexity.

Prérequis :

basic skills in administering oracle 12C, and relational database.

Participants:

Utilisateurs finals
Développeurs d'appli-

Programme

5 jours

Introduction to Oracle Business Intelligence Enterprise Edition	tional Views in Analyses	Ask, and Data Mashup
Working with Oracle Business Intelligence Analyses	Visualizing Data on Maps	Working with BI Content in Smart View
Filtering Data for Analyses	Creating Oracle Business Intelligence Dashboards	Oracle Business Intelligence Mobile
Selecting and Grouping Data for Analysis	Configuring Oracle Business Intelligence Dashboards	Using Oracle Business Intelligence Delivers
Advance Visualization in Analysis	Creating Dashboard Prompts and Variables	Administering the Presentation Catalog
Showing Results with Pivot Tables	Scorecards and Key Performance Indicators	Direct Database Analysis Expressions
Working with Addi-	Visual Analyzer, BI	

Implementing Data Models and Reports with SQL Server 2014 [20466DC]

Objectifs :

creating managed enterprise BI solutions. implement multidimensional and tabular data models, deliver reports with SQL Server® Reporting Services, create dashboards with SharePoint Server PerformancePoint Services, and discover business insights by using data mining.

Prérequis

At least 2 years' experience of working with relational databases .

Participants

Programme

5 jours

Module 1: Introduction to Business Intelligence and Data Modeling

Lab : Exploring a BI Solution

Module 2: Creating Multidimensional Databases

Lab : Creating a Multidimensional Database

Module 3: Working with Cubes and Dimensions

Lab : Working with Cubes and Dimensions

Module 4: Working with Measures and Measure Groups

Lab : Configuring Measures and Measure Groups

Module 5: Introduction to MDX

Lab : Using MDX

Module 6: Customizing Cube Functionality

Lab : Customizing a Cube Implement Translations.

Module 7: Implementing an

Analysis Services Tabular Data Model

Lab : Implementing an Analysis Services Tabular Data Model

Module 8: Introduction to Data Analysis Expression (DAX)

Lab : Creating Calculated Columns and Measures by using DAX

Module 9: Implementing Reports with SQL Server Reporting Services

Lab : Creating a Report with

Designing Business Intelligence Solutions with Microsoft SQL Server 2014 [20467DC]

Objectifs :

Plan and design a BI solution that is based on SQL Server 2014 .

Prérequis :

A basic understanding of dimensional modeling , server hardware knowledge, Integration Services packages that include control flows and data flows, multidimensional cube with Analysis Services, tabular model with PowerPivot and Analysis Services .

Participants:

BI professionals with experience of implementing solutions with SQL

Programme

5 jours

Module 1: Planning a BI Solution

Lab : Planning BI Solutions

Module 2: Planning SQL Server Business Intelligence Infrastructure

Lab : Planning BI Infrastructure

Module 3: Designing a Data Warehouse

Lab : Designing a Data Warehouse Logical Schema

Module 4: Designing an

ETL solution

Lab : Designing an ETL Solution

Module 5: Designing Analytical Data Models

Lab : Designing Analytical Data Model
Lab : Designing Dimensions and Hierarchies

Module 6: Planning a BI Delivery Solution

Lab : Planning a BI Delivery Solution

Module 7: Designing a Reporting Services Solution

tion

Lab : Designing a Reporting Services Solution

Module 8: Designing an Excel Based reporting Solution

Lab : Designing and Excel Reporting Solution

Module 9: Planning a SharePoint Server BI Solution

Lab : Implementing a SharePoint Server BI Solution

Lab : Implementing PerformancePoint Services

Les fondamentaux du big data : Enjeux et Perspectives du Big Data

3 jours

Objectifs :

Présenter les enjeux de l'analyse des données numériques à grande échelle, Comprendre le Big Data et son impact dans l'entreprise.

Comment faire du Big Data ? Sur quelle plateforme technologique s'appuyer ? les outils de l'écosystème Hadoop pour chaque cas d'usage.

Prérequis

Chefs de projet, responsables marketing, développeurs Big Data, Consultant BI, directeur de projet.

Programme

Comprendre le BIG Data

- Introduction (Jeu : Donnez un titre aux images)
- De la donnée au Big Data (l'histoire de Google qui compare un grain de riz à un octet)
- Quelles avancées technologiques ont permis l'avènement du Big Data
- Quelles évolutions comportementales ont accéléré l'avènement du Big Data
- Les définitions du Big Data

Les 5V du Big Data Les sources de données

- Pourquoi capturer les données ?
- Les données internes de l'entreprise
- Les données externes : Web
- Les données externes : Open Data

Les données externes : Les réseaux sociaux Big data en entreprise

- Les secteurs impactés
- Cas d'usages du Big Data

La gouvernance des données

Les métiers du Big Data

- Administrateur, Architecte, Développeur, Data Analyst, Data Scientist, DPO, Data Owner, CDO, ..
- Quel profil pour quel poste ?

Le Cadre juridique des données : GDPR

Big Data vs Business Intelligence

NOSQL

Panorama des outils Big Data

Types d'architectures

Bases de données NOSQL

3 jours

Objectifs :

Comprendre les différents concepts des bases avancées et limites du NoSQL, Savoir choisir une base NoSQL (clé/valeur, document, colonne, graphe) en fonction des cas d'usages, créer, manipuler et interroger des données dans HBase.

Prérequis :

Aucune connaissance sur Hadoop n'est requise.

Certains exemples et exercices sont présentés en Java, SQL et Shell donc une connaissance d'un de ces langages est un plus.

