Course : Big Data, technical overview

Synthesis course - 2d - 14h00 - Ref. BAG
Price : 2020 CHF E.T.

★★★★★   **5 / 5**

This summary will introduce you to the challenges and benefits of big data, and the technologies available for its implementation. You'll follow the steps involved in a massive data project, from setting up a big data platform, ingesting and processing the data, to visualizing the results.

Teaching objectives

At the end of the training, the participant will be able to:

- ✓ Discover the key concepts of big data
- ✓ Understanding the technological ecosystem of a big data project
- ✓ Evaluate techniques for managing massive data flows
- ✓ Implement statistical analysis models to meet business needs
- ✓ Discover data visualization tools

Intended audience

Dataminers, statistical researchers, developers, project managers, business intelligence consultants.

Prerequisites

Basic knowledge of relational models, statistics and programming languages. Basic knowledge of business intelligence concepts.

Practical details

Demonstration

Present the Hadoop platform and its basic components, use an ETL to manage data, create analysis models and dashboards.

Course schedule

PARTICIPANTS

Dataminers, statistical researchers, developers, project managers, business intelligence consultants.

PREREQUISITES

Basic knowledge of relational models, statistics and programming languages. Basic knowledge of business intelligence concepts.

TRAINER QUALIFICATIONS

The experts leading the training are specialists in the covered subjects. They have been approved by our instructional teams for both their professional knowledge and their teaching ability, for each course they teach. They have at least five to ten years of experience in their field and hold (or have held) decision-making positions in companies.

ASSESSMENT TERMS

The trainer evaluates each participant's academic progress throughout the training using multiple choice, scenarios, hands-on work and more.
Participants also complete a placement test before and after the course to measure the skills they've developed.

1   Understanding the key concepts and challenges of big data

- The origins of big data.
- The value of data: an important change.
- Data as raw material.
- Key market figures worldwide and in France.
- The challenges of big data: ROI, organization, data confidentiality.

Demonstration
Introduction to big data architecture.

2   Big data technologies

- Architecture and components of the Hadoop 2 platform.
- Storage modes (NoSQL, HDFS).
- How MapReduce and Yarn work...
- Main Hadoop distributions: Hortonworks, Cloudera, MapR...
- Technologies: Spark, Storm, Databrick, Azure Machine Learning...
- How to install a Hadoop platform.
- Presentation of specific Big Data technologies (Talend, Tableau, QlikView...).

Demonstration
Installation of a complete big data platform.

3   Big data processing

- How Hadoop Distributed File System (HDFS) works.
- Import data to HDFS.
- Data processing with PIG.
- SQL queries with HIVE.
- Create massive data flows with an ETL.

Demonstration
Implementation of massive data flows.

4   Data analysis and processing methods for Big Data

- Exploration methods.
- Segmentation and classification.
- machine learning, estimation and prediction.
- Real time, artificial intelligence.
- Model implementation.

Demonstration
Introduction to the Spark environment, Jupyter Notebook, R Notebook and Shiny. Setting up machine learning analyses with R, Python and Scala.

TEACHING AIDS AND TECHNICAL RESOURCES

• The main teaching aids and instructional methods used in the training are audiovisual aids, documentation and course material, hands-on application exercises and corrected exercises for practical training courses, case studies and coverage of real cases for training seminars.

• At the end of each course or seminar, ORSYS provides participants with a course evaluation questionnaire that is analysed by our instructional teams.

• A check-in sheet for each half-day of attendance is provided at the end of the training, along with a course completion certificate if the trainee attended the entire session.

TERMS AND DEADLINES

Registration must be completed 24 hours before the start of the training.

ACCESSIBILITY FOR PEOPLE WITH DISABILITIES

Do you need special accessibility accommodations? Contact Mrs. Fosse, Disability Manager, at psh-accueil@orsys.fr to review your request and its feasibility.

5 Data visualization, representing data visually

- Market-leading solutions.
- Going beyond static reports.
- Data visualization and the art of telling numbers in a creative and entertaining way.
- Measure e-reputation, brand awareness, customer experience and satisfaction...

Demonstration
Presentation and use of a data visualization tool to create dynamic analyses.

6 Conclusion

- Conditions for success.
- Synthesis of best practices.
- Bibliography.

Dates and locations

REMOTE CLASS
2026 : 2 Apr., 26 May, 8 Sep., 17 Nov.