DATAMARCOS

LIVE ONLINE · MENTOR-LED PROGRAM

# Data Science & AI with GenAI

24 weeks · 8–10 hours a week · Part-time, live online

A 24-week, part-time live program that takes you from first steps in Python to statistics, machine learning, an introduction to deep learning, and practical GenAI — including prompting, RAG and a simple agent — ending in a capstone you can show. Built for a wide audience: no prior programming is required. This is a data-science path, not a BI-only analytics course and not a production LLM engineering program.

**Fees and EMI options shared on a counsellor call.**

Program guide · Curriculum, projects and format. No prices are listed in this brochure.

Noida One, 303-304, Rasoolpur Nawada, Industrial Area, Sector 62, Noida, Uttar Pradesh 201309, India

www.datamarcos.com · hello@datamarcos.com · +91 88519 30644

# About this program

## Who it is for

- Career starters and career-switchers who want a structured path into data science and applied AI
- Analysts who already work with spreadsheets or dashboards and want to add statistics, machine learning and GenAI
- Professionals who want to go beyond using ChatGPT and learn to analyse data, build models and apply LLMs carefully
- Learners who can commit 8–10 hours a week for 24 weeks, joining live from anywhere

## Who it is not for

- People who mainly want spreadsheets, SQL and BI dashboards. That shorter path is Data Analytics with AI (16 weeks)
- Engineers who already want to ship production RAG systems, multi-agent apps or LLMOps. See GenAI & Agentic AI Engineer or LLMOps & AI Evaluation
- People seeking placement or salary guarantees — we do not offer those

## Starting point

- No prior programming or machine-learning experience is required; Python is taught from the start
- Comfort with school-level maths (percentages, averages and simple graphs) is helpful
- You will need a laptop that can run a browser and a Python environment, plus 8–10 hours a week

# What you will be able to do

These are skills and work products from the program — not placement, salary or promotion promises.

- 1. Write Python to clean, explore and visualise a dataset, and keep a reproducible notebook
- 2. Use basic statistics honestly: describe uncertainty, and avoid over-claiming from a single number
- 3. Train supervised and unsupervised models, split data properly, and evaluate them with the right metrics
- 4. Explain, in plain language, what a model can and cannot support for a decision
- 5. Build a small GenAI prototype: a prompted workflow, a simple RAG path, and a simple tool-using agent with a human review step
- 6. Leave with a capstone that combines analysis or a model with a written narrative — not a placement promise

# Projects

## Project 1: Cleaned dataset and exploratory brief

A documented, cleaned table, charts and a question the data can actually support. This is the foundation for later modelling.

## Project 2: Evaluated machine-learning model

A supervised model with a proper split, a baseline, metrics, error analysis and a written statement of limits.

## Project 3: Capstone: analysis, model and optional GenAI

A reproducible notebook, a short written narrative, and — if it earns its place — a simple RAG or agent prototype with an evaluation note. This is coursework, not a job guarantee.

DATAMARCOS

# Curriculum · 24 weeks

8–10 hours a week, live and mentor-led. Open each week in the live sessions for labs and feedback.

## Week 1: Python start for data work

- Notebooks, variables, types and control flow
- Reading and writing tabular files
- How a script differs from a spreadsheet, and why that helps
- Using an assistant to explain an error, then fixing it yourself

Lab: Load a sample table in a notebook and print a sound overview.
You ship: A starting notebook with checks.

## Week 2: Functions, data structures and tidy code

- Lists, dictionaries and functions you can reuse
- Reading errors and writing small tests
- Keeping a notebook runnable from top to bottom
- A short personal coding checklist

Lab: Refactor messy exploration code into small functions.
You ship: A tidier notebook you can rerun.

## Week 3: Wrangling with pandas

- Selecting, filtering and deriving columns
- Missing values, types, duplicates and joins
- A reproducible cleaning log, not ad-hoc edits
- When a spreadsheet is enough, and when Python is better

Lab: Clean a messy table and record every step.
You ship: A wrangled dataframe and a cleaning log.

## Week 4: Exploratory analysis and charts

- Asking a precise question before you plot
- Charts for comparisons, trends and distributions
- Labelling, units and honest scales
- A one-sentence claim that the chart actually supports

Lab: Build a small chart set and drop any chart that does not earn its place.
You ship: An exploratory chart pack tied to your question.

## Week 5: SQL enough to get the data out

- SELECT, WHERE, JOIN and GROUP BY for analysis work
- Checking a query result against a known sample
- When to stay in SQL and when to continue in Python
- This is enough SQL to feed a data-science workflow, not a full BI-dashboard program

Lab: Pull a joined, aggregated table and continue the analysis in pandas.
You ship: A SQL-plus-Python analysis slice.

