

PREPNPLACED

1:1 MENTORSHIP PLAN · 2026

1:1 DATA ENGINEERING MENTORSHIP WITH DURGESH YADAV

SQL to Python to Spark, Snowflake, Kafka and AWS,
taught 1:1 in order, with an assignment after every class.

30

DE classes and
mentorship sessions

3

portfolio projects
on GitHub

10

mock interviews
+ 10 live support

28

SQL + Python classes
if you need them

COMPLETE PLAN

₹2,00,000

SQL + Python + Data Engineering + 10 mocks

DE PLAN

₹1,30,000

Already know SQL and Python? Start at DE

THE STACK

SQL · Python · PySpark · Databricks · Delta Lake · Snowflake · dbt
Airflow · Kafka · Structured Streaming · AWS · Git · Docker

OVERVIEW

TWO PLANS, ONE MENTOR

Pick by one question: do you already know SQL and Python?

This is a 1:1 mentorship with Durgesh Yadav that takes you from where you are today to building, running and defending production-style pipelines on the modern data stack. It comes in two plans. The **Complete plan** teaches SQL and Python in live 1:1 classes before Data Engineering starts. The **DE plan** is for people who already know SQL and Python: it skips those two tracks and starts straight at Data Engineering, so it costs less and finishes sooner. The projects, mentorship, mocks and interview support are the same in both.

	Complete plan	DE plan
Who it is for	New to SQL or Python, or not yet at interview level in them	Already comfortable with SQL and Python
SQL, 16 live 1:1 classes	Included	Not included
Python, 12 live 1:1 classes	Included	Not included
Data Engineering, 15 classes + 15 mentorship sessions	Included	Included
3 portfolio projects on GitHub	Included	Included
10 mock interviews	Included	Included
Live support for 10 real interviews	Included	Included
Paid live hours	68	40
Assignments reviewed by Durgesh	58	30
Free doubt lectures	8	4
Duration at 3 sessions a week	About 26 weeks	About 15 weeks
Fee	₹2,00,000	₹1,30,000

Choose the DE plan if

You can already write joins, GROUP BY, window functions and CTEs in SQL without looking them up, and in Python you are comfortable with functions, dictionaries, file handling and basic pandas.

The skill check in the DE plan

In the DE plan, the SQL mock (MI01) and the Python mock (MI02) run in weeks 1 and 2. They confirm you are ready to skip those tracks. If a gap shows up, you can add only the track you need, SQL for ₹40,000 or Python for ₹30,000, at the same hourly rates.

FEES

FEES AND DISCOUNT

Hourly rates · both plan prices · up to 10% off

Fee breakdown

Component	Sessions	Rate	Complete plan	DE plan
SQL	16 × 1 hr	₹2,500 / hr	₹40,000	Not included
Python	12 × 1 hr	₹2,500 / hr	₹30,000	Not included
Data Engineering 15 learning + 15 mentorship	30 × 1 hr	₹3,500 / hr	₹1,05,000	₹1,05,000
Mock interviews	10 × 1 hr	₹2,500 / hr	₹25,000	₹25,000
Total	68 / 40 hrs		₹2,00,000	₹1,30,000

Discount: up to 10%, never more

If you are a serious learner and the fee is a stretch, Durgesh can take up to 10% off either plan. It is decided on the intro call, before you enrol, and 10% is the ceiling.

Plan	Full fee	Maximum discount	Lowest possible fee
Complete plan	₹2,00,000	₹20,000 (10%)	₹1,80,000
DE plan	₹1,30,000	₹13,000 (10%)	₹1,17,000

Included at no extra cost

What you get	Details
Assignments	One after every class and mentorship session, reviewed in writing by Durgesh: 58 in the Complete plan, 30 in the DE plan.
Doubt support	Live doubt time in every class, 1:1 doubt calls when you are stuck, and chat support for quick questions.
Additional doubt lectures	1-hour doubt lectures at the end of each track block: 8 in the Complete plan, 4 in the DE plan.
Live interview support	For up to 10 of your real interviews: a prep call before and a debrief after. Both plans.
All content and material	Class notes, datasets, assignment sheets, project templates, cheat sheets, the PrepnPlaced interview question bank and class recordings.

