Information Booklet cum Syllabus

Of

Data Analyst Trainee

April 2026

National Institute of Electronics and Information Technology

An Autonomous Scientific Society under

Ministry of Electronics and Information Technology, Government of India

NIELIT Gorakhpur

M.M.M.U.T. Campus, Deoria Road

Gorakhpur (U.P.) -273010

CONTENTS

1. About Course

This course provides a strong foundation in Data Science, Relational Databases, and Statistical Analysis. Students gain hands-on experience with modern analytical tools such as SQL, Power BI, and Python through practical implementation and real-world projects.

By the end of the program, learners will be able to perform complex data cleaning, build interactive business intelligence dashboards, and execute exploratory data analysis to drive decision-making. This course equips students with the essential skills required to pursue a career in Data Analytics and Business Intelligence.

2. NIELIT

National Institute of Electronics and Information Technology, NIELIT, (Erstwhile DOEACC Society) is an autonomous scientific society of the Ministry of Electronics& Information Technology, Government of India. The Society is registered under the Societies Registration Act, 1860. NIELIT was set up to carry out Human Resource Development and related activities in the area of Information, Electronics & Communications Technology (IECT). NIELIT is engaged both in Formal & Non-Formal Education in the areas of IECT besides development of industry oriented quality education and training programmes in the state-of-the-art areas. NIELIT has endeavored to establish standards to be the country's premier institution for Examination and Certification in the field of IECT. It is also one of the National Examination Body, which accredits institutes/organizations for conducting courses in IT and Electronics in the non-formal sector.

3. Objective of Course

The objective of this course is to equip learners with strong foundations in **Statistical Analysis, Relational Database Management, and Business Intelligence**, along with hands-on experience in building automated data workflows and interactive dashboards for real-world decision-making.

After completing the course, learners will be able to:

- **Develop a strong understanding** of data lifecycle fundamentals, including data collection, cleaning, exploratory analysis (EDA), and interpretation.
- **Master Advanced Spreadsheet Logic** to perform complex data manipulation using logical functions, array formulas, and dynamic pivot tables.
- **Write Optimized SQL Queries** to extract and transform data from relational databases using joins, aggregations, and subqueries.
- **Understand Statistical Modeling** and the application of descriptive and inferential statistics to validate data trends and patterns.
- **Implement Data Cleaning Pipelines** using Python (Pandas/NumPy) to handle missing values, outliers, and inconsistent data formats.
- **Build Interactive Dashboards** using Power BI or Tableau, incorporating DAX/calculated fields, slicers, and drill-through capabilities.
- **Design Exploratory Data Analysis (EDA) Reports** that utilize Matplotlib and Seaborn for advanced visual storytelling.
- **Apply Predictive Analytics** basics to forecast trends using linear regression and time-series analysis.

4. Job Roles of Course

This course is designed to equip learners with the necessary skills for the following job roles:

- Junior Data Analyst
- Business Intelligence (BI) Trainee
- Data Associate
- Reporting Analyst
- Data Analyst Trainee

5. Eligibility

Pursing Diploma /Graduation

6. Total duration of the course

90 Hours (Theory: 35 Hrs, Practical/Tutorial: 55 Hrs)

7. Course Details

7.1. Course Outline and Objective of Each Unit

S. No.	Module Name	Duration (Theory) in Hours	Duration (Practical) in Hours	Total Learning Hrs.	Learning Objectives
1	**Foundations of Data & Advanced Excel**	5	10	15	• Understand the Data Lifecycle and Business Intelligence basics. • Master advanced logical functions (XLOOKUP, IFS, Nested IF). • Use Pivot Tables and Power Query for data cleaning. • Apply descriptive statistics to summarize business data.
2	**Data Analysis & SQL**	10	12	22	• Understanding Data analytics Techniques • Data exploration and statistical analysis. • Understand RDBMS concepts, schemas, and primary/foreign keys. • Write efficient queries using SELECT, WHERE, and ORDER BY. • Implement complex Joins and Aggregations (GROUP BY, HAVING). • Use subqueries and CTEs to solve business logic problems.
3	**Data Visualization & BI**	5	10	15	• Connect BI tools (Power BI) to multiple data sources. • Build data models and establish relationships between tables. • Create interactive dashboards using slicers, filters, and DAX. • Apply visual best practices for data related reporting.

