

## ✓ 6-Month (26-Week) Data Science & Analytics Program

By Alpha IT Managed Services, Mohali

Plot ITC 15, IT Park, Sector 67, Mohali, Punjab

---

### Module 1: Advanced Excel for Data Analytics (Weeks 1–2)

#### Week 1

- Excel interface, data cleaning, data types
- Lookup functions: VLOOKUP, HLOOKUP, XLOOKUP
- Text, date, logical functions (IF–AND–OR)
- Conditional formatting & data rules
- Pivot tables, slicers & pivot charts

#### Week 2

- Advanced formulas (INDEX–MATCH, SUMIFS, COUNTIFS)
  - What-If analysis: Goal Seek, Scenario Manager
  - Power Query fundamentals
  - Building KPI dashboards
  - Excel Analytics Mini Project
- 

### Module 2: Python for Data Science (Weeks 3–6)

#### Week 3: Python Foundations

- Python basics: variables, data types, operators
- Conditional statements, loops
- Working with lists, sets, tuples, dictionaries
- Functions & lambda functions
- File handling & JSON/CSV reading

#### Week 4: Data Analysis with NumPy

- Arrays, indexing, slicing
- Vectorized operations
- Broadcasting & reshaping
- Aggregations & mathematical operations
- Working with numeric datasets

#### Week 5: Pandas for Data Analysis

- DataFrames, cleaning, handling missing values
- Filtering, sorting, merging, joining
- GroupBy, pivot tables, aggregations
- Creating reusable preprocessing pipelines
- Mini project: Data cleaning & preparation

#### Week 6: Visualization with Python

- Matplotlib charts: bar, line, scatter
- Seaborn: heatmaps, pairplots, boxplots
- Outlier detection through visuals
- Creating storytelling visuals
- Python Visualization Project

---

## Module 3: SQL for Data Analytics (Weeks 7–8)

### Week 7: SQL Foundations

- Installing MySQL / SQL basics
- SELECT queries, WHERE, ORDER BY
- Joins (INNER/LEFT/RIGHT)
- DISTINCT, LIMIT, GROUP BY
- Hands-on analysis using sample datasets

### Week 8: Advanced SQL

- Subqueries & nested queries
  - Window functions (ROW\_NUMBER, RANK)
  - CTEs, Views
  - Case statements for reporting
  - SQL Analytics Project
- 

## Module 4: Statistics & Probability for Data Science (Weeks 9–10)

### Week 9: Descriptive & Inferential Statistics

- Mean, median, mode, variance
- Outliers (IQR, Z-score)
- Sampling & distributions
- Confidence intervals
- Hypothesis testing (Z-test, t-test)

### Week 10: Probability for Data Science

- Probability basics
  - Bayes theorem
  - Probability distributions (normal, binomial, poisson)
  - Correlation vs causation
  - Statistics mini project (EDA + stats interpretation)
- 

## Module 5: Tableau for Data Visualization (Weeks 11–12)

### Week 11

- Tableau interface & data connection
- Dimensions vs Measures
- Filters, sorting, groups, hierarchies
- Basic charts: bar, line, area, maps
- Publishing your first dashboard

### Week 12

- Calculated fields & LOD expressions
- Dual-axis charts & advanced visuals
- Parameter controls & interactivity
- Dashboard design principles
- Tableau Analytics Project

---

## Module 6: Power BI for Business Intelligence (Weeks 13–14)

### Week 13

- Power BI Desktop overview
- Data loading, cleaning with Power Query
- Creating relationships & data modeling
- Basic visuals & formatting
- Slicers, filters, bookmarks

### Week 14

- DAX basics (CALCULATE, SUMX, FILTER)
  - Advanced visuals & dashboard layout
  - KPI design & storytelling reports
  - Publishing to Power BI Service
  - Power BI Project
- 

## Module 7: Machine Learning for Data Science (Weeks 15–20)

### Week 15: ML Foundations

- ML workflow: train–test–validate
- Data preprocessing pipeline
- Label encoding, scaling
- Feature selection basics
- Evaluating ML results

### Week 16: Regression Models

- Linear Regression
- Polynomial Regression
- Regularization (Ridge, Lasso)
- Evaluating regression models
- Sales forecasting mini project

### Week 17: Classification Models

- Logistic Regression
- K-Nearest Neighbors
- Naive Bayes
- Decision Trees
- Classification metrics (F1, ROC-AUC)

### Week 18: Advanced ML for Analytics

- Random Forest
- Gradient Boosting
- XGBoost (analytics-level)
- Hyperparameter tuning
- Customer churn project

### Week 19: Clustering & Dimensionality Reduction

- K-Means clustering
- Hierarchical clustering
- PCA for dimensionality reduction

- Silhouette score
- Customer segmentation project

### **Week 20: Model Deployment (Basic Level)**

- Saving models (pickle/joblib)
  - Loading predictions in Python
  - Introduction to Streamlit dashboards
  - Creating simple ML apps
  - Deployment mini project
- 

## **Module 8: Real-World Data Analytics Projects (Weeks 21–26)**

### **Week 21: Capstone 1 – Excel/SQL**

- Problem definition
- Data cleaning & preparation
- KPI creation
- Dashboard development
- Final presentation

### **Week 22: Capstone 2 – Tableau**

- Storyboard planning
- Visual design
- Multi-sheet dashboard
- Filters, interactivity
- Publishing & sharing

### **Week 23: Capstone 3 – Power BI**

- Problem framing
- ETL workflow in Power BI
- DAX-based KPIs
- Dashboard creation
- Insights & business decisions

### **Week 24: Capstone 4 – Python EDA**

- Data acquisition
- End-to-end EDA
- Statistical insights
- Visualization storytelling
- Summary report

### **Week 25: Capstone 5 – Machine Learning**

- Feature engineering
- Model comparison
- Final model selection
- Interpretation of results
- Documentation

### **Week 26: Portfolio & Career Preparation**

- Resume building (Data Analyst / Data Scientist)
- GitHub portfolio setup
- Mock interview preparation
- Case study practice

- Final certification