YOUR MENTOR

DURGESH YADAV

Data Engineer II at 7-Eleven · Founder, PrepNPlaced · Bangalore

Durgesh builds data pipelines for a living. At 7-Eleven he works on supply-chain and vendor-payment reconciliation pipelines in PySpark on Databricks: a medallion setup on Delta tables, a config-driven ETL framework, data-quality controls, and Airflow-orchestrated Python Dash KPI dashboards for the Product and Demand Chain teams. Before that, at Target, he worked on e-commerce data at the 1B+ row scale with advanced SQL, PostgreSQL, Hive and DOMO. Most of what this plan teaches is what he uses at work every week. The mock interviews come from rounds he has faced in his own 20+ interviews, and mentees who went through the same loops after him keep reporting the same patterns.

5+

years teaching data engineering and analytics

20+

DA and DE interviews he has sat himself

3+

data engineering cohorts delivered

4

certifications: Azure, Databricks, AWS, Google

Work experience

Period	Role	Company
Aug 2024 to now	Data Engineer II (Data Analytics)	7-Eleven
Oct 2022 to Jul 2024	Product Analyst	Target
Nov 2021 to Oct 2022	SE Data Analyst	Coforge

Teaching and mentoring

Where	Role
PrepNPlaced	Founder. Runs the Data Engineering and Data Analytics cohorts and webinars.
Scaler	Data Analytics Instructor, since Mar 2026
BossCoder Academy	Data Engineering Mentor and Instructor, since Jun 2025
GeeksforGeeks	Data Analytics Instructor and Mentor, since May 2024
Analytics Vidhya	Data Engineering Instructor
Topmate	1:1 mentorship, since Sep 2022

Certifications and education

Microsoft Azure DP-203, Databricks / Spark, AWS Data Analytics, Google Data Analyst. B.Tech in Computer Science, Dr. APJ Abdul Kalam Technical University (2017 to 2021). GitHub: github.com/analyticsdurgesh

HOW IT WORKS

HOW THE MENTORSHIP RUNS

Class format · Assignments · Doubt support · Mocks · Interview support

Every class, the same shape

Minutes	What happens
0 to 5	Quick recap and review of the last assignment.
5 to 50	The topic, taught live with hands-on coding on real datasets.
50 to 60	Open doubts, then the next assignment is handed over and explained.

Assignments

There is one assignment after every class and mentorship session: 58 in the Complete plan, 30 in the DE plan. Each one is due before the next session and goes into your own GitHub repo, one folder per track, so your practice turns into a visible body of work. Durgesh sends written feedback within 48 hours and you get one resubmission. Completion is tracked in a shared progress sheet.

Doubt support, four layers

Layer	How it works
Live, in class	The last 10 minutes of every class are kept for doubts from that class or the previous assignment.
Doubt calls	Stuck on a concept or an assignment? Ask for a 1:1 call. It gets scheduled within 24 to 48 hours.
Chat support	Send code, screenshots or errors on WhatsApp. Replies within 24 hours on working days.
Doubt lectures	Extra 1-hour lectures at fixed points in the plan, 8 in the Complete plan and 4 in the DE plan (listed on the doubt lectures page). Free.

Mock interviews and live interview support

Each mock is 45 minutes of interview and 15 minutes of feedback. You get a written scorecard on problem solving, code quality, communication and depth, plus the three things to fix before the next mock. Once real interviews start, live support covers up to 10 of them: a prep call before each one on the JD, the company's round pattern and likely questions, and a debrief after on what was asked, where you slipped and what to fix.

ROADMAP

THE ORDER YOU WILL FOLLOW

Complete plan: 76 sessions (68 paid + 8 free) · DE plan: 44 sessions (40 paid + 4 free)

Topics are sequenced so that each block uses what the previous one taught. SQL comes first because everything after it depends on it, projects start only once the matching tools have been covered, and most of the mocks sit at the end, when you are closest to real interviews. The DE plan follows the same order but skips Phases 1 and 2, so Data Engineering starts in week 1.

Phase	Focus	Weeks: Complete	Weeks: DE plan	Sessions	Outcome
Phase 0	Kickoff	1	1	M01	Skill audit, target roles, personal roadmap, tools setup
Phase 1	SQL	1-7	Skipped	S01-S16, DL1, DL2, MI01	Querying to modeling to tuning, then the first mock
Phase 2	Python	8-12	Skipped	P01-P12, DL3, DL4, MI02	Core Python, pandas, APIs, a tested mini ETL
Phase 3	Spark and lakehouse	13-16	1-5	D01-D06, M02-M04, DL5, MI03	Spark, Databricks, Delta, medallion. Project 1 shipped
Phase 4	Modern stack	17-18	6-7	D07-D09, M05-M06, DL6	Snowflake, dbt, Airflow. Project 2 shipped
Phase 5	Streaming and cloud	19-20	8-9	D10-D12, M07-M08, DL7	Kafka, Structured Streaming, AWS. Project 3 shipped
Phase 6	Production and design	21-22	10-11	D13-D15, M09, DL8	Data quality, CI/CD, system design, portfolio polish
Phase 7	Job ready	23-26	12-15	M10-M15, MI04-MI10	Resume, profiles, outreach, 7 mocks, live interview support