# Curriculum (continued)

## Week 6: Phase checkpoint: a question and a clean dataset

- Putting the question, the cleaned table and the first charts together
- Peer and mentor feedback
- What statistics will add next

Lab: Present your brief and cleaned data.
You ship: A phase-one packet ready for statistical work.

## Week 7: Descriptive statistics that stay honest

- Averages, spreads and simple distributions, in plain language
- Why a single headline number is often misleading
- Comparisons across groups without over-claiming
- Choosing a summary that matches the question

Lab: Compute and chart a small set of descriptive measures with caveats.

## Week 8: Probability, sampling and bias

- Samples versus populations, in practical terms
- Common sampling and selection problems
- Independence and why it is often a hopeful assumption
- What your dataset cannot represent

Lab: Write down how the data were collected and what that allows.

## Week 9: Inference and uncertainty

- Confidence intervals as a range, not a ritual
- Standard errors in plain language
- What 'significant' does and does not mean
- Reporting uncertainty next to a point estimate

Lab: Add an interval or a clear uncertainty statement to one claim.

## Week 10: Comparisons, tests and a statistics checkpoint

- A careful group comparison or A/B-style readout
- Multiple testing and fishing for a result
- Peer and mentor review
- What machine learning will add — and what it will not replace

Lab: Present a statistical comparison and the checks you ran.

## Week 11: What machine learning can and cannot do

- Prediction versus explanation
- Train, validation and test splits, and leakage
- A baseline before a fancy model
- Choosing a first supervised problem from your data

Lab: Define a target, a split and a simple baseline.
You ship: A problem statement and a baseline score.

# Curriculum (continued)

## Week 12: Supervised learning: regression

- Linear models and when they are enough
- Features, scaling and regularisation at a practical level
- Residual plots and obvious misspecification
- Comparing a model with the baseline

Lab: Train a regression model and inspect residuals.
You ship: A regression notebook with a fair comparison.

## Week 13: Supervised learning: classification

- Logistic regression and a tree-based model as a comparison
- Class imbalance, in practical terms
- Decision thresholds and who is affected by an error
- A first classification report

Lab: Train two classifiers and compare them on the same split.
You ship: A classification notebook with a threshold note.

## Week 14: Model evaluation that a colleague can read

- Accuracy, precision, recall, ROC and calibration, used with care
- Why a leaderboard metric can mislead
- Error analysis: looking at the cases the model gets wrong
- A written evaluation note, not only a score

Lab: Write an evaluation note and an error-analysis sample.
You ship: An evaluation report a non-specialist can follow.

## Week 15: Unsupervised learning

- Clustering and what a cluster is not
- Dimensionality reduction as a view, not a proof
- Checking stability and sanity of groupings
- When unsupervised work should stay exploratory

Lab: Try a clustering or a reduced view and stress-test it.
You ship: An unsupervised exploration note.

## Week 16: Phase checkpoint: an evaluated model

- Assembling the model, the metrics and the limits
- Peer and mentor review
- What an introduction to deep learning will and will not add

Lab: Present the model and the evaluation evidence.
You ship: A phase-three ML packet.

# Curriculum (continued)

## Week 17: Neural networks in practical terms

- Layers, loss and optimisation, without a research lecture
- Overfitting and regularisation in this setting
- When a simpler model is still the better first choice
- Setting up a small experiment

Lab: Train a small network on a familiar task and compare it with a baseline.
You ship: An introductory deep-learning notebook.

## Week 18: A small deep-learning experiment

- Transfer learning or a slightly larger architecture, used carefully
- Compute, data size and diminishing returns
- What this introduction does not cover: production training at scale
- A written note on when you would stay with classical ML

Lab: Complete a small experiment and write the comparison.
You ship: An experiment note plus the notebook.

## Week 19: LLMs and prompting for data work

- What current models are good at, and where they invent
- Prompt patterns for analysis, explanation and code drafts
- Verifying every number and citation yourself
- When not to use a model at all

Lab: Draft an analysis or a code snippet with an assistant and document what you accepted or rejected.
You ship: A prompting playbook for your capstone domain.

## Week 20: A simple RAG prototype

- Retrieval-augmented generation in outline: documents, chunks, queries, answers
- Measuring whether the retrieved context actually supports the answer
- Common failure modes: missing sources and confident invention
- This is an introductory prototype, not a production RAG service

Lab: Build a small question-answering path over a document set you control.
You ship: A simple RAG notebook with a few evaluated questions.

