

— TR-02

Data Science & Machine Learning

COURSE BRIEF

SKILLS & TOOLS YOU WILL MASTER

 Python + SQL

 Pandas + NumPy

 Power BI (DAX + Power Query)

 Streamlit + Plotly

 Scikit-learn

10 Reasons This Works.

Why students choose Saarathi – and how each card helps you ship a real career.

01

10 Max / Batch

Not 50. Not 30. Ten. Every student seen, heard, and unblocked – every class.

02

Sunday Open Classroom

Every Sunday the classroom is open. Come in, pair-program, get unstuck, or just focus.

03

Weekly Detailed Syllabus

Full syllabus from day one. Every week: what to read, what to build, what the outcome is.

04

Revision Every 5th Day

Every fifth class is a revision session. Nothing piles up or slips through.

05

Week 1: Mindset

The first week is about getting into the right headspace before touching a line of code.

06

3 Ways of Learning

Self-learning + collaboration + mentorship. All three built into every course.

07

1:1 Mock Interview

Real mock interview with your mentor in the final week. Actual hiring prep, not fluff.

08

CV + LinkedIn

Resume writing and LinkedIn optimisation. Leave with a profile that gets replies.

09

Real Projects

Every course ends with a deployed, publicly accessible project you can show employers.

10

Gate Assessment

Aptitude test before Day 1. Maps your strengths, weaknesses, learning style, and pace – then builds your personal plan.

saarathiacademy.com.np/data-science-and-machine-learning-course-in-nepal

+977-9744442469 · +977-9761095364

Data Science & Machine Learning

Why This Course

This course trains the junior data analyst / data scientist profile employers actually want: someone who can combine Python, SQL, statistical reasoning and stakeholder communication to deliver honest analysis. Employers want analysts who can clean data, write SQL, run rigorous statistics and explain findings in plain language, not tool-name collectors. Pay scales with your skills, portfolio, and whether you work locally or for international/remote clients. We don't promise a salary.

Most data science courses rush to machine learning before foundational analyst skills are solid, and skip the BI tooling employers actually hire for. This Python-first course follows the realistic analyst sequence, Python fundamentals before NumPy, SQL before modelling, EDA and visualisation, then a full Power BI week (Power Query, data modelling, DAX, Row-Level Security), then starter ML, so graduates leave with one defensible, reproducible body of work that demonstrates data reasoning from first principles to a completed project. Power BI gets a dedicated week because it is the most-named BI tool in Kathmandu job postings.

AI tools are woven into every phase the way working analyst teams actually use them in 2026: Claude for exploring messy datasets, Copilot for accelerating Pandas code and drafting DAX in Power BI (DAX generation is now generally available), ChatGPT as a rubber duck during EDA interpretation. These are not gimmicks, they are productivity tools that require judgment to use well: every AI-drafted query or measure is read and corrected before it is trusted, and this course builds that judgment.

Who Is This Course For

- **Absolute beginners** starting from the fundamentals who want a real path into data roles
- **Spreadsheet-heavy professionals** (finance, operations, marketing) ready to switch from Excel to Python
- **Fresh graduates** from any stream, commerce, engineering, stats, economics, targeting an analyst role
- **Working professionals who make data-driven decisions** and want to test claims more rigorously
- **CS or IT learners** who want applied industry work instead of academic-only theory
- **Career changers** targeting analyst or junior data scientist roles who want a structured, beginner-paced entry into Python, SQL and analytics

Course Snapshot

PARAMETER	DETAILS
Course Code	TR-02
Duration	12 Weeks (3 Months)
Schedule	Monday to Friday, 2 Hours/Day
Total Hours	120 Hours
Batch Size	Maximum 10 Students
Course Fee	NPR 35,500
Level	Beginner-friendly, taught from the fundamentals (light pre-course prep listed below) → job-ready Data / BI Analyst
Outcome	Job-ready Data Analyst / Power BI Analyst (with the foundations to grow into a Junior Data Scientist role)

Prerequisites

- No prior Python, statistics or maths required, all three start from zero in Week 1
- A laptop with 8GB+ RAM
- Comfortable using a computer daily, files and folders
- Basic English reading
- Saarathi Gate Assessment before Day 1, diagnostic, no pass or fail

Nothing to install beforehand. Python, VS Code and Jupyter go on your machine together in the first class.

