Data Science & Big Data Analytics

Data Science & Big Data Analytics

5 days training course

For detailed information on training course dates, please click the link:

[Data Science & Big Data Analytics](#).

Course Overview

The **Data Science & Big Data Analytics** course provides participants with the essential tools and techniques for analyzing and deriving actionable insights from vast amounts of data. It covers the basics of data science, data visualization, machine learning algorithms, and big data tools. Participants will also learn how to apply these techniques in real-world scenarios to address business challenges, improve decision-making, and foster innovation.

Objectives

- Gain a comprehensive understanding of data science principles and methodologies.
- Learn how to process, analyze, and visualize large datasets using popular data science tools and techniques.
- Understand the fundamentals of machine learning and how to apply algorithms to big data.
- Develop the ability to work with big data platforms and technologies, including Hadoop and Spark.
- Master data cleaning, feature engineering, and building predictive models.
- Learn how to use data-driven insights to inform business strategies and decisions.

Who Should Attend?

This course is designed for professionals in data-driven roles who are seeking to enhance their skills in data science and big data analytics. It is ideal for data analysts, business analysts, data engineers, IT professionals, and managers who want to understand how to extract valuable insights from large datasets to drive business decisions. Individuals from various industries, such as finance, healthcare, marketing, and manufacturing, will benefit from this course.

Course Outline:

Day 1: Introduction to Data Science and Big Data

- Overview of data science and its applications in various industries
- Key concepts: Data collection, processing, and analysis
- Big Data: Definition, sources, and importance
- Types of data: Structured, unstructured, and semi-structured
- Introduction to data science tools: Python, R, SQL
- Practical exercise: Setting up a data science environment (e.g., installing Python, R, and Jupyter notebooks)
- Case study: Real-world applications of data science in business

Day 2: Data Processing and Data Cleaning

- The data preparation pipeline: From raw data to ready-to-use data
- Data wrangling techniques using Python and R
- Handling missing data and outliers
- Data transformations and normalization
- Feature engineering and creating new variables from existing data
- Practical exercise: Cleaning and preparing a messy dataset for analysis
- Case study: Data preprocessing for machine learning

Day 3: Data Visualization and Exploratory Data Analysis (EDA)

- Introduction to data visualization principles and best practices
- Tools for data visualization: Matplotlib, Seaborn, Tableau
- Creating various types of charts and plots: Bar, line, scatter, and heat maps
- Performing exploratory data analysis (EDA) to identify patterns and insights
- Visualizing and interpreting data distributions and relationships
- Practical exercise: Building visualizations to explore a dataset
- Case study: Using EDA to uncover business insights

Day 4: Introduction to Machine Learning and Algorithms

- Overview of machine learning: Supervised vs unsupervised learning
- Introduction to regression, classification, and clustering algorithms
- Building a simple linear regression model
- Decision trees, k-NN, k-means clustering, and random forests
- Model evaluation metrics: Accuracy, precision, recall, and F1 score
- Practical exercise: Building and evaluating a machine learning model
- Case study: Predicting customer behavior using machine learning

Day 5: Big Data Technologies and Advanced Analytics

- Introduction to Big Data technologies: Hadoop, Spark, NoSQL databases
- Working with large datasets using Hadoop and Spark
- Parallel processing and distributed computing with Spark
- Introduction to advanced analytics: Natural language processing (NLP) and deep learning
- Applying big data tools to analyze and process unstructured data
- Practical exercise: Using Spark for large-scale data processing
- Case study: Applying big data analytics to solve business problems

DOCUMENTATION

The **MTC team** has meticulously prepared **high-quality training materials** for distribution to all delegates.

CERTIFICATES

An **accredited Certificate of Completion** will be awarded to participants who successfully attend and complete the program.

SCHEDULE

Course sessions are scheduled as follows:

- **Morning Session:** 09:00 AM – 1:00 PM
- **Afternoon Session:** 01:00 PM – 05:00 PM

REGISTRATION & PAYMENT

To register, please complete the **registration form** available on the course page and submit it with your **preferred payment method**. Alternatively, you can contact us via **email or WhatsApp** for assistance.

TRAVEL & TRANSPORT

We ensure a **seamless travel experience** by providing **airport-hotel-airport** transfers for all participants.