

CAREER-SWITCH PROGRAM FOR WORKING PROFESSIONALS

Data Analytics using NumPy, Pandas & Matplotlib

with **Python, SQL, Statistics & Docker**: the common foundation for
Data Analyst, Data Engineer and **Data Scientist** careers.

Python

NumPy

Pandas

Matplotlib

Seaborn

Statistics

PostgreSQL

Docker

Git & GitHub



Prof. Ramdas Biradar
Lead Mentor · 14+ yrs



Prof. Neha Biradar
Co-Mentor · 14+ yrs

STARTS

Fri, 16 Oct 2026

Completes March 2027

SCHEDULE

Mon · Wed · Fri

9:30 – 10:30 PM IST

FORMAT

58 Live Sessions

incl. 8 hands-on labs

SEATS

Only 30

From ₹15,999

AT A GLANCE

Switch into data, **without quitting your job.**

A 5-month live online program, taught in Marathi, built for working professionals. You'll learn the core toolkit every data role uses (Python, NumPy, Pandas, Matplotlib, SQL, Statistics and Docker) and prove it with three projects on your GitHub.

58

Live sessions, including
8 hands-on labs

3

GitHub projects: 2 minor
+ 1 major

9

Modules, from Python to
Docker

30

Seats only, so every
participant gets attention

Batch details

- **Start:** Friday, 16 October 2026
- **Days:** Monday, Wednesday & Friday
- **Time:** 9:30 PM – 10:30 PM IST
- **Mode:** Online live only
- **Duration:** ~5 months, 60-minute sessions
- **Expected completion:** March 2027
- **Language:** Marathi, with Hindi & English where needed

Who this batch is for

- Working professionals (IT or non-IT) planning a move into data roles
- Testers, support engineers, developers and MIS/Excel users who want real data skills
- Finance, operations and business professionals who work with data daily
- **No prior coding required.** Python is taught from the ground up in Module 1

What this batch is, honestly

- Prepares you for **Data Analyst** roles with a job-relevant toolkit and portfolio
- Builds the **shared foundation** that Data Engineering and Data Science are built on
- Gives you real exposure to pipelines (Docker + Postgres) and statistics used in DS
- It is **not** a complete Data Engineering course: Spark, Airflow and cloud come next
- It is **not** a Machine Learning course: scikit-learn and ML come next
- No placement guarantee. We give you the skills, portfolio and interview prep

In the final module you get a clear roadmap for the next step toward Data Engineer or Data Scientist roles.

How each week works

Three live classes combining concept and live coding, a practice set after every class, and a hands-on lab after about every 6 sessions.

Your time commitment

3 hours of live classes plus **3–4 hours of practice** each week. Consistency matters more than speed.

Missed a class?

Every session is recorded and shared, so you can catch up at your own pace.

LEARNING PATH

From zero code to an end-to-end data pipeline.

Each module builds on the previous one. Hands-on labs and project reviews are included in the session counts. Months are approximate.

M0	Orientation & Setup Data careers, tools, Git & GitHub workflow	2 sessions ~Oct 2026
M1	Python for Data Just the Python you need for data, from zero	7 sessions ~Oct – Nov 2026
M2	Mathematics for Data + NumPy Business maths, algebra, linear algebra + fast arrays	7 sessions ~Nov 2026
M3	Data Wrangling with Pandas Load, clean, reshape, combine and summarise data	10 sessions ~Nov – Dec 2026
M4	Visualisation & Storytelling Matplotlib, Seaborn and one-page data stories	5 sessions ~Dec 2026 – Jan 2027
M5	Statistics & Probability Descriptive stats, probability, testing, A/B tests	7 sessions ~Jan 2027
M6	SQL with PostgreSQL Queries to window functions, design and Python	11 sessions ~Jan – Feb 2027
M7	Docker for Data Containers, Postgres in Docker, Python connections	4 sessions ~Feb 2027
M8	Capstone & Career Capstone labs, portfolio, interviews, demo day	5 sessions ~Mar 2027

