PY-DS: Python for Data Science

Course Overview

This Python for Data Science (PY-DS) course outline provides a complete roadmap for mastering data manipulation, cleaning, and visualization using industry-standard libraries. It typically covers programming fundamentals, NumPy, Pandas, and data visualization tools, culminating in a practical capstone project

Course Description

This course equips you with the foundational skills to extract, clean, analyze, and visualize data using Python. You will transition from writing basic scripts to building full data analysis pipelines using industry-standard libraries like NumPy, Pandas, Matplotlib, and Seaborn. Through hands-on projects, you will learn how to turn raw, messy data into actionable business insights and predictive models.

Learning Curriculum

Module 1

Python Foundations

Basic Syntax: Variables, data types (strings, integers, floats, booleans), and operators.Control Structures: if-else statements, for and while loops.Functions & OOP: Creating custom functions (with arguments/keywords), lambda functions, and basic class structures.

Module 2

Python Data Structures & File Handling

Core Structures: Lists, dictionaries, tuples, and sets.Comprehensions: Using list and dictionary comprehensions for efficient code.File I/O: Reading and writing .csv, .txt, and .json files

Module 3

Data Analysis & Manipulation (NumPy & Pandas)

NumPy: Introduction to arrays, vectorization, and mathematical operations (mean, median, standard deviation).Pandas Series & DataFrames: Data loading, inspection, and indexing.Data Preparation: Handling missing values, filtering, grouping (groupby), merging, and reshaping messy datasets.

Module 4

Data Visualization & Storytelling

Matplotlib: Creating foundational charts like line graphs, scatter plots, and histograms.Seaborn: Advanced statistical graphics, correlation heatmaps, and pair plots

Module 5

Applied Data Science & Machine Learning

Scikit-Learn Basics: Building, training, and evaluating predictive models (e.g., Linear Regression, Classification).Exploratory Data Analysis (EDA) Pipeline: Applying the data science lifecycle from data cleaning to evaluation

© 2026 Proasia Upskilling Sdn Bhd | www.proasia.my