**Annexure -III**

**Detailed Curriculum**

**Name of Unit of Qualification** : Configure Deployment Platform

**Duration** : 15 Hours

**Topics** : Ubuntu

| Outcome | Contents | Hrs. |
|---|---|---|
| Preparation of platform for analysing big data | Introduction to Virtual Machine, creating and configuring Virtual Machine, Installing Ubuntu Operating System on Virtual Machine | 3 |
| Acquiring skills to interact with prepared platform | Operating System Concepts: Linux History, Benefits of Linux, Different Flavors of Linux, Introducing Ubuntu, Installing Ubuntu: Starting Up, Logging in, Exploring the Desktop, Ubuntu Basics, Browsing the File System, Understanding File System Concept, Managing Files, Real and Virtual Files, Mounting, File Searches, File Size, File Space<br>Understanding Linux Files/Directories: Viewing Text Files, Using a Command Line Text Editor, Creating Files, Searching through Files, Comparing Text Files, Copying, Moving, Managing Files.<br>Ubuntu Commands, Running Basic commands, Piping and Filtering Commands, Directory and File handling commands | 6 |
| Acquiring skills to secure files by managing users, groups and their privilege | Users, Groups and Permissions, Root and Other Users, Adding and Deleting Uses and Groups, Adding and Changing Passwords, Users and File Permissions<br>Managing and Handling Processes: Viewing Processes, Controlling Processes, Controlling Jobs, Background and Foreground Jobs.<br>Scheduling Tasks, Installing and Removing Software | 6 |

**Name of Unit of Qualification** : Analyse and Define Business Requirements

**Duration** : 15 Hours

**Topics** : Data Base Management Systems and MySQL

| Performance Criteria(OUTCOME) No. | Contents | Hrs. |
|---|---|---|
| Acquiring knowledge on Data and Data Analysis | Introduction to data, data analysis and data analyst. Difference between Qualitative data and quantitative data, Data analysis as a process and as a cycle. Data summarization and visualization, Data analysis techniques: Linear regression, classification, clustering, decision tree, probability and time series analysis, | 3 |
| Selection of database based on Requirements | Introduction to database, characteristics of data in database, DBMS, advantages of DBMS, file-oriented approach versus Database-oriented approach to Data Management, disadvantages of file- oriented approach. A brief overview of relational model. Definition of relation, properties of relational model, Concept of keys: candidate key, primary key, alternate key, Foreign key, Fundamental integrity rules: entity integrity, referential integrity. | 1 |
| Acquiring the skills on designing database | Entity-relationship model as a tool for conceptual design-entities, attributes and relationships, ER-diagram, Cardinality Ratio, strong and weak entities, converting an E-R model into relational schema, Examples of E-R Model | 3 |

| | | |
|---|---|---|
| Acquiring the skills to manipulate data | SQL, characteristics & Advantages of SQL, SQL data types, SQL constructs: select-from-where, insert, delete, and update. SQL constructs: group by, having, order by. | 3 |
| Acquiring the advanced skills to manipulate data | Nested queries, joins, union, intersection, correlated nested queries, views and indexes. Practical hands on SQL statements using MySQL | 3 |
| Acquiring the skills to manage large scale data warehouse and eliciting hidden information | Knowledge Discovery in Databases, Data Mining, Data warehouse. Migrating data from source to data warehouse, cleaning, aggregation operations. | 2 |

**Name of Unit of Qualification** : Design and Develop Presentation Layer
**Duration** : 45 Hours
**Topics** : Java Programming

| Learning Outcome (NO) | Topics | Hours |
|---|---|---|
| Acquiring fundamental software developing skills | OOPS Principles, an Overview of Java Object-Oriented Programming, Data Types, Variables, and Arrays, Operators-Arithmetic Operators, The Bitwise Operators, Relational Operators, Boolean, Logical Operators, Programming Constructs, Methods and Inheritance | 9 |
| Acquiring skills on handling unusual situations at runtime | The basic Java I/O Classes and String Handling<br>Exception-Handling Fundamentals, Exception Types, Uncaught Exceptions, Using try and catch, displaying a Description of an Exception, Multiple catch Clauses, Nested try Statements, Throw throws finally Java's Built-in Exceptions | 6 |
| Acquiring skills on development software with latest practices | Packages, Access Protection, Importing Packages and Interfaces | 6 |
| Acquiring skills on architecture of front-end application | Applet Fundamentals Applet Architecture an Applet Skeleton Applet Initialization and Termination, Simple Applet Display, | 6 |
| Acquiring skills on developing front-end application | Controls: Labels, TextField and Buttons. Handling Buttons and TextField and basic Layout Managers | 6 |
| Attaining skills on integrating application with backend database | Java Database Connectivity JDBC-ODBC Bridge JDBC Drivers Creating DSN DriverManager, Connection, Statement, ResultSet, ODBC Database URL Statement and it usage with Applet | 12 |

