COURSE PROSPECTUS

Name of the Group: IT

Name of the Course: Junior Data Analyst

Course Code: QG-4.5-IT-00349-2023-V1-NIELIT

Starting Date: 15/07/2024 (Tentative)

Duration: 540 Hours

Course Coordinator: Ms. Manjiri Lavadkar, Project Engineer, Nielit Auarangabad.

Course Description:

The Junior Data Analyst Qualification is meant to develop the front end and Implementation of use cases of Data Analytics. The Data Analyst Associate Qualification is meant to inculcate the basic skills of data handling, processing and visualization. Inculcate the basic skills of Hadoop Framework and programming experience of core Java. Training in different tools of Hadoop like Map Reduce, Hive will increase the exposure to the analytics field.

Course Objectives:

The course aims to equip students with the skills necessary to perform fundamental arithmetic and statistical calculations using spreadsheet software. Students will learn how to effectively organize, store, and manipulate structured data, ensuring data integrity and accessibility. Additionally, the course will guide students through the installation and configuration of the Hadoop framework and Java programming environment, laying the groundwork for big data processing.

Course Outcome:

- ✓ Execute fundamental arithmetic and statistical calculations in spread sheet applications, utilizing functions and formulas to process and analyse data efficiently.
- ✓ Create and interpret graphical data visualizations such as charts, graphs, and dashboards to effectively communicate data insights and trends.
- ✓ Design experiments to evaluate and compare different machine learning techniques on real world problems.
- ✓ Successfully install and configure the Hadoop framework, alongside Java programming setup, to support big data processing and management.
- ✓ Design and implement front-end applications, employing web development technologies and best practices to create user-friendly interfaces.
- ✓ Apply data analytics techniques to real-world scenarios, solving practical problems by implementing use cases and presenting actionable insights derived from data analysis.

Expected Job Roles:

- Database Administrator
- Data analyst
- Data Engineer
- Data Architect
- Data Scientist

Course Structure:

Module	Topics	HRs
Perform basic calculation Using spreadsheet	Features of Spreadsheet Creating and saving worksheet and workbook Layouts, text formats alignment. Basic functions and formulas, sorting Graphs	30
Manage structured data	Introduction to database, advantages of DBMS, Concept of keys: candidate key, primary key, alternate key, Foreign key, Fundamental integrity rules: entity integrity, referential integrity. Entity-relationship model. ER-diagram, SQL, Knowledge Discovery in Databases, Data Mining, Data warehouse. Migrating data from source to data warehouse, cleaning, aggregation operations. Knowledge Discovery in Databases, Data Mining, Data warehouse. Migrating data from source to data warehouse, cleaning, aggregation operations.	30
Analyses data using spreadsheet tool	Data, table & formatting Conditional formatting Advanced charting Pivot Table. Data Validation Filtering, What if analysis, Role of Maths and Statistical techniques in Data Science, Probability andStatistics, Regression analysis, Descriptive statistics, Linear Regression	90
Manage data in Open source tool	Introduction to Virtual Machine, creating and configuring Virtual Machine, 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 Viewing Text Files, Using a Command Line Text Editor, Creating Files, Searching through Files, Comparing Text Files, Copying, Moving, Managing Files.	30

Visualize data graphically	Introduction Tableau , Connecting to Excel, CSV Text Files,Connecting to Databases, Analyzing,Formatting,Sorting and charts	60
Installation of Hadoop Framework and Java Programming	Installation of Hadoop, Java concepts, Oops concepts, Looping techniques, Methods,Method Overloading, Method Over-riding, Arrays, Java Programming HDFS, Architecture of Hadoop, Configuring Hadoop Cluster & Hadoop Commands.	30
Manage big data using hadoop	Big Data Concepts, Need for analyzing Big Data, its roles in Business Intelligence and decision making. Analysis of Data Using Mapreduce and Hive. Implementation with examples	60
Front end application development Front end application development	Advanced Java, Applets, Swings - Front end Java Database Connectivity, JDBC-ODBC Bridge, Creating JDBC Drivers, DSN DriverManager, Connection, Statement, Result Set, ODBC Database, URL Statement and itsuse with Applets	60
Implementation of use cases of Data Analytics	Student will be to implement various use cases in data analytics.	60
Employability Skills	Students will be able to get the additional skills apart from the technical skills, to be jobready	90

Other Contents:

I. **Course Fees: Course** fee is Rs 24,300+ GST (* Nil for SC/ST Candidates)
II. **Registration Fee:** An amount of Rs.1000/- (including all taxes as applicable) (non-refundable) should be paid at the time of registering for the course.
III. **Course Fee Instalment Structure: -** Can be paid in two instalments
IV. **Eligibility:**
- Completed 1st year of UG
- Pursuing 1st year of UG and continuous education
- Pursuing 3rd year of 3-year diploma after 10th and continuous education
- Completed 3-year diploma after 10
- Completed 2nd year of 2 year diploma after 12th
- Pursuing 2nd year of 2- year diploma after 12 and continuous education
- 10th Grade pass plus 3 years of vocational education & Training
- Previous relevant Qualification of NSQF Level 4 and with minimum education as 8th Grade pass

V. **Number of Seats :30**

VI. **Selection of candidates:** The candidates passed in the qualifying examination will be based on their marks obtained, subject to eligibility and availability of seats.

VII. **Important Dates:**

Starting Date for Registration	**01/07/2024**
Last date to submit application form:	**10/07/2024(Tentative)**
Counselling/Admission	**11/08/2024(Tentative)**
Last Date for Payment of Fee	**15/08/2024(Tentative)**
Commencement of class work:	**15/07/2024 (Tentative)**

VIII. **Course Timings:** 3:00 Hrs. in week days (Mon-Fri).

IX. **Lab Facilities:** LIST OF EQUIPMENT (For a batch of 30 students)

Sr. NO.	Description	Qty
1	Classroom	2
2	Student Chair	30
3	Student Table	15
4	Smart Interactive Display	2
5	White Board	2
6	Desktop computer with Accessories: installed with: Microsoft Excel Mysql Ubuntu Tableau Java Hadoop	30