Website : www.pineappleeducation.com

PINEAPPLE EDUCATION

Innovation, Research, Academic & Training Institute with Placement Support

(Mentored by Lawyers, Chartered Accountants, IT Professionals & Academicians)

21 Years

Powered by IIMMI INTERNATIONAL INSTITUTE OF MANAGEMENT, MEDIA & IT

PROFESSIONAL DATA SCIENCE PROGRAM | CAREER-FOCUSED CURRICULUM

Learning Roadmap

A focused path through the complete 10-module Data Science curriculum.

8 Months	10 Modules	Projects + Career Prep
Program Duration	Complete Curriculum	Applied Learning

01 Computer Fundamentals & Career Orientation Foundation	**02** Microsoft Excel (Basic to Advanced) Analytics Tooling
03 SQL (Beginner to Advanced) Databases & Querying	**04** Python for Data Science Programming & Analysis
05 Statistics Statistical Foundations	**06** Power BI Business Intelligence
07 Data Science & Machine Learning Modeling & Deployment	**08** Real-World Projects Applied Practice
09 Interview Preparation Career Readiness	**10** Capstone Project End-to-End Delivery

8-Month Training Plan

Month-wise schedule showing what is covered and what you build.

Month	Focus	Modules	Hands-on Output
Month 1	Computer Fundamentals + Excel	Modules 01, 02	Excel dashboard on Office Employee dataset
Month 2	SQL (Beginner to Advanced)	Module 03	HR, Sales and Banking SQL projects
Month 3	Python for Data Science	Module 04	Cleaning + EDA notebooks (Pandas, Seaborn)
Month 4	Statistics	Module 05	Hypothesis testing and A/B testing case study
Month 5	Power BI + Real-World Projects	Modules 06, 08	KPI dashboards (Sales, HR, Retail, Finance)
Month 6	Data Science & ML (Part 1)	Module 07	Regression and classification models, Churn analysis
Month 7	Data Science & ML (Part 2)	Modules 07, 08	Tuning, Time Series, Flask/FastAPI deployment
Month 8	Interview Prep + Capstone	Modules 09, 10	Capstone, final presentation, GitHub portfolio

Tools Covered

Microsoft Excel	**MySQL / PostgreSQL**	**Python**	**Pandas & NumPy**
Matplotlib & Seaborn	**Power BI**	**Git & GitHub**	**Flask / FastAPI**

MODULE 01

01

Computer Fundamentals & Career Orientation

Foundation-level orientation to data science, essential computing concepts, software setup, and career readiness.

Topics Covered

- Introduction to Data Science & Analytics
- Data, Information, and Business Intelligence
- Data Analyst vs Data Scientist vs Data Engineer
- SDLC and Data Life Cycle
- Types and Categories of Data
- Common Data File Formats (CSV, Excel, JSON, XML, Parquet)

Software to Install

- Microsoft Excel
- MySQL or PostgreSQL
- Python
- Visual Studio Code / Jupyter Notebook
- Power BI
- Git and GitHub

Assignment

- Install and verify all required software
- Create and configure a GitHub account
- Download and organize practice datasets from Kaggle

02

MODULE 02

Microsoft Excel (Basic to Advanced)

Spreadsheet analysis, formula development, data preparation, pivot reporting, and dashboard-oriented Excel skills.

Part 1: Excel Basics

- Interface, Workbooks & Worksheets
- Data Entry & Formatting
- Relative, Absolute, Mixed Referencing
- Number Formats, Borders, Fill Series
- Freeze Panes, Hide / Unhide, Printing
- Practice: Office Employee Dataset

Part 2: Basic Formulas

- SUM, AVERAGE, COUNT, COUNTA
- MIN, MAX
- ROUND, ROUNDUP, ROUNDDOWN
- SUBTOTAL, AGGREGATE

Part 3: Logical Functions

- IF, Nested IF, IFS
- AND, OR, NOT
- IF with AND / IF with OR
- Hands-on practice questions

Part 4: Text Functions

- LEFT, RIGHT, MID, LEN, TRIM
- CONCAT, CONCATENATE, TEXTJOIN
- UPPER, LOWER, PROPER
- FIND, SEARCH
- SUBSTITUTE, REPLACE, TEXT

Part 5: Date Functions

- TODAY, NOW, DATE
- DAY, MONTH, YEAR
- DATEDIF, EDATE, EOMONTH
- NETWORKDAYS, WORKDAY, WEEKDAY

Part 6: Lookup Functions

- VLOOKUP, HLOOKUP, XLOOKUP
- INDEX, MATCH
- INDEX + MATCH
- Dynamic Lookup Techniques