Mathematics

Percentages, growth, algebra, logs, sets, vectors, matrices and calculus intuition

Statistics

Descriptive stats, probability, distributions, CLT, confidence intervals, tests, A/B testing

SQL

PostgreSQL from SELECT to window functions, CTEs, schema design and Python

Docker

Postgres + pgAdmin in containers, docker-compose, Python connections

Tools you'll use

Python 3 Jupyter VS Code NumPy
 Pandas Matplotlib Seaborn SciPy
 PostgreSQL pgAdmin SQLAlchemy Docker
 Docker Compose Git GitHub

Project milestones

- **Minor Project 1** (Pandas EDA): starts in Lab 4 · review ~Jan 2027
- **Minor Project 2** (SQL + Stats): starts in Lab 7 · review ~Feb 2027
- **Major Project** (Data Pipeline): released in Module 7 · Demo Day ~Mar 2027

DETAILED SYLLABUS · SESSION BY SESSION

What you'll learn, **class by class.**

M0 Orientation & Setup

2 sessions · Led by Ramdas Sir · approx. Oct 2026

Understand the three data careers, set up a professional toolkit, and learn the GitHub workflow used to submit every project.

S01 Welcome to Data Careers & Your Toolkit

- Data Analyst vs Data Engineer vs Data Scientist: day-to-day work, skills and hiring expectations
- How this batch runs: live classes, practice sets, projects, doubt support
- Install Python (Miniconda), VS Code and Jupyter; your first notebook

S02 Git & GitHub for Data Professionals

- Repositories, commits, push/pull, .gitignore and branches (basics)
- Writing a professional README in Markdown
- Setting up your GitHub profile; how projects will be submitted and reviewed

M1 Python for Data

7 sessions · Led by Ramdas Sir · approx. Oct – Nov 2026

Learn exactly the Python you need for data work: no prior coding required. Every session ends with practice problems.

S03 Python Basics

- Variables, data types, operators and expressions
- Input/output and f-strings
- How Python runs code; writing clean, readable code

S04 Strings & Type Conversion

- String methods for cleaning text: strip, split, replace, lower/upper
- Indexing and slicing
- Converting between types; common conversion errors

S05 Lists, Tuples & Sets

- Creating, slicing and modifying lists
- Tuples and sets: when to use which
- List comprehensions for quick transformations

S06 Dictionaries, Conditions & Loops

- Dictionaries and nested data; working with JSON-like records
- if/elif/else and logical operators
- for/while loops, enumerate, zip, break/continue

S07 Functions & Modules

- Defining functions; parameters, return values, *args/**kwargs
- lambda, map and filter
- Importing modules, pip and virtual environments

S08 Files, Errors & Dates

- Reading and writing CSV and JSON files
- Exception handling with try/except/finally
- datetime basics; mini-exercise: analyse a CSV by hand (and see why Pandas exists)

S09 Hands-on Lab 1: Python Practice

LAB

- Guided problem-solving on strings, collections, loops and functions
- File-handling exercises with real CSV/JSON files
- Doubt clearing on Module 1

M2 Mathematics for Data + NumPy

7 sessions · Led by Ramdas Sir · approx. Nov 2026

Build the practical mathematics behind data work and apply it immediately with NumPy's fast arrays.

S10 Maths I: Numbers for Business Data

- Percentages, percentage change, ratios and proportions
- Growth rates and CAGR; weighted averages
- Summation (sigma) notation; rounding and precision

S11 Maths II: Algebra, Functions & Logarithms

- Functions and graphs; linear equations, slope and intercept
- Exponents and logarithms; why log scales matter
- Sets and Boolean logic: the maths behind filters and joins

S12 NumPy I: Arrays

- ndarray: creation, dtype, shape and reshape
- Indexing, slicing and boolean masks
- Array vs list: speed and memory

S13 Maths III: Linear Algebra & Calculus Intuition

- Vectors and matrices; shape, transpose, dot product and matrix multiplication
- A dataset as a matrix: rows, features and observations
- Rate of change and minima: the intuition behind how ML models learn

