BIG DATA USAGE AND BUSINESS PRACTICES

COURSE OVERVIEW

The course provides participants with a clear and practical understanding of what big data is and why it plays a critical role in modern business success. The course covers how large and diverse datasets are collected, stored, processed, analyzed, and explored to support smarter decisions, enhance performance, and uncover new opportunities. Participants will learn how organizations transform raw data into valuable insights that improve customer understanding, operational efficiency, and strategic planning. In addition, the course introduces essential concepts in big data processing, large-scale analytics, and data visualization to help participants interpret and communicate insights effectively. By the end of the course, learners will gain foundational knowledge and practical skills to confidently engage in big data initiatives and contribute to data-driven business improvement.

TARGET COMPETENCIES

- Big Data Fundamentals
- Data Collection and preprocessing
- Exploratory and statistical Data Analysis
- Machine Learning Fundamentals
- Predictive Analytics
- Data Visualization
- Big Data Technologies

TARGET AUDIENCE

This course is ideal for individuals who are looking to transition their careers into the field of data science and analytics. It is suitable for professionals from non-technical backgrounds interested in exploring data science as a career. Business analysts aiming to enhance their analytical capabilities. IT professionals interested in expanding their skill set to include big data technologies.

COURSE OBJECTIVES

By completely attending this course, participants will be able to:

- Understand the fundamental concepts of big data analytics and its significance in various industries.
- Utilize essential tools and technologies for data collection, storage, and preprocessing.
- Perform exploratory data analysis to uncover patterns, trends, and outliers within datasets.
- Apply statistical methods and machine learning techniques for predictive and prescriptive analysis.
- Communicate data-driven insights effectively through data visualization and storytelling.
- Collaborate in a team-based environment to solve real-world data challenges.
- Develop critical thinking and problem-solving skills specific to big data scenarios.

NOTE

The course requires using laptop running on Microsoft Windows OS with Excel and Power BI Desktop fully installed and with connection to internet. Participants must bring their own laptops.

COURSE METHODOLOGY

The course will be delivered through a combination of instructor-led lectures and demonstrations. Hands-on lab sessions to practice concepts using real-world datasets. Group discussions and case studies for practical application. Assignments and projects to reinforce learning.

COURSE OUTLINE

INTRODUCTION TO BIG DATA ANALYTICS

- Understanding the big data landscape.
- Role of data scientists in different industries.
- Exploring the data science workflow.

DATA COLLECTION AND PREPROCESSING

- Data sources and types.
- Data cleaning and quality assessment.
- Data transformation and feature engineering.

EXPLORATORY DATA ANALYSIS

- Descriptive statistics and data distribution.
- Data visualization techniques.
- Identifying patterns and outliers.

STATISTICAL ANALYSIS FOR DATA SCIENCE

- Introduction to basic statistical concepts.
- Hypothesis testing and p-values.
- Correlation and causation.

INTRODUCTION TO MACHINE LEARNING

- Supervised vs. unsupervised learning.
- Overview of popular machine learning algorithms.
- Model selection and evaluation.

PREDICTIVE ANALYTICS WITH REGRESSION

- Linear regression and its applications.
- Logistic regression for classification.
- Model evaluation metrics.

DATA VISUALIZATION AND COMMUNICATION

- Principles of effective data visualization.
- Creating plots and charts using Python libraries.
- Storytelling with data.

INTRODUCTION TO BIG DATA TECHNOLOGIES

- Overview of Hadoop and MapReduce.
- Introduction to Apache Spark and distributed computing.
- Handling big data challenges.

FUTURE TREND IN DATA ANALYTICS

- Real-time Analytics Growth
- AI-Driven Data Processing
- Expansion of Data Privacy & Governance