Part 7: Conditional Functions

- SUMIF, SUMIFS
- COUNTIF, COUNTIFS
- AVERAGEIF, AVERAGEIFS

Part 8: Data Analysis Tools

- Sorting, Filtering, Advanced Filter
- Remove Duplicates
- Text to Columns, Flash Fill
- Data Validation
- Conditional Formatting

Part 9: Pivot Tables

- Pivot Tables & Pivot Charts
- Slicers & Timelines
- Dashboard Design

Part 10: Advanced Excel

- Dynamic Arrays: FILTER, SORT, UNIQUE, SEQUENCE
- Named Ranges
- Error Handling
- Power Query
- Basic VBA & Macros

03

MODULE 03

SQL (Beginner to Advanced)

Relational fundamentals and query syntax through joins, window functions, optimization, and applied SQL projects.

Phase 1: Database Fundamentals

- Data vs. Information
- DBMS & RDBMS
- Database Architecture
- Keys, Constraints, Schema
- Normalization
- ACID Properties

Phase 2: SQL Basics

- CREATE, ALTER, DROP, TRUNCATE
- INSERT, UPDATE, DELETE
- SELECT, WHERE, ORDER BY, LIMIT
- DISTINCT, LIKE, IN, BETWEEN
- NULL Handling, CASE

Phase 3: Aggregate Functions

- COUNT, SUM, AVG, MIN, MAX
- GROUP BY
- HAVING

Phase 4: Joins

- INNER JOIN, LEFT JOIN
- RIGHT JOIN, FULL JOIN
- SELF JOIN
- CROSS JOIN

Phase 5: Intermediate SQL

- Subqueries, Correlated Subqueries
- CTE & Recursive CTE
- UNION, UNION ALL
- INTERSECT, EXCEPT

Phase 6: Window Functions

- ROW_NUMBER, RANK, DENSE_RANK
- NTILE
- LEAD, LAG
- FIRST_VALUE, LAST_VALUE

Phase 7: Advanced SQL

- Views & Indexes
- Stored Procedures & Functions
- Transactions
- Query Optimization
- Performance Tuning

SQL Projects

- HR Database
- Sales Analysis
- Customer Retention
- Banking Database
- Hospital Database

GOAL

QUERY

Build accurate business queries using joins, CTEs, aggregates and window functions.

VALIDATE

Check logic, edge cases and performance so results are reliable.

EXPLAIN

Turn query outputs into clear insights, decisions and actions.

04

MODULE 04

Python for Data Science

Practical Python programming for data handling, analysis, visualization, and reusable analytical workflows.

Python Basics

- Variables, Data Types, Operators
- Strings
- Lists, Tuples
- Dictionaries, Sets

Control Flow

- IF, ELIF, ELSE
- Loops
- Break, Continue, Pass

Functions

- User-Defined Functions
- Lambda Functions
- Map, Filter, Reduce

OOP

- Classes & Objects
- Constructors
- Inheritance
- Polymorphism

File Handling

- Read & Write Files
- CSV, Excel, JSON Files

Exception Handling

- Try
- Except
- Finally

NumPy

- Arrays
- Broadcasting
- Mathematical Operations

Pandas

- Series, DataFrame
- Merge, GroupBy, Pivot
- Missing Values
- Data Cleaning

Matplotlib & Seaborn

- Line, Bar, Pie Charts
- Histogram, Scatter Plot
- Heatmap, Pairplot
- Boxplot, Violin Plot

ANALYTICAL WORKFLOW

01 LOAD
Import trusted data sources.

02 CLEAN
Fix types, duplicates, missing values.

03 EXPLORE
Find patterns and outliers.

04 VISUALIZE
Choose clear charts.

05 EXPLAIN
Link evidence to action.

05

MODULE 05

Statistics

Statistical foundations to interpret data, evaluate relationships, test assumptions, and support evidence-based decisions.

Descriptive Statistics

- Mean, Median, Mode
- Variance
- Standard Deviation
- Percentiles, Quartiles

Probability & Sampling

- Probability
- Sampling
- Inferential Statistics

Testing & Relationships

- Hypothesis Testing
- Correlation
- Regression
- A/B Testing

WHAT THIS MODULE SUPPORTS

SUMMARIZE

Select the right measures of centre, spread, percentiles and probability to describe data accurately.