S14 NumPy II: Vectorisation & Broadcasting

- Vectorised operations and ufuncs
- Broadcasting rules
- Aggregations along axes: sum, mean, std, argmax

S15 NumPy III: Random, Linear Algebra & Missing Values

- Random numbers and simple simulations
- Matrix operations with @ and np.linalg
- Handling NaN; performance tips

S16 Hands-on Lab 2: Maths & NumPy

LAB

- Business-maths problems solved in Python
- NumPy array exercises: masks, broadcasting, aggregations
- Doubt clearing on Module 2

M3 Data Wrangling with Pandas

10 sessions · Led by Ramdas Sir · approx. Nov – Dec 2026

Master the everyday toolkit of every data role: load, clean, reshape, combine and summarise real-world data.

S17 Series & DataFrames

- Reading CSV, Excel and JSON
- Inspecting data: head, info, describe, dtypes
- Understanding the index

S18 Selecting & Filtering

- loc and iloc; boolean filtering and query()
- Sorting and ranking
- Creating and modifying columns

S19 Data Cleaning I

- Missing values: detect, drop or fill
- Duplicates and type conversions
- Text cleaning with the .str accessor

S20 Data Cleaning II

- Detecting and treating outliers
- Categorical data, replace and map
- apply(); binning with cut and qcut

S21 Hands-on Lab 3: Data Cleaning

LAB

- Clean a messy real-world dataset step by step
- Pair-review of cleaning decisions
- Doubt clearing

S22 GroupBy & Aggregation

- Split-apply-combine
- agg with multiple functions; named aggregations
- transform and filter on groups

S23 Combining & Reshaping

- merge, join and concat (with SQL analogies)
- pivot_table, melt and crosstab
- Avoiding duplicate rows after merges

S24 Time Series

- Parsing datetimes; the .dt accessor
- resample, rolling windows and shift
- Month-on-month and year-on-year comparisons

S25 Pandas Best Practices

- Method chaining and readable pipelines
- Memory optimisation and reading large files in chunks
- Writing to CSV, Excel and Parquet · **Minor Project 1 released**

S26 Hands-on Lab 4: Pandas + Minor Project 1 Work Session

LAB

- End-to-end Pandas exercise
- Start Minor Project 1 with mentor guidance
- Repository setup and README checklist

M4 Visualisation & Storytelling

5 sessions · Led by Ramdas Sir · approx. Dec 2026 – Jan 2027

Turn analysis into clear, honest charts and one-page reports that decision-makers understand.

S27 Matplotlib Fundamentals

- Anatomy of a figure: Figure and Axes (object-oriented API)
- Line, bar, scatter and histogram
- Plotting directly from Pandas

S28 Customising Charts

- Titles, labels, annotations, colours and styles
- Subplots and dual axes
- Exporting high-resolution images

S29 Seaborn Quick Wins

- Distribution plots, box and violin plots
- Heatmaps and correlation matrices
- Pair plots for exploring relationships

S30 Data Storytelling

- Choosing the right chart for the question
- Decluttering and highlighting the insight
- Building a one-page analysis report; common charting mistakes

S31 Minor Project 1: Review & Showcase

- Selected participants present their EDA
- Mentor feedback on code, charts and README
- Common mistakes and how to fix them

M5 Statistics & Probability

7 sessions · Led by Neha Ma'am · approx. Jan 2027

Learn the statistics used to make decisions from data, with every concept demonstrated in Python (NumPy, Pandas, SciPy).

