ANNEXURE -1
Detailed syllabus : I016

Name of the Group: Data Science

Course Name: Certified Data Scientist

Course Code:

Module No	Topic	Details
1	Configure Deployment Platform	Introduction to Virtual Machine, creating and configuring Virtual Machine, Installing Ubuntu Operating System on Virtual Machine 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 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. Ubuntu Commands, Running Basic commands, Piping and Filtering Commands, Directory and File handling commands Users, Groups and Permissions, Root and Other Users, Adding and Deleting Uses and Groups, Adding and Changing Passwords, Users and File Permissions Managing and Handling Processes: Viewing Processes, Controlling Processes, Controlling Jobs, Background and Foreground Jobs. Scheduling Tasks, Installing and Removing Software
2	Analyze and Define Business 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. 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 Normalization concepts in relational model, normal forms (1NF, 2NF, 3NF) SQL, characteristics & Advantages of SQL, SQL data types, SQL constructs: select-from-where, insert, delete, and update. SQL constructs: group by, having, order by. Nested queries, joins, union, intersection, correlated nested queries, views and indexes. Practical hands on SQL statements using MySQL/Oracle database .Knowledge Discovery in Databases, Data Mining, Data warehouse. Migrating data from source to data warehouse, cleaning, aggregation operations.
3	Design and Develop Presentation Layer	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 .The basic Java I/O Classes and String Handling ,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 ,Packages, Access Protection, Importing Packages and Interfaces Applet Fundamentals Applet Architecture An Applet Skeleton Applet Initialization and Termination ,Simple Applet Display , Controls: Labels , TextField and Buttons .Handling Buttons and

		TextField and basic Layout Managers Java Database Connectivity JDBC-ODBC Bridge JDBC Drivers Creating DSN DriverManager, Connection, Statement, ResultSet ,ODBC Database URL Statement and it usage with Applet
4	Analyze Big Data in Cluster Environment	Big Data Concepts, Need for analyzing Big Data, its roles in Business Intelligence and decision making. Big Data, Hadoop Architecture, Hadoop ecosystem components, storage, Hadoop Distributed File System (HDFS), Single node installation. Multi node installations. Cluster Architecture, Cluster configuration files Hadoop commands, Hadoop Server Role, name Node, secondary node, data node, file write and read.Shell commands, Accessing files on HDFS and local machine 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. Sorting, reverse sorting, secondary sorting, Compression Techniques, Working with sequential files.
5	Analyze Data using Big Data Analytic tools	Introduction to HIVE, installing HIVE, Data types, HIVE shell, HIVE commands, Complex Data types, UDF in Hive using Java HBASE introduction and installation in Ubuntu, integration with Hadoop, HBase Shell, HBase storage techniques, HBase-Java connectivity, Writing programs in Java using HBase to handle big data PIG installation and configuration in Ubuntu, data types, commands: group, filter, order, Distinct, Join, union. UDF in Pig using Java JAQL data model, Jaql syntax, jaqlshell, core operators, input /

		output adapter, jaql build in function, jaql statement: assignment, explains, import, quit etc. Embedding jaql in java
6	Manage Real World Data Analytic	Identification of requirements of analytics Requirement Analysis Preparation Design of Real World Data Analytic Application Develop Real World Data Analytic Application Test Real World Data Analytic Application Implement Real World Data Analytic Application Demonstration
7	Enhancing Communication Skill	Communication , verbal and non-verbal communication Building professional relationship, Relationship at work , Making the most of personal and professional relationships, Competency Description, Managing Difficult Business Relationships Interview Techniques: Planning For The Interview, Preparing for an Interview, Interview Formats, Stages Of The Interview, Types Of Interview Questions Best Bet for Interview Preparation: Mock Interviews, The Benefits of Mock Interviews Experience & Skills, Curriculum Vitae: Overview, types of CV, Covering letter, Writing a Resume, Acceptance Letter, Thank You Letter