Pace

Complete plan: about 26 weeks (6 months) at 3 sessions a week, roughly 4 months at 4 to 5 a week. DE plan: about 15 weeks at 3 sessions a week, roughly 10 weeks at 4 to 5 a week, with the SQL and Python mocks in weeks 1 and 2 as the skill check. In both plans, live interview support starts whenever your first real interview is scheduled, even if that happens before Phase 7.

Session codes used in this plan

S = SQL class · P = Python class · D = Data Engineering class · M = Data Engineering mentorship session · MI = mock interview · DL = additional doubt lecture

TRACK 01 · COMPLETE PLAN ONLY · WEEKS 1-7

SQL

16 classes · 1 hr each · ₹2,500 / hr · ₹40,000

From the first SELECT to the modeling, tuning and interview patterns that data engineering rounds test. Doubt lectures DL1 (after S08) and DL2 (after S16), then mock MI01. Not part of the DE plan.

S01	SQL foundations: SELECT, WHERE, ORDER BY
COVERS	How a query actually runs (FROM › WHERE › GROUP BY › HAVING › SELECT › ORDER BY), SELECT and aliases, WHERE with AND / OR / IN / BETWEEN / LIKE, how NULL behaves (IS NULL, COALESCE), ORDER BY and LIMIT.
ASSIGNMENT	Set up PostgreSQL (or MySQL) and load the course e-commerce dataset: customers, orders, order_items, products. Write 12 queries covering filters, pattern matches, NULL checks and the top 10 orders by amount.
S02	Aggregations and GROUP BY
COVERS	COUNT, SUM, AVG, MIN, MAX. COUNT(*) vs COUNT(col) vs COUNT(DISTINCT). Grouping on several columns, HAVING vs WHERE, conditional aggregation with CASE.
ASSIGNMENT	Answer 10 business questions: revenue by month, orders per city, customers with more than 3 orders, average basket size by category, and the share of cancelled orders.
S03	Joins, and the duplicate-row trap
COVERS	INNER, LEFT, RIGHT, FULL, SELF and CROSS joins. Choosing join keys, fan-out duplicates that inflate totals, anti-joins (customers with no orders), joining 3+ tables.
ASSIGNMENT	Solve 12 join questions, including one anti-join and one self-join (employee and manager). Then write one query where a bad join inflates revenue and explain the bug in 3 lines.
S04	Subqueries
COVERS	Scalar subqueries, IN and EXISTS, correlated subqueries, derived tables, and when EXISTS, IN or a JOIN is the better choice.
ASSIGNMENT	Solve 8 questions twice: once with a subquery and once with a join. For each pair, write which version you would ship and why.
S05	CTEs and recursive CTEs
COVERS	The WITH clause for readable multi-step logic, chaining CTEs, recursive CTEs for hierarchies and for generating date series.
ASSIGNMENT	Rewrite 3 long queries from S04 using CTEs. Build an org chart with levels using a recursive CTE, and generate a calendar table for 2026.