S32 Descriptive Statistics

- Types of data; mean, median and mode
- Variance, standard deviation, percentiles and IQR
- Skewness, outliers and box plots

S33 Probability

- Sample spaces, events and probability rules
- Conditional probability and independence
- Bayes' theorem intuition with real examples

S34 Probability Distributions

- Binomial and Poisson distributions
- Normal distribution, z-scores and the empirical rule
- Simulating distributions with NumPy

S35 Sampling & Estimation

- Populations vs samples; sampling bias
- Central Limit Theorem and standard error
- Confidence intervals and bootstrapping

S36 Hypothesis Testing

- Null and alternative hypotheses; p-values; Type I and II errors
- t-tests and chi-square tests with SciPy
- Interpreting results in plain business language

S37 Correlation & A/B Testing

- Covariance; Pearson and Spearman correlation
- Correlation vs causation; simple linear regression intuition
- Designing and analysing an A/B test end to end

S38 Hands-on Lab 5: Statistics

LAB

- Descriptive stats and distributions on real data
- Run and interpret hypothesis tests with SciPy
- Doubt clearing on Module 5

M6 SQL with PostgreSQL

11 sessions · Led by Ramdas Sir · approx. Jan – Feb 2027

Write the SQL that analysts and engineers use daily, from basic queries to window functions and schema design.

S39 Relational Databases & First Queries

- Tables, rows, keys and schemas; installing PostgreSQL and pgAdmin
- CREATE, INSERT and data types
- SELECT, WHERE, ORDER BY and LIMIT

S40 Functions & Conditional Logic

- String, date and numeric functions
- CASE WHEN
- NULL handling with COALESCE; DISTINCT

S41 Aggregation

- COUNT, SUM, AVG, MIN and MAX
- GROUP BY and HAVING
- Business KPIs in SQL

S42 Joins

- INNER, LEFT, RIGHT, FULL, SELF and CROSS joins
- Join pitfalls: fan-out and duplicate counts
- Multi-table business questions

S43 Hands-on Lab 6: SQL I

LAB

- Query practice: filtering, aggregation and joins
- Business questions on a sample database
- Doubt clearing

S44 Subqueries, CTEs & Set Operations

- Scalar, correlated and IN/EXISTS subqueries
- Common Table Expressions (WITH)
- UNION, INTERSECT and EXCEPT

S45 Window Functions I

- OVER and PARTITION BY
- ROW_NUMBER, RANK and DENSE_RANK
- Top-N per group

S46 Window Functions II

- LAG and LEAD; period-over-period change
- Running totals and moving averages
- Cohort and retention analysis

S47 Design, Performance & SQL from Python

- Keys, normalisation (1NF-3NF) and an intro to star schemas
- Indexes, EXPLAIN and views
- Querying from Python with SQLAlchemy and pandas.read_sql · **Minor Project 2 released**

S48 Hands-on Lab 7: SQL II + Minor Project 2 Work Session

LAB

- Window-function and CTE practice
- Start Minor Project 2 with mentor guidance
- Loading data into PostgreSQL

S49 Minor Project 2: Review & Showcase

- Selected participants present their SQL analysis
- Review of query quality, correctness and statistical reasoning
- SQL interview-style problem walkthrough

M7 Docker for Data

4 sessions · Led by Ramdas Sir · approx. Feb 2027

Run databases and data scripts in containers, and connect Python to a containerised PostgreSQL. This is the bridge to data engineering.

S50 Docker Fundamentals

- Why containers; images vs containers
- Docker Desktop setup (WSL2 on Windows)
- run, ps, logs, exec; ports and volumes

S51 PostgreSQL + pgAdmin in Docker

- Running Postgres in a container with persistent volumes
- Environment variables and configuration
- docker-compose.yml for a multi-service setup

S52 Connecting Python to Containerised Postgres

- SQLAlchemy / psycopg connections; secrets in .env
- Loading DataFrames with to_sql; reading back with read_sql
- A Dockerfile for a Python ETL script · **Major Project released**

S53 Hands-on Lab 8: Docker

LAB

- Everyone gets Postgres + pgAdmin running in Docker
- Troubleshooting on Windows/WSL2, macOS and Linux
- Connect Python to the container and load a DataFrame

M8 Capstone & Career

5 sessions · Led by Ramdas Sir · approx. Mar 2027

Build the major project with mentor support, present it, and prepare to apply for data roles.

