

# Data Science & Data Analytics — 6-Month Project-Based Curriculum

*Code the Future*

**Format:** Live, instructor-led | **Duration:** 6 Months | **Mode:** Online, hands-on, project-driven

**Outcome:** Job-ready Data Analyst / Data Scientist portfolio with 6+ real-world projects.

## MONTH 1 — Foundations: Python, Statistics & Excel

*Goal: Build the core programming and statistical thinking needed for all later modules.*

| Week   | Topics                                                                                 |
|--------|----------------------------------------------------------------------------------------|
| Week 1 | Python basics — variables, data types, loops, functions, control flow                  |
| Week 2 | Data structures (lists, dicts, tuples, sets), file handling, NumPy arrays              |
| Week 3 | Descriptive statistics, probability basics, distributions, hypothesis testing          |
| Week 4 | Excel for analysts — pivot tables, formulas (VLOOKUP/XLOOKUP, INDEX-MATCH), dashboards |

**Project:** Excel-based Sales/Retail Dashboard with pivot tables and KPI tracking

**Tools:** Python, Jupyter Notebook, NumPy, MS Excel

## MONTH 2 — Data Wrangling, SQL & Databases

*Goal: Learn to extract, clean, and manipulate data at scale.*

| Week   | Topics                                                                                         |
|--------|------------------------------------------------------------------------------------------------|
| Week 5 | SQL fundamentals — SELECT, JOINS, GROUP BY, subqueries                                         |
| Week 6 | Advanced SQL — window functions, CTEs, stored procedures, query optimization                   |
| Week 7 | Data cleaning with Pandas — missing values, duplicates, outliers, data types                   |
| Week 8 | Data wrangling — merging, reshaping, feature engineering, EDA with Pandas & Matplotlib/Seaborn |

**Project:** End-to-end SQL + Python data cleaning pipeline on a messy real-world dataset (e.g., e-commerce or HR data)

**Tools:** MySQL/PostgreSQL, Pandas, Seaborn, Matplotlib

## MONTH 3 — Data Visualization & Business Intelligence

*Goal: Turn cleaned data into decision-ready visual stories.*

| Week    | Topics                                                                    |
|---------|---------------------------------------------------------------------------|
| Week 9  | Power BI — data modeling, DAX basics, interactive dashboards              |
| Week 10 | Tableau — calculated fields, filters, storytelling with data              |
| Week 11 | Advanced visualization — dashboards for executives, KPI design principles |

|         |                                                             |
|---------|-------------------------------------------------------------|
| Week 12 | Storytelling with data, presenting insights to stakeholders |
|---------|-------------------------------------------------------------|

**Project:** Interactive Power BI/Tableau dashboard for a business case (e.g., sales performance, marketing analytics, or finance KPIs)

**Tools:** Power BI, Tableau

## MONTH 4 — Machine Learning Fundamentals

*Goal: Understand and apply core supervised & unsupervised ML algorithms.*

| Week    | Topics                                                                              |
|---------|-------------------------------------------------------------------------------------|
| Week 13 | ML overview, supervised vs unsupervised, train/test split, model evaluation metrics |
| Week 14 | Regression — linear, multiple, polynomial regression                                |
| Week 15 | Classification — logistic regression, decision trees, random forest, KNN            |
| Week 16 | Clustering & dimensionality reduction — K-means, hierarchical clustering, PCA       |

**Project:** Predictive model (e.g., customer churn prediction or house price prediction) with full ML pipeline

**Tools:** Scikit-learn, Pandas, Matplotlib

## MONTH 5 — Advanced Analytics, Deep Learning Intro & Model Deployment

*Goal: Go beyond basic ML into advanced techniques and real deployment.*

| Week    | Topics                                                                                                                          |
|---------|---------------------------------------------------------------------------------------------------------------------------------|
| Week 17 | Model tuning — cross-validation, hyperparameter tuning (GridSearch/RandomSearch), ensemble methods (XGBoost, Gradient Boosting) |
| Week 18 | Introduction to Deep Learning — neural network basics, TensorFlow/Keras intro                                                   |
| Week 19 | NLP basics — text preprocessing, sentiment analysis, intro to LLMs/AI tools in analytics                                        |
| Week 20 | Model deployment — Flask/Streamlit app, deploying models via APIs, intro to cloud (AWS/GCP basics)                              |

**Project:** Deployed ML web app (e.g., sentiment analysis tool or churn predictor with a live Streamlit interface)

**Tools:** TensorFlow/Keras, Streamlit/Flask, Git, basic cloud deployment

## MONTH 6 — Capstone Project, Portfolio & Career Preparation

*Goal: Consolidate all skills into an industry-grade capstone and prepare for job placement.*

| Week    | Topics                                                                                        |
|---------|-----------------------------------------------------------------------------------------------|
| Week 21 | Capstone project kickoff — problem selection, dataset sourcing, planning                      |
| Week 22 | Capstone development — data pipeline, modeling/dashboarding, iteration with mentor feedback   |
| Week 23 | Capstone finalization — documentation, GitHub portfolio setup, resume & LinkedIn optimization |

|         |                                                                                            |
|---------|--------------------------------------------------------------------------------------------|
| Week 24 | Mock interviews, case-study rounds, project presentation day, placement assistance kickoff |
|---------|--------------------------------------------------------------------------------------------|

**Project:** End-to-end Data Science/Analytics project chosen by the learner (e.g., business intelligence dashboard + predictive model combined), presented to a panel

**Tools:** All tools covered in the program

## Skills Covered End-to-End

Python | SQL | Excel | Pandas/NumPy | Power BI | Tableau | Statistics & Probability | Machine Learning | Deep Learning Basics | NLP Basics | Model Deployment | Git/GitHub | Data Storytelling

## Deliverables by Program End

- 6 hands-on projects (portfolio-ready, GitHub hosted)
- 1 industry-grade capstone project with live presentation
- Resume, LinkedIn, and GitHub profile built for job applications
- Mock interview experience and placement support

## Suggested Weekly Structure

- 2 live instructor-led sessions per week
- 1 doubt-clearing/mentorship session per week
- Weekly assignment + monthly project submission
- Peer code reviews and collaborative learning sessions

*Note: This curriculum is a template and can be adjusted in pacing/depth based on batch level (beginner vs. working professional), and whether the track is positioned as "Data Analytics" (Months 1–3 emphasis) or full "Data Science" (all 6 months with ML/DL depth).*