Optional warm-up. None of this is required and Week 1 starts from zero, but arriving with it makes the first weeks easier:

- The first few lessons of freeCodeCamp Scientific Computing with Python, free at freecodecamp.org/learn
- Khan Academy Statistics and Probability units 1-2, and Algebra Basics (variables and equations)
- Practise 10 Google Sheets formulas: SUM, AVERAGE, IF, VLOOKUP, COUNTIF, SUMIF, MIN, MAX, SORT, FILTER

Assessments at a glance

Two exams bracket the course. The **Saarathi Gate Assessment** runs before Day 1 as a diagnostic with no pass or fail; it maps your learning style and pace. The **Saarathi Certification Exam** is sat at the end of the course, 2 hours and around 30 questions specific to this course in mixed formats (single choice, multi-select, true/false, numeric, ordering, matching, fill-in-the-blank, short or long answer, code where relevant, and short case studies). Pass mark is 60 percent and a pass is mandatory for the Saarathi Certificate of Completion. The exam is proctored: it runs in fullscreen, switching tab or app and pasting are logged, and three strikes end it. You can answer the questions in any order and come back to any of them, but every question needs an answer before you submit. Recommendation letters are discretionary, mentor-issued only to learners who pass the Certification Exam, and never guaranteed.

Your Learning Week

DAY	ACTIVITY
Mon–Fri	2-hour live class session (hands-on, project-driven) with mentor
Mon–Fri	2-hour self-study at home (assignments, practice, debugging)
Sunday	Open classroom for focused build time or rest
Day 5 (every week)	Within the 2-hour session, the last 1 hour is a built-in revision block for consolidation and doubt-clearing

Real improvement happens when you build, break, debug and re-ship between classes. The classroom gives structure; the hours between classes give depth.

Week-by-Week Curriculum

Phase 1 • Analyst Foundations & Data Handling (Weeks 1–4, 40 Hours)

The first four weeks (a third of the course) are pure base-building. Python alone gets a deliberate **1.5 weeks** (Week 1 + the first half of Week 2) before any data library, so the foundation actually holds.

WEEK	FOCUS AREA	TOPICS
Week 1	Python Fundamentals I:	- What data roles actually do (analyst, scientist, engineer) and where you fit - Setup from zero: Python 3.13, Jupyter (run code cell by cell),

WEEK	FOCUS AREA	TOPICS
	Setup & Core Language	VS Code, Git, a virtual environment and a first GitHub commit - Variables, numbers, text and operators - Decisions (<code>if / elif / else</code>) and loops (<code>for , while</code>) - Functions: parameters, return values and scope - Containers: lists (ordered), dictionaries (labelled values), tuples (fixed), sets - Spreadsheet-to-Python bridge: SUM, IF, VLOOKUP, COUNTIF and their code equivalents - Notebook hygiene and AI intro together (one talk-segment block, room to settle in): markdown cells, clean cell order, Restart-and-Run-All, debugging habits, then ChatGPT and Copilot as a rubber duck, never run unread
Week 2	Python Fundamentals II + NumPy	<i>Python (first half, completing 1.5 weeks):</i> - List and dictionary comprehensions to fluency - Full File I/O with the standard library: <code>open()</code> , context managers, <code>pathlib</code> , <code>csv.DictReader</code> - Deeper error handling: <code>try / except / else / finally</code> , reading tracebacks - Functions deeper, modules and debugging together (extra hands-on time): default args, <code>*args / **kwargs</code> , reusable modules and imports, then print vs breakpoint and isolating a bug - Pure-Python CSV mini-analyzer mini-project (read, clean, summarise, no library) <i>NumPy (back half):</i> - NumPy arrays: vectorisation (fast column math), boolean masks, one broadcasting example - Why arrays beat Python lists for numbers (speed/memory), bridging into Pandas
Week 3	Pandas & AI-Assisted Data Cleaning	- Pandas Series and DataFrames (<code>.loc</code> vs <code>.iloc</code> , dtype awareness) - Reading CSV and Excel (JSON, Parquet awareness) - Analyst Excel: pivot tables, XLOOKUP/SUMIFS, basic charts, mirrored in <code>pivot_table / groupby</code> - Filtering, sorting, groupby, aggregation - Missing values, duplicates, type conversion - String and date cleaning (<code>.str</code> , <code>to_datetime</code> , timezone awareness) - Merging and reshaping together: <code>merge</code> , <code>concat</code> , picking join keys, then <code>melt</code> (wide to long) and <code>pivot</code> (long to