**Name of Unit of Qualification** : Analyse Big Data in Cluster Environment

**Duration** : 30 Hours

**Topics** :Hadoop and Map Reduce Programming

| Performance Criteria(OUTCOME) No. | Contents | Hrs. |
|---|---|---|
| Acquiring skills on platform preparation for managing big data | Big Data Concepts, need for analyzing Big Data, its roles in Business Intelligence and decision making.<br>Big Data, Hadoop Architecture, Hadoop ecosystem components, storage, Hadoop Distributed File System (HDFS), Single node installation. | 6 |
| Acquiring skills on platform preparation for managing big data in grid | Multi node installations. Cluster Architecture, Cluster configuration files Hadoop commands, Hadoop Server Role, name Node, secondary node, data node, file write and read. | 6 |
| Acquiring skills on interacting with big data file system | Shell commands, Accessing files on HDFS and local machine | 3 |
| Acquiring skills on analysing data using conventional style of programming | Map Reduce Framework, Examples, Developing Map Reduce Programs, structure of Map Reduce program, Map and Reduce Tasks, Life cycle methods, Data types, data loading, Map and Reduce Tasks, partitioner, combiners, input formats, output formats, Custom input format, Error Handling, Tuning Advance Map Reduce, Fair and Capacity Scheduler. Running Map Reduce on local and Distributed modes. | 9 |
| Acquiring advance skills on analysing data using conventional style of programming | Sorting, reverse sorting, secondary sorting, Compression Techniques, Working with sequential files, | 6 |

**Name of Unit of Qualification** : Analyze Data using Big Data Analytic Tools

**Duration** : 60 Hours

**Topics** : Big Data Analysis using HBase, HIVE and PIG

| Performance Criteria (OUTCOME) No. | Contents | Hrs. |
|---|---|---|
| Use of data warehouse facility for analysing big data | Introduction to HIVE, installing HIVE, Data types, HIVE shell, HIVE commands, Complex Data types, | 9 |
| Use of Programming language to Analyse big data stored at data warehouse | UDF in Hive using Java | 6 |
| Use of column database for analysing big data | HBASE introduction and installation in Ubuntu, integration with Hadoop, HBase Shell, HBase storage techniques, HBase-Java connectivity, | 9 |
| Use of programming language to analyse stored in column database | Writing programs in Java using HBase to handle big data | 6 |
| Use of high level tool to analyse big data | PIG installation and configuration in Ubuntu, data types, commands: group, filter, order, Distinct, Join, union. | 9 |
| Use of programming language to analyse big data stored in high level tool | UDF in Pig using Java | 6 |
| Use of Big Data Analytic tool to analyse semi-structured/unstructured data | JAQL data model, Jaql syntax, jaqlshell, core operators, input / output adapter, jaql build in function, jaql statement: assignment, explains, import, quit etc. | 9 |
| Use of programming language to analyse Semi Structured/ unstructured data | Embedding jaql in java | 6 |

**Name of Unit of Qualification** : Manage Real World Data Analytic Application
**Duration** : 60 Hours

**Topics** : Project

| Performance Criteria (OUTCOME) No. | Contents | Hrs. |
|---|---|---|
| Identify big data Requirements | Identification of requirements of analytics | 9 |
| Document big data requirements | Requirement Analysis Preparation | 9 |
| Design big data application | Design of Real-World Data Analytic Application | 9 |
| Develop big data application | Develop Real World Data Analytic Application | 18 |
| Test big data application | Test Real World Data Analytic Application | 9 |
| Steps to implement the developed application | Implement Real World Data Analytic Application | 3 |
| Demonstrate big data application | Demonstration | 3 |

**Name of Unit of Qualification** : Enhancing Communication Skill

**Duration** : 15 Hours

**Topics** : Soft Skills and Communication

| Performance Criteria(OUTCOME) No. | Contents | Hrs. |
|---|---|---|
| Acquiring Communication Skill | Communication, verbal and non-verbal communication | 5 |
| Managing career, staff and professional relationships | Building professional relationship, Relationship at work , Making the most of personal and professional relationships, Competency Description, Managing Difficult Business Relationships | 5 |
| Preparing for interview | **Interview Techniques**: Planning For The Interview, Preparing for an Interview, Interview Formats, Stages Of The Interview, Types Of Interview Questions<br><br>**Best Bet for Interview Preparation**: Mock Interviews, The Benefits of Mock Interviews Experience & Skills,<br><br>**Curriculum Vitae**:Overview, types of CV, Covering letter, Writing a Resume, Acceptance Letter, Thank You Letter | 5 |