4 **7**	**Python for Data Analysis**	**10**	**15**	**25**	• Understand Python syntax, data types, and list comprehensions. • Perform data manipulation and cleaning using **Pandas** and **NumPy**. • Create statistical visualizations using **Matplotlib** and **Seaborn** . • Handle missing data and outliers programmatically.
5	**Career Practicum & Capstone Project**	**5**	**8**	**13**	• Execute an end-to-end data project (Cleaning → Analysis → Insights). • Understand data ethics, privacy, and governance standards. • Practice data storytelling and insight presentation techniques.

7.2. Detailed Course

Module Name	Unit	Contents	Hrs.
Foundations of Data & Advanced Excel	**Data Analytics Lifecycle & Stats**	• Understanding Descriptive, Diagnostic, Predictive, and Prescriptive analytics • **Business Statistics:** Mean, Median, Mode, Standard Deviation, and identifying Outliers	**15**
	Professional Spread sheet Mastery	• **Advanced Logic:** XLOOKUP, INDEX/MATCH, Nested IF, and IFS functions • Power Query for automated cleaning, Pivot Tables, and Data Validation.	
Data Analysis & SQL	**Data Analysis Techniques**	• Introduction to descriptive, diagnostic, predictive, and prescriptive analytics. • Techniques for identifying patterns, trends, and anomalies in datasets. • Applying basic statistical measures to summarize data and drive decision-making. • Using analytical frameworks to translate raw data into actionable business insights.	**22**
	SQL Querying & Data Extraction	• **RDBMS Basics:** Primary/Foreign Keys, Schemas, and ACID properties. • Mastering SELECT, JOINs, GROUP BY, and Subqueries for complex data retrieval.	

Data Visualization & BI	**Business Intelligence & Storytelling**	• Import/export data using PowerBI. • Star Schemas, DAX/Calculated fields, and interactive filtering. • Building drill-down reports in Power BI to communicate insights.	**15**
Python for Data Analysis	**Python Programming Foundations**	• Variables, Lists, Dictionaries, and Control Flow (Loops/Conditionals). • Writing custom functions for repetitive data tasks.	**25**
	Data Science Libraries (Pandas & NumPy)	• Vectorized operations with NumPy and DataFrame cleaning with Pandas. • Advanced charting with Matplotlib and Seaborn for trend discovery	
Career Practicum & Capstone Project	**Applied Project & Portfolio**	• Developing an end-to-end analytical project from a raw dataset. • Data Ethics, and professional presentation skills.	**13**

8. Reference Books/ Study Materials

1. Python for Data Analysis – Wes McKinney
2. SQL for Data Analysis: Advanced Techniques for Transforming Data – Cathy Tanimura
3. Storytelling with Data: A Data Visualization Guide for Business Professionals – Cole Nussbaumer Knaflic
4. Practical Statistics for Data Scientists – Peter Bruce & Andrew Bruce
5. Microsoft Excel Data Analysis and Business Modeling – Wayne Winston

9. Practical Assignments

1. Create a student grade sheet in Excel that calculates the total and average marks for five subjects, and use conditional formatting to highlight scores below 40 in red.

2. Build a simple "Product Search" tool in Excel where entering a Product ID uses the VLOOKUP or XLOOKUP function to display the item name and its current price from a separate list.

3. Write a Python script that asks a user to input their monthly salary and monthly rent, then calculates what percentage of their income goes toward rent.

4. Use Python to create a basic calculator that can add, subtract, multiply, and divide two numbers based on user input.

5. Write a SQL query using SELECT and WHERE to list all customers from a "Sales" table who live in a specific city and have made a purchase in the last 30 days.
6. Load a small CSV file containing a list of employees into Python using the Pandas library and print the first 10 rows and the total number of employees in the list.
7. Use Pandas to find and remove any duplicate rows in a dataset of "Store Orders" and fill in any blank "Discount" cells with a value of 0.
8. Create a simple Bar Chart in Excel or Python that compares the monthly sales of four different branches of a shop to identify which branch performed the best.
9. Build a basic one-page dashboard in Power BI or Tableau that shows "Total Revenue" as a big number and a "Sales by Category" chart that changes when you select a specific month.
10. Use a dataset of "Daily Steps" for one month to calculate the average steps taken per day and create a Line Chart to show if the person is becoming more active over time.