## Week 21: A simple tool-using agent

- When a single prompt is enough, and when a tool call helps
- A minimal agent that can call a function (for example a query or a calculator) and return a checked result
- Human review before any action that would matter at work
- What this course does not cover: multi-agent production systems

Lab: Wire a simple tool-using agent and log each step.
You ship: A small agent prototype with a review step.

# Curriculum (continued)

## Week 22: Evaluating GenAI and knowing its limits

- A small evaluation set for your prototype
- Quality, cost and when a rule or a classical model is better
- Privacy: what must not be pasted into a public tool
- Choosing what belongs in the capstone

Lab: Score the prototype on a short question set and write residual risks.
You ship: A GenAI evaluation note.

## Week 23: Build the capstone

- Choosing a scope you can finish: one question, one main method, a clear narrative
- Reproducible notebooks and a short method note
- Optional: include the GenAI prototype only if it earns its place
- Mentor review of claims versus evidence

Lab: Assemble the notebook, the figures and the first draft of the narrative.

## Week 24: Present, review and next steps

- Live capstone presentations and mentor feedback
- A written brief a hiring manager or a colleague could read without opening the notebook
- Honest next steps: more statistics, more engineering, or a later GenAI engineering program if that is the right path
- No placement or salary promises

Lab: Present the capstone and the checks you ran.

# Format, fees and refunds

Part-time and live online, in a small mentor-led cohort. Weekend-morning batches are offered for learners in the US. Evening batches (India time) are offered for learners in the Gulf and Africa. Europe-oriented groups typically meet after work and on Saturdays. Recordings are shared if you miss a live session. Exact days, times and the next cohort date are confirmed on a counsellor call.

Join live from anywhere. Weekend-morning batches are offered for learners in the US; evening batches (India time) for the Gulf and Africa; Europe-oriented groups typically meet after work and on Saturdays.

A Datamarcos certificate of completion is awarded when you finish the program work. Cohort start dates are shared on a counsellor call — we do not publish a date until then.

### Fees and EMI options shared on a counsellor call.

7-day refund policy: if the program is not right for you, ask within 7 days of enrollment.

## Mentor

Live sessions are mentor-led by Deepesh Mishra, founder of Datamarcos, a BITS Pilani alumnus, and a solution architect and corporate trainer in data and AI who has trained teams at 200+ organizations.

# Questions, answered

### How is this different from Data Analytics with AI?

Data Analytics with AI is a 16-week program focused on spreadsheets, SQL, Python for analysis, BI dashboards and careful AI assistance. Data Science & AI with GenAI is a 24-week path that goes further into statistics, machine learning, an introduction to deep learning, and practical GenAI (prompting, a simple RAG prototype and a simple agent), ending in a data-science capstone.

### How is this different from GenAI & Agentic AI Engineer?

GenAI & Agentic AI Engineer is an engineering program for people who want to build and deploy LLM applications, production RAG and multi-agent systems. This program is a data-science path: Python, statistics and machine learning come first, and GenAI is taught as a practical additional tool, not as a production engineering syllabus.

### Do I need prior programming experience?

No. Python is taught from the start. Comfort with school-level maths is helpful. You will need a laptop and 8–10 hours a week for 24 weeks.

### How long is the program?

Twenty-four weeks, part-time and live online, at about 8–10 hours a week including live sessions and labs.

### Are the sessions live?

Yes. Sessions are live and mentor-led, with recordings if you miss a class. Weekend-morning batches are offered for learners in the US, with other regional times available. Exact times are confirmed on a counsellor call.

### Will this get me a job or a salary increase?

We do not promise placement, salary or promotion outcomes. The program is designed so you can show a capstone — a notebook, a narrative and the checks you ran.

### Is this an official university or vendor certificate?

No. Completing the program leads to a Datamarcos certificate of completion. We do not award a university, government or vendor certification.

### What about fees?

Fees and EMI options shared on a counsellor call.

### Is there a refund policy?

Yes. Datamarcos has a 7-day refund policy. If the program is not right for you, ask within 7 days of enrollment.

DATAMARCOS

# Talk to a counsellor

Share a few details on datamarcos.com and a counsellor will talk through program fit, the next cohort, fees and enrollment.

### Fees and EMI options shared on a counsellor call.

7-day refund policy: if the program is not right for you, ask within 7 days of enrollment.

Noida One, 303-304, Rasoolpur Nawada, Industrial Area, Sector 62, Noida, Uttar Pradesh 201309, India

www.datamarcos.com · hello@datamarcos.com · +91 88519 30644

This brochure does not promise placement, salary or promotion outcomes. It is not legal, immigration or official-government certification advice.