WEEK	FOCUS AREA	TOPICS
		wide) - End-to-end cleaning on a curated messy dataset, with AI as a partner: Claude to explore datasets, Copilot for Pandas code, always verified
Week 4	SQL, Window Functions & Python-SQL Integration	- Relational thinking, keys, normalisation and PostgreSQL setup - Single-table reads built up slowly in one block: SELECT, WHERE, ORDER BY, LIMIT, then summaries with GROUP BY, HAVING, aggregates, <code>date_trunc</code>, CASE bucketing - INNER JOIN and LEFT JOIN to fluency (RIGHT/FULL/self/anti awareness) - CTEs (WITH, named query stages) to fluency, a core SQL-screen skill beside joins and windows; correlated/nested subqueries awareness - Three core window functions to fluency: ROW_NUMBER, RANK, LAG (deeper fluency in the Week 12 drill) - Python-SQL integration: <code>read_sql</code>, SQLAlchemy, parameterised queries, avoiding SQL injection - Multi-source ingest: the same question from CSV, Excel, JSON API and Postgres, then reconciled

Phase 2 · Statistics, EDA & Decision Framing (Weeks 5–7, 30 Hours)

WEEK	FOCUS AREA	TOPICS
Week 5	Descriptive Statistics & Distribution Thinking	- Summary numbers: mean, median, mode, variance, standard deviation, percentiles - Distribution shape: histograms, box plots, one KDE, skewness (violin/ECDF awareness) - Outlier detection (z-scores, IQR) and handling - Bivariate relationships and correlation-vs-causation together: scatterplots, relationship thinking, and why "moves together" never proves "causes" (confounders, reverse causation, selection bias) - Probability and sampling together: conditional/Bayes intuition, then sampling and bias - Central Limit Theorem intuition via simulation - Structured EDA summary habits: write down what you checked and found

WEEK	FOCUS AREA	TOPICS
Week 6	Probability, Sampling & Hypothesis Testing	- Confidence intervals: a defensible range, not false certainty (bootstrap awareness) - 2-sample t-test to fluency: real difference or chance? - Chi-square test to fluency: the same question for categories - p-values and Type I/II errors (multiple-comparison/Bonferroni awareness) - Effect size and practical significance - Mann-Whitney / ANOVA / Bayesian testing (awareness)
Week 7	EDA, Visual Storytelling & Streamlit Dashboards	- Chart choice framework: comparison, distribution, relationship, composition - Seaborn to fluency (Matplotlib fine-control awareness) - One interactive Plotly chart (choropleth awareness) - North-Star KPI with leading/lagging indicators - A/B testing fundamentals and experiment framing - Streamlit: notebook to shareable dashboard with filters and caching (deploys in Week 11) - 10-minute filter-context primer (a slicer narrows what a number sums over), priming Week 8 DAX - Power BI vs Streamlit decision matrix - Analysis narrative and recommendations with honest uncertainty, drafted with AI for EDA: ChatGPT as a rubber duck, when to trust and when to push back

Phase 3 · Business Intelligence & Starter Machine Learning (Weeks 8–10, 30 Hours)

WEEK	FOCUS AREA	TOPICS
Week 8	Power BI: End-to-End Business Intelligence	- Power Query / M for ETL: profile, clean, fix types, pivot/unpivot, merge vs append, fact + dimension tables - Star-schema modelling: relationships, cross-filter direction, marked Date table, role-playing dimensions - DAX foundations: one slow context-transition build ("CALCULATE turns the current row into a filter"), measures vs calculated columns - DAX to fluency: CALCULATE with ALL/ALLEXCEPT, SUMX, VAR/RETURN, time intelligence (TOTALYTD, SAMEPERIODLASTYEAR) - DAX awareness: calculation groups, visual calculations, RANKX/AVERAGEX, REMOVEFILTERS/KEEPFILTERS, quick measures