S06	Window functions I: ranking
COVERS	OVER, PARTITION BY and ORDER BY. ROW_NUMBER vs RANK vs DENSE_RANK, NTILE, and top-N per group.
ASSIGNMENT	Find the top 3 products per category by revenue and the second-highest salary per department. Deduplicate a table so only the latest row per key survives.
S07	Window functions II: analytics
COVERS	LAG and LEAD, running totals, moving averages, frame clauses (ROWS vs RANGE), FIRST_VALUE and LAST_VALUE.
ASSIGNMENT	Month-over-month revenue growth %, a 7-day moving average of orders, days between consecutive orders per customer, and cumulative revenue share.
S08	CASE, strings and dates
COVERS	CASE for buckets and flags. String functions (SUBSTRING, CONCAT, TRIM, REPLACE, basic regex). Date functions (DATE_TRUNC, EXTRACT, intervals, date differences), type casting, time zones.
ASSIGNMENT	Clean a messy customers table (phone formats, name casing, invalid emails) and add a segment column (new / active / dormant) using date logic.
S09	Deduplication, pivots and set operations
COVERS	Finding and removing duplicates, UNION vs UNION ALL, INTERSECT and EXCEPT, pivoting with conditional aggregation, unpivoting.
ASSIGNMENT	Deduplicate a raw events table, build a month x category revenue pivot, and reconcile two tables with EXCEPT, reporting every mismatch.
S10	OLTP vs OLAP, and normalization
COVERS	Transactional vs analytical systems, primary and foreign keys, constraints, 1NF / 2NF / 3NF with examples, and when to denormalize on purpose.
ASSIGNMENT	Normalize a flat 20-column sales sheet to 3NF, draw the ER diagram, and write the CREATE TABLE statements with keys and constraints.
S11	Dimensional modeling
COVERS	Facts vs dimensions, declaring the grain, star vs snowflake schema, surrogate keys, fact table types (transaction, periodic snapshot, accumulating snapshot), conformed dimensions.
ASSIGNMENT	Design a star schema for a food-delivery app: state the grain, list the facts and dimensions, then write 5 analytical queries against it.
S12	Slowly changing dimensions and MERGE
COVERS	SCD Type 1 vs Type 2 (and where Type 3 shows up), effective dates and current flags, MERGE and upsert patterns, idempotent loads.
ASSIGNMENT	Implement SCD Type 2 on a customer dimension using 3 daily change files and MERGE (or INSERT ... ON CONFLICT). Prove the history is correct with queries.

S13	Indexes and query tuning
COVERS	How B-tree indexes work, column order in composite indexes, reading EXPLAIN and EXPLAIN ANALYZE, sargable filters, the usual causes of slow queries.
ASSIGNMENT	Take 5 slow queries on a 1M-row table, capture EXPLAIN before and after indexing or rewriting, and write a one-page tuning note.
S14	Partitioning and clustering in warehouses
COVERS	Table partitioning and partition pruning, clustering keys, Snowflake micro-partitions, the cost of full scans, materialized views.
ASSIGNMENT	Partition the orders table by month, compare how filtered queries scan it, and write down when you would and would not cluster a table.
S15	Interview query patterns
COVERS	Gaps and islands, consecutive-day streaks, top-N per group, retention and cohort queries, median and percentiles, sessionization.
ASSIGNMENT	Solve 15 questions taken from product-company interview loops, spread across the six patterns, with a short explanation for each.
S16	SQL inside pipelines
COVERS	Transactions and ACID, isolation levels in plain terms, views vs materialized views, incremental loads using a watermark column, data checks written in SQL (row counts, null %, duplicates, reconciliation).
ASSIGNMENT	Write an incremental load from a staging table to a target using a last-updated watermark, plus 6 SQL data checks that fail loudly when something is off.

TRACK 02 · COMPLETE PLAN ONLY · WEEKS 8-12

PYTHON

12 classes · 1 hr each · ₹2,500 / hr · ₹30,000

Python the way data engineers use it at work: files, APIs, pandas, tests and clean repos. Doubt lectures DL3 (after P06) and DL4 (after P12), then mock MI02. Not part of the DE plan.

P01	Python basics for data work
COVERS	Setup (Python 3.11+, VS Code, Jupyter), variables and types, strings and f-strings, lists and tuples, slicing, if / for / while.
ASSIGNMENT	15 short exercises on string cleanup, list slicing and looping over order records. Push them to a new GitHub repo.
P02	Dictionaries, sets and comprehensions
COVERS	Dictionaries for lookups and counting, sets for dedup and membership checks, list / dict / set comprehensions, sorting with key functions.
ASSIGNMENT	Parse 1,000 order records (a list of dicts) and compute revenue per city, unique customers and the top 5 products without using pandas.
P03	Functions
COVERS	Parameters and defaults, *args and **kwargs, return values, scope, lambda, map / filter / sorted, type hints and docstrings.
ASSIGNMENT	Build a utility module of 8 reusable cleaning functions (trim, normalize phone, parse date, safe cast and so on) with type hints and docstrings.
P04	Files, errors and logging
COVERS	Reading and writing CSV and JSON, context managers, try / except / finally, custom exceptions, using logging instead of print.
ASSIGNMENT	Write a script that reads a folder of CSVs, logs and skips bad rows, and writes one clean JSON