S54 Capstone Lab I: ETL & Data Quality

- Extract, validate and clean the chosen dataset
- Data-quality checks and logging
- Loading into PostgreSQL

S55 Capstone Lab II: Analysis Layer

- SQL views and analysis tables
- Statistical analysis of the key question
- Report charts

S56 Portfolio & Personal Branding

- Writing a README that recruiters read
- A resume for career switchers; LinkedIn and GitHub profile
- Telling your project story in an interview

S57 Interview Preparation & Roadmap

- Python, Pandas, SQL and statistics interview questions
- Group mock-interview practice
- Next steps: roadmap to Data Engineer (Spark, Airflow, cloud) and Data Scientist (scikit-learn, ML)

S58 Capstone Demo Day

- Participants present their end-to-end projects
- Final feedback
- Certificates and the way forward

PROJECTS YOU PUSH TO GITHUB

Three projects that **prove your skills.**

Every project is a public GitHub repository with clean code, a clear README and reproducible setup. These repositories are your portfolio for interviews.

MINOR PROJECT 1 · INDIVIDUAL

Retail Sales Exploratory Data Analysis

Starts in Lab 4
Review ~Jan 2027

PROBLEM

A retailer wants to understand what drives its revenue. You get a messy real-world transactions dataset (e.g. the public UCI *Online Retail II* dataset) and turn it into clear business insights.

WHAT YOU'LL DO

- Clean the data: missing customers, cancelled orders, negative quantities, duplicates and wrong types
- Engineer features: revenue, month, weekday, basket size, customer tenure
- Analyse top products, countries, seasonality, repeat vs one-time customers
- Produce 8–10 publication-quality Matplotlib/Seaborn charts
- Write 5 actionable business recommendations

SKILLS PROVEN

Python · NumPy · Pandas cleaning, GroupBy, time series · Matplotlib · Storytelling

REPOSITORY STRUCTURE

```
retail-sales-eda/
├── data/raw/           (download script, not data)
├── data/processed/
├── notebooks/
│   └── 01_cleaning_and_eda.ipynb
├── src/cleaning.py
├── reports/figures/
├── requirements.txt
└── README.md          (questions, insights, charts)
```

How every project is evaluated

CRITERION	WEIGHT
Correctness of analysis & results	30%
Code quality & structure	20%
Insights & storytelling	20%
README & documentation	15%
Reproducibility (runs on a fresh machine)	15%

MINOR PROJECT 2 · INDIVIDUAL

E-commerce SQL Analytics & Hypothesis Testing

Starts in Lab 7
Review ~Feb 2027

PROBLEM

Analyse a multi-table e-commerce database (e.g. the public *Olist Brazilian E-Commerce* dataset: orders, customers, items, payments, reviews, sellers) in PostgreSQL, then test a business hypothesis statistically.

WHAT YOU'LL DO

- Design the schema and load all tables into PostgreSQL
- Answer 15 business questions in SQL: revenue trends, top categories, seller performance, delivery delays by region
- Use CTEs and window functions for month-over-month growth, cohort retention and RFM customer segments
- Test a hypothesis in Python with SciPy, e.g. *"Do late deliveries lead to lower review scores?"*

SKILLS PROVEN

SQL joins, CTEs, window functions · Schema design · SQLAlchemy · Hypothesis testing

REPOSITORY STRUCTURE

```
ecommerce-sql-analytics/
├── sql/
│   ├── 01_schema.sql
│   ├── 02_load.sql
│   └── 03_business_questions.sql
├── notebooks/
│   └── hypothesis_test.ipynb
├── results/                (query outputs, charts)
├── requirements.txt
└── README.md              (ER diagram, questions, findings)
```

GitHub submission rules

- One public repository per project, with meaningful commit history
- No raw data or passwords committed: use download scripts and .env
- README with problem, approach, key charts, findings and how to run
- Submit the repository link before the review session
- Mentors review and give written feedback

MAJOR PROJECT · CAPSTONE · INDIVIDUAL

End-to-End Dockerised Data Pipeline & Analytics Report

Released in Module 7
Demo Day ~March 2027

ARCHITECTURE



PROBLEM

Build a small but real data system that anyone can run with one command. Choose one mentor-approved public dataset (for example Indian air-quality data, NYC taxi trips, or Indian stock-market prices) and answer a meaningful question end to end.