Programme

Les Bases de données relationnelles

- Formes normales
- Modèle structuré
- Transaction ACID
- Forces et limites

Les Principes du NoSQL

- Pourquoi est né le NoSQL?
- CAP vs ACID
- Données structurées vers non-structurées
- Évolutivité

- Les principes d'une architecture distribuée

Les bases Clé/Valeur

- Avantages/inconvénients
- Riak : Présentation et mise en oeuvre

Les bases document

- Avantages/inconvénients
- MongoDB: Présentation et mise en oeuvre

Les bases orientées colonne

- Avantages/

inconvénients

- Hbase & Cassandra: Présentation et mise en oeuvre

Les bases graphes

- Avantages/inconvénients
- neo4j: Présentation et mise en oeuvre

Cas Pratique Big Data SGBDR VS NoSQL

- Différents cas d'usage : Comment y répondre? Avec un SGBDR ou avec un système NoSQL?
- Conception du modèle de données dans les 2

Formation Big Data, méthodes et solutions pratiques pour l'analyse des données

5 jours

Objectifs :

comprendre les enjeux et les apports du Big Data , intégrer des volumétries massives de données structurées et non structurées via un ETL, analyser grâce à des modèles statistiques et des dashboards dynamiques.

Prérequis

Connaissances de base des modèles relationnels, des statistiques et des langages de programmation. Connaissances de base des concepts de la BI .

Participants

Programme

Comprendre les concepts et les enjeux du Big Data

- Origines et définition du Big Data.
- Les chiffres clés du marché dans le monde et en France.
- Les enjeux du Big Data : ROI, organisation, confidentialité des données.

Un exemple d'architecture Big Data.

Les technologies du Big Data

- Description de l'architecture et des composants de la plateforme Hadoop.
- Les modes de stockage (NoSQL, HDFS).
- Principes de fonctionne-

ment de MapReduce, Spark, Storm...

- Principales distributions du marché (Hortonworks, Cloudera, MapR, Elastic Map Reduce, Biginsights).
- Installer une plateforme Hadoop.
- Les technologies du datascientist.

Présentation des technologies spécifiques pour le Big Data (Tableau, Talend, Qlikview ...).

Exercice

Installation d'une plateforme Big Data Hadoop (via Cloudera QuickStart ou autre).

Gérer les données structurées et non structurées

- Principes de fonctionne-

ment de Hadoop Distributed File System (HDFS).

- Importer des données externes vers HDFS.
- Réaliser des requêtes SQL avec HIVE.
- Utiliser PIG pour traiter la donnée.
- Le principe des ETL (Talend...).

Gestion de streaming de données massive (NIFI, Kafka, Spark, Storm...)

Exercice

Implémentation de flux de données massives.

Technique et méthodes Big data analytics

Exercice

Mise en place d'analyses avec une des outils

Analyzing Big Data with Microsoft R [20773AC]

3 jours

Objectifs :

use Microsoft R Server to create and run an analysis on a large dataset, and show how to utilize it in Big Data environnements, such as a Hadoop or Spark cluster, or a SQL Server database.

Prérequis :

Programming experience using R, Knowledge of common statistical methods and data analysis best practices.

Participants:

people who wish to analyze large datasets , developers who need to integrate R analyses into their solutions.

Programme

Module 1: Microsoft R Server and R Client Explain how Microsoft R Server and Microsoft R Client work.Lessons

Explain the purpose of the ScaleR functions.

Lab : Exploring Microsoft R Server and Microsoft R Client

- **Module 2:** Exploring Big DataAt the end of this module the student will be able to use R Client with R Server to explore big data held in different data stores.

Lab : Exploring Big Data

Module 3: Visualizing Big DataExplain how to visualize data by using graphs and plots.

Lab : Visualizing data

Module 4: Processing Big DataExplain how to transform and clean big data sets.

Lab : Processing big data

Module 5: Parallelizing Analysis OperationsExplain how to implement options for splitting analysis jobs

Lab : Using rxExec and RevoPemaR to parallelize operations

Module 6: Creating and Evaluating Regression ModelsExplain how to build and evaluate regression models generated from big data

Lab : Creating a linear regression model

Module 7: Creating and Evaluating Partitioning ModelsExplain how to create and score partitioning models generated from big data.

Perform Cloud Data Science with Azure Machine Learning [20774AC]

Objectifs :

analyze and present data by using Azure Machine Learning, and to provide an introduction to the use of machine learning with big data tools such as HDInsight and R Services .

Prérequis

Programming experience using R, Knowledge of common statistical methods and data analysis best practices. Relational database

Participants

Programme

5 jours

Module 1: Introduction to Machine Learning This module introduces machine learning and discussed how algorithms and languages are used.

Lab : Introduction to machine Learning

Module 2: Introduction to Azure Machine Learning- Describe the purpose of Azure Machine Learning, and list the main features of Azure Machine Learning Studio.

Lab : Introduction to Azure machine learning

Module 3: Managing Datasets At the end of this module the student will be able to

upload and explore various types of data in Azure machine learning.

Lab : Managing Datasets

Module 4: Preparing Data for use with Azure Machine Learning This module

Lab : Preparing data for use with Azure machine learning

Module 5: Using Feature Engineering and Selection

Lab : Using feature engineering and selection

Module 6: Building Azure Machine Learning

Lab : Building Azure ma-

chine learning models

Module 7: Using Classification and Clustering with Azure machine learning models

Lab : Using classification and clustering with Azure machine learning models

Module 8: Using R and Python with Azure Machine Learning

Lab : Using R and Python with Azure machine learning

Module 9: Initializing and Optimizing Machine Learning Models

Lab : Initializing and optimizing machine learning