AFRICAN ADVANCED LEVEL TELECOMMUNICATIONS INSTITUTE (AFRALTI)

COURSE OUTLINE

Title:	**Big Data Analytics (BDA)**
Duration:	**5 Days**

OVERVIEW:

Big Data analytics allow organizations to make better business decisions by deriving value from vast amounts of untapped data. This course provides the knowledge and skills to build competitive strategies around data-driven insights. Further, you leverage an overall lifecycle approach using sophisticated data modeling within your Big Data analytics projects.

TARGET AUDIENCE:

This course is intended for technical and database professionals, managers, data analysts, data scientists, businesses analysts and database professionals. The course will be useful for those professional involved in forecasting and trends management.

PRE-REQUISITE:

Programming and a background in statistics is helpful, but not required.

LEARNING OUTCOMES

- Optimize business decisions and create competitive advantage with Big Data analytics
- Perform predictive analysis using the appropriate data mining approach
- Use Python to create code that reads data from sensors and stores it in a SQL database.
- Visualize, clean, manipulate and integrate data sets.
- Learn fundamental principles of Big Data platforms like Hadoop.
- Use storytelling to present insights gained from extracted data.
- Derive business benefit from unstructured data
- Apply the data analytics life cycle approach to Big Data

COURSE CONTENTS:

Big Data & Analytics

Course Outline

Course Introduction
Big Data and Analytics
A Global Community
More Than Just Information
Course Overview

Chapter 1: Data and the Internet of Things
1.0 Introduction
1.0.1 Welcome
1.0.1.1 Chapter 1: Data and the Internet of Things
1.1 Value of Data
1.1.1 The Data Aspect of a Connected World
1.1.1.1 Unique Role of Data
1.1.1.2 What is Data?
1.1.2 Data is Growing Exponentially
1.1.2.1 Estimating Exponential Growth
1.1.2.2 Growth of Data
1.1.2.3 Lab – Demonstrate the Growth of Data with VNI
1.1.3 Data Growth Changes Our Lives
1.1.3.1 Data Growth Impact
1.1.3.2 Business Example: Kaggle
1.1.3.3 Social Example: DrivenData
1.1.3.4 Environmental Example: Planetary Skin Institute
1.2 Data and Big Data
1.2.1 Where Does Big Data Come From
1.2.1.1 Defining Big Data
1.2.1.2 Big Data Characteristics
1.2.1.3 Sources of Big Data
1.2.1.4 Real-World Examples of Big Data Sources
1.2.2 Open Data and Private Data
1.2.2.1 What is Open Data?
1.2.2.2 What is Private Data?
1.2.2.3 Lab – Explore Sources of Open Data
1.2.2.4 Lab – Where is My Data?
1.2.3 Structured and Unstructured Data

Lab Activities

Chapter 1 Data and the Internet of Things
1.1 Value of Data
1.1.2 Data is Growing Exponentially
1.1.2.3 Lab – Demonstrate the Growth of Data with VNI
1.2 Data and Big Data
1.2.2 Open Data and Private Data
1.2.2.3 Lab – Explore Sources of Open Data
1.2.2.4 Lab – Where is My Data?
1.3 Managing Big Data
1.3.2 Basic Data Management Technologies
1.3.2.6 Lab – Demonstrate Spreadsheet Limitations in Data Analysis
1.3.2.8 Labs – Introduction to PL-App
1.3.2.9 Lab – Take the Python Challenge

Chapter 2 Fundamentals of Data Analysis
2.2 Using Big Data
2.2.2 Types of Data Analysis
2.2.2.5 Lab – Basic Data Analysis
2.2.4 Data Analysis Lifecycle
2.2.4.5 Lab – San Francisco Crime
2.3 Data Acquisition and Preparation
2.3.2 Data Preparation
2.3.2.7 Lab – Preparing Data
2.4 Big Data Ethics
2.4.1 What are the Ethical Concerns?
2.4.1.5 Lab - Big Data Ethics
2.5 Preparation for Chapter 2 Internet Meter Labs
2.5.1 Part 1
2.5.1.4 Lab - Internet Meter Data Analysis
2.5.2 Part 2
2.5.2.4 Lab – Working with Python and SQLite
2.5.2.5 Lab – Internet Meter SQL

Chapter 3 Data Analysis
3.1 Analyzing Data
3.1.4 Analysis Using Descriptive Statistics
3.1.4.6 Lab - Descriptive Statistics in Python
3.1.5 Analysis Using Correlation
3.1.5.5 Lab – Correlation Analysis in Python
3.2 Preparation for Chapter 3 Internet Meter Lab
3.2.1 Basic Analysis with pandas.
3.2.1.6 Lab – Internet Meter Visualization

Chapter 4 Advanced Data Analytics and Machine Learning
4.1 Predictive Analytics
4.1.2 Regression

4.1.2.4 Lab – Simple Linear Regression in Python
4.1.3 Classification
4.1.3.5 Lab – Decision Tree Classification
4.2 Model Evaluation
4.2.2 Error in Analyses
4.2.2.5 Lab – Evaluating Fit Errors in Linear Regression
4.3 Preparation for Chapter 4 Labs
4.3.1 Regression and Prediction Lab
4.3.1.4 Lab - Internet Meter Linear Regression
4.3.2 Lab - Internet Meter Anomaly Detection
4.3.2.4 Lab - Internet Meter Anomaly Detection

Chapter 5 Storytelling with Data
5.2 The Power of Visualization
5.2.3 Choosing the Right Visualization for the Job
5.2.3.8 Lab - Visualizing Data in Excel
5.3 Preparation for Chapter 5 Labs
5.3.1 Chapter 5 Labs
5.3.1.4 Lab - Advanced Data Visualization
5.3.1.5 Lab - Internet Speed Compliance

Chapter 6 Architecture for Big Data and Data Engineering
6.1 Scaling Data Analytics
6.1.3 The Virtualized Data Center
6.1.3.4 Lab – Install a Virtual Machine on a Personal Computer
6.4 The Image Processing Labs
6.4.1 Digital Images as Data
6.4.1.2 Lab - Install and Test the Raspberry Pi Camera
6.4.1.3 Lab - Image Processing Change Detection
6.4.1.4 Lab – Smile Detection

For more information, please contact us on
Tel: +254 710 207 061, + 254 733 444 421
Email: training@afralti.org