

BIG DATA USAGE AND BUSINESS PRACTICES

COURSE OVERVIEW

The Introduction to Big Data Analytics course is designed for individuals seeking a career transition into the field of data science and analytics. This comprehensive course provides a solid foundation in the principles, techniques, and tools used in big data analytics. Participants will gain hands-on experience with industry-standard tools and technologies to process, analyze, and derive insights from large and complex datasets. By the end of the course, participants will be equipped with the essential knowledge and skills to embark on a successful journey as a data scientist.

TARGET COMPETENCIES

- Data Collection and Preprocessing: Understanding data sources, data cleaning, transformation, and feature engineering.
- Exploratory Data Analysis: Descriptive statistics, data visualization, and identifying data quality issues.
- Statistical Analysis: Hypothesis testing, confidence intervals, and correlation analysis.
- Machine Learning Fundamentals: Introduction to supervised and unsupervised learning algorithms.
- Predictive Analytics: Regression models, classification models, and model evaluation techniques.
- Data Visualization: Creating meaningful visual representations of data using tools like Matplotlib and Seaborn.
- Big Data Technologies: Introduction to Hadoop, Spark, and distributed computing concepts.
- Ethics and Privacy: Understanding ethical considerations and data privacy in big data analytics.

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.

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.

ETHICS AND PRIVACY IN DATA ANALYTICS

- Ethical considerations in data collection and analysis.
- Data privacy regulations and compliance.
- Responsible data handling practices.

CAPSTONE PROJECT

- Applying learned concepts to solve a real-world data challenge.
- Data analysis, interpretation, and presentation.
- Collaborative team project.

To register or for complete course information

Office: +971 4 430 8394 | WhatsApp: +971 50 454 9895 | Email: courses@viftraining.com

web: www.viftraining.com