WEEK	FOCUS AREA	TOPICS
		- Report design: slicers, drillthrough, bookmarks, tooltips, mobile, accessibility - Power BI Service: publish, scheduled refresh, gateway, static Row-Level Security; beginners ship Import mode (DirectQuery/Direct Lake, dynamic RLS awareness); Copilot in Power BI (GA DAX generation, report-grounded chat) for drafting measures, every one read and corrected, never run unread - Revision + one published end-to-end report, capstone mentor 1:1
Week 9	ML Workflow, Regression & Classification	- ML lifecycle: split, prepare, train, evaluate; capstone topic signed off in the Week-8 Day-5 1:1 - Leakage prevention, scaling, encoding, scikit-learn Pipeline - Linear regression (predict a number) + RMSE, MAE, R² - Logistic regression (predict yes/no) + confusion matrix, precision, recall, F1 to fluency (ROC-AUC/PR-AUC awareness) - Cross-validation: k-fold to fluency (stratified/time-series awareness)
Week 10	Feature Engineering, Clustering & Model Explanation	- Tree models and class-imbalance together: decision trees and random forests + class-imbalance handling - Feature engineering: ratios, bins, date parts, interactions - Feature selection and leakage checks - K-Means clustering to fluency (elbow, silhouette, scaling); hierarchical/dendrograms awareness - Model explanation: SHAP global summary + one force plot + one PDP (permutation importance, gini pitfall awareness) - Writing limitations and caveats - One ~20-minute "what's next / lives in TR-10" slide (time series, NLP, deep learning, Spark, all awareness)

Phase 4 · Validation, Capstone & Career (Weeks 11–12, 20 Hours)

WEEK	FOCUS AREA	TOPICS
Week 11	Data Validation, Reproducibility & Capstone Sprint	- Day 1, lock: problem, dataset, success metric and a "do-nothing" baseline - Data validation and pytest together (one capstone-discipline block): completeness, consistency, freshness, uniqueness (<code>pandera</code> / <code>assert</code>), then fixtures and shape/dtype/null-rate assertions

WEEK	FOCUS AREA	TOPICS
		- Reproducible structure: <code>data/</code> , <code>src/</code> , <code>notebooks/</code> , pinned <code>requirements.txt</code> , fixed <code>random_state</code> , <code>.env</code> - Lightweight ETL thinking and delivery choice: notebook vs dashboard vs endpoint - Build days together: clean, EDA, model or experiment, then the dashboard or report; ship, peer review, demo rehearsal and final deliverables - MLOps awareness pointer (covered further in TR-10)
Week 12	Portfolio, SQL Interviews & Career Launch	- Resume and LinkedIn for data roles - GitHub portfolio and README polish - SQL case-style interview practice: deep window-function fluency (LEAD, SUM OVER running totals, moving averages, partitioned ranking) - Mock analyst + data scientist interviews - Job search planning and application strategy - Demo day presentation

Skills You Will Gain

CATEGORY	TECHNOLOGIES
Languages	Python 3, SQL (PostgreSQL)
Data Handling	pandas, NumPy, data cleaning, Excel (pivot tables, XLOOKUP/SUMIFS, charts), ETL pipeline basics
Visualisation	Matplotlib, Seaborn, Plotly
Business Intelligence	Power BI: Power Query (M), star-schema data modelling, DAX, time intelligence, Row-Level Security, Power BI Service
Dashboards	Power BI, Streamlit
Machine Learning	scikit-learn (regression, classification, clustering, cross-validation)
Testing	pytest for data validation
Tools	Jupyter, Git, GitHub, VS Code

CATEGORY	TECHNOLOGIES
AI Assistants	GitHub Copilot, ChatGPT, Claude

Topic Depth and Awareness

Depth means repeated practice until you can apply the concept confidently. **Awareness** means you are introduced clearly enough to read and compare it later.