WHAT YOU'LL BUILD

- **docker-compose** stack running PostgreSQL + pgAdmin with persistent storage
- A Python **ETL script** that downloads, validates, cleans (Pandas) and loads data via SQLAlchemy
- **Data-quality checks** (row counts, nulls, ranges) with logging
- **SQL views** for the analysis layer
- A **statistical analysis** answering your key question (trend, comparison or hypothesis test)
- An **analytics report** with charts and recommendations
- Config via **.env**, no passwords in code

STRETCH GOALS (OPTIONAL)

- Incremental loads, pytest unit tests, or scheduling with cron

WHY IT MATTERS

This one project shows analyst skills (SQL, charts, insights), engineering skills (ETL, Docker, databases) and data-science thinking (statistics). It's exactly the "hand-in-hand" foundation of all three careers.

REPOSITORY STRUCTURE

```

data-pipeline-capstone/
├── docker-compose.yml
├── .env.example
├── etl/
│   ├── Dockerfile
│   ├── extract.py
│   ├── transform.py
│   ├── load.py
│   └── quality_checks.py
├── sql/
│   ├── schema.sql
│   └── views.sql
├── analysis/
│   └── analysis.ipynb
├── reports/
│   ├── figures/
│   └── report.md
├── run_pipeline.py
├── requirements.txt
└── README.md
    (architecture, setup,
     how to run, findings)
  
```

RUN IT IN TWO COMMANDS

```

docker compose up -d
python run_pipeline.py
  
```

YOUR MENTORS

Learn from people who **teach for a living.**

Ramdas Sir leads the program and teaches every module except Statistics, which is taught by Neha Ma'am.



LEAD MENTOR · @LOGIC_OVERFLOW

Prof. Ramdas Biradar

14+ years teaching experience

QUALIFICATION	MCS, BCS
EXPERTISE	Python, SQL, C, Java, DSA, Web Technologies, Operating Systems, Computer Networking
STUDENTS	5,000+ students trained; @logic_overflow on Instagram
IN THIS BATCH	Teaches Python, Mathematics, NumPy, Pandas, Visualisation, SQL and Docker; leads projects, labs and career preparation



CO-MENTOR · STATISTICS

Prof. Neha Biradar

14+ years teaching experience

QUALIFICATION	MCS, BCS
EXPERTISE	Mathematics, Statistics, Python, C, C++, Java
IN THIS BATCH	Teaches the Statistics & Probability module

WHAT YOU GET

Everything included, **one clear fee.**

58 live sessions

60-minute live classes with live coding, including 8 hands-on labs.

Session recordings

Every class is recorded and shared for revision.

Practice sets

Assignments after every session to build real fluency.

3 reviewed projects

Two minor and one major GitHub project with mentor feedback.

Doubt support

Batch community group for questions between classes.

Certificate

Certificate of completion from Logic Overflow.

FEES

EARLY-BIRD

₹15,999

For the **first 15 registrations** or until **Sunday, 4 October 2026, 11:59 PM IST**, whichever comes first.

REGULAR

₹17,999

After the early-bird offer closes, until all **30 seats** are filled.

Before you join

- Laptop with Windows 10/11 (64-bit, WSL2 supported), macOS or Linux
- 8 GB RAM minimum (16 GB recommended) and 20 GB free disk space, needed for Docker
- Stable internet connection
- 3–4 hours per week for practice outside class

Policies

- Admissions are first-come, first-served; registration closes when 30 seats are filled
- Fees are **non-refundable and non-transferable**
- If Logic Overflow cancels or postpones the batch, a full refund will be given
- The session schedule is shared in the batch group; any changes are communicated in advance

Reserve your seat before the early-bird offer closes.

logic-overflow.com/batch/da-2026 →

Tap the link to register and pay online.

WhatsApp: +91 97303 81255

Email: ramdas@logic-overflow.com

Website: logic-overflow.com

Instagram: @logic_overflow