COMPARE

Evaluate groups and relationships using sampling, hypothesis tests, correlation, regression and A/B testing.

DECIDE & COMMUNICATE

Interpret uncertainty responsibly, separate evidence from assumption, and state clear conclusions.

06

MODULE 06

Power BI

Business intelligence reporting, data modeling, DAX, interactive visualization, and secure dashboard publication.

Data Preparation

- Power BI Interface
- Data Import
- Power Query
- Data Cleaning

Modeling & DAX

- Data Modeling & Relationships
- Star Schema
- DAX Basics
- Time Intelligence

Reporting & Publishing

- Visualizations
- KPI Dashboard
- Drill-through Analysis
- Bookmarks
- Publishing Reports
- Row-Level Security (RLS)

BUSINESS INTELLIGENCE OUTCOME

PREPARE & MODEL

Clean data in Power Query and build a clear star-schema model.

MEASURE & ANALYZE

Create DAX measures, time intelligence and KPIs.

VISUALIZE & PUBLISH

Design dashboards, apply RLS and share insights.

07

MODULE 07

Data Science & Machine Learning

Applied machine learning spanning exploratory analysis, feature engineering, model building, evaluation, and deployment.

Data Preparation

- Exploratory Data Analysis (EDA)
- Feature Engineering
- Train-Test Split
- Cross-Validation

Supervised Learning

- Linear Regression
- Logistic Regression
- Decision Tree
- Random Forest
- KNN, SVM
- Naive Bayes

Unsupervised Learning

- K-Means Clustering

Evaluation & Tuning

- Model Evaluation Metrics
- Hyperparameter Tuning

Time Series

- Time Series Analysis
- Trend & Seasonality
- Forecasting Basics

Deployment

- Model Deployment (Flask / FastAPI)
- Packaging a model behind an API

WHAT THIS MODULE SUPPORTS

BUILD

Engineer features and train regression, classification and clustering models.

EVALUATE

Validate with the right metrics and tune for reliable results.

DEPLOY

Serve trained models through an API for real applications.

08

MODULE 08

Real-World Projects

Applied portfolio work reinforcing business analysis, modeling, insight generation, and presentation skills.

Projects 1 - 5

- HR Analytics Dashboard
- Sales Dashboard
- Finance Dashboard
- Customer Churn Analysis (ML)
- Retail Dashboard

Projects 6 - 10

- E-commerce Dashboard
- Hospital Dashboard
- Banking Dashboard
- IPL Analysis
- Netflix Dashboard

PORTFOLIO PRACTICE

Build sector-based projects across HR, sales, finance, retail, e-commerce, healthcare, banking, sports and media. Each project emphasizes KPI selection, analysis or modeling, insight generation and presentation.

09

MODULE 09

Interview Preparation

Structured preparation for technical interviews, business case discussions, portfolio presentation, resume development, and mock interviews.

Technical Preparation

- Excel Interview Questions
- SQL Interview Questions
- Python Coding Questions
- Power BI Interview Questions
- Statistics Questions
- Machine Learning Questions

Career Readiness

- Case Studies
- Resume Building
- Mock Interviews
- Portfolio Presentation Practice
- Explain Projects in Business Language

INTERVIEW GOAL

PROBLEM

Business question, audience, data context and success metric.

METHOD

Tool choice, data preparation, analysis steps and why.

RESULT

Quantified findings, KPIs and business impact.

REFLECTION

Limitations, trade-offs and next action.

10

MODULE 10

Capstone Project

A complete end-to-end project integrating data preparation, SQL analysis, Python EDA, modeling, dashboard development, insights, and professional documentation.

Stages 1 - 4

- 01 Data Collection
- 02 Data Cleaning
- 03 SQL Analysis
- 04 Python EDA & Machine Learning Model

Stages 5 - 8

- 05 Power BI Dashboard
- 06 Business Insights
- 07 Final Presentation
- 08 GitHub Documentation

PROGRAM OUTCOME

END-TO-END DELIVERY

Combine cleaning, SQL, Python EDA, modeling and Power BI in one workflow.

DECISION-READY INSIGHTS

Turn findings into clear conclusions and recommendations.

PORTFOLIO-READY

Finish with a dashboard, presentation and organized GitHub repo.

This curriculum is designed for an 8-month professional Data Science training program, progressing from fundamentals to industry-level projects and interview preparation.