FOCUS AREA	LEARNING FOCUS
Python, Pandas, NumPy, SQL	Depth
Analyst Excel (pivot tables, XLOOKUP/SUMIFS, basic charts)	Depth
SQL INNER/LEFT joins, three core window functions (ROW_NUMBER, RANK, LAG), Python-SQL integration	Depth
Deep SQL window-function fluency (LEAD, SUM OVER, running totals, partitioned ranking)	Depth (Week 12 interview drill)
Descriptive stats, 2-sample t-test, chi-square, CLT intuition	Depth
Seaborn, one Plotly chart, local Streamlit	Depth
Matplotlib fine control, Plotly choropleth	Awareness
Power BI: Power Query, data modelling, DAX (context transition, CALCULATE, SUMX, time intelligence), static RLS, publishing in Import mode	Depth
Regression, linear/logistic models, confusion matrix + precision/recall/F1, k-fold CV, trees/forests, K-Means, feature engineering, SHAP	Depth
Pytest and reproducible project structure	Depth
Analysis narrative and stakeholder communication	Depth
Power BI storage-mode choice (Import vs DirectQuery vs Direct Lake; beginners ship Import), dynamic RLS, calculation groups, visual calculations, RANKX/AVERAGEX, paginated reports	Awareness

FOCUS AREA	LEARNING FOCUS
ROC-AUC/PR-AUC (read the chart), stratified/time-series CV, hierarchical clustering, permutation importance	Awareness
Tableau as a resume hedge	Awareness
TikTok Analytics, social platform APIs	Awareness
ETL pipeline tooling (Airflow, Prefect)	Awareness
Time-series forecasting (Prophet, ARIMA)	Awareness (covered in depth in TR-10)
Deep learning, TensorFlow, PyTorch	Awareness (covered in depth in TR-10)
Big Data tools (Spark)	Awareness
MLOps and model deployment	Awareness (covered further in TR-10)

Project Pool

Each student picks **one** capstone with their mentor in the Week-8 Day-5 1:1, just before the Week-9 ML lifecycle begins. The pool rotates per batch:

#	PROJECT	STACK FOCUS
1	Customer churn predictor (telecom / SaaS)	Python, SQL, scikit-learn
2	Banking / fintech Power BI dashboard	SQL, Power Query, data modelling, DAX, RLS
3	E-commerce funnel analysis	SQL, window functions, Power BI / visualisation
4	NGO impact tracker	Python, Pandas, Streamlit
5	Healthcare outcomes study	Python, statistics, SHAP

#	PROJECT	STACK FOCUS
6	Trainer-scoped client brief	Whatever the brief requires

Career Paths & Trajectory

ROLE	TIMELINE	FOCUS
Data Analyst / Power BI Analyst	0–2 yrs	SQL, Power BI, cleaning, stakeholder explaining
BI Analyst / Analytics Associate	1–3 yrs	Power BI dashboards, data modelling, metric design, storytelling
Business Analyst	1–3 yrs	Requirements gathering, process mapping, data-driven decisions
Junior Data Scientist / Product Analyst	2–4 yrs	Problem framing, modelling, evaluation honesty
Senior Data Analyst / Data Scientist	4+ yrs	Judgment, reproducibility, high-trust analysis

FAQ

What does Data Science & Machine Learning cost and how do I pay?

NPR 35,500. Pay by cash at our Old Baneshwor campus, bank transfer, or Fonepay. Email info@saarathiacademy.com.np for the payment plan.

Do I need prior experience?

Prerequisites are listed above. The Saarathi Gate Assessment runs before Day 1 to map your starting point, diagnostic, not pass-or-fail. Below prerequisites? We suggest a short foundation track first.

How long is the course?

12 Weeks (3 Months). The capstone sprint runs in Week 11, with mock interviews and demo day in Week 12.

What tools will I use?

Python, SQL (PostgreSQL), pandas, Power BI (Power Query, DAX, data modelling), Streamlit, scikit-learn. Every module ships a public, deployed artefact for your portfolio, no toy projects.

What roles does this prepare me for?

Roles are listed in the Career Paths section. We do not promise jobs. We build your portfolio, run mock interviews, polish your CV and LinkedIn, and refer you into our hiring network.

How does the batch and mentorship work?

Maximum 10 students per batch, every student gets direct mentor time and project review. Hybrid by default, in-person at Old Baneshwor with online recordings and office hours. Sunday Open Classroom is free for enrolled students: pair-program, get unstuck, or just focus.

Enrollment

- **Next Batch:** see the live course page
 - **Location:** Old Baneshwor Chowk, Kathmandu, Nepal
 - **Mode:** Online + Offline
 - **Contact:** 9744442469, 9761095364
 - Limited to 10 seats per batch
-

Data Science & Machine Learning — full details, fees and next batch dates:
saarathiacademy.com.np/data-science-and-machine-learning-course-in-nepal

Version 1.0 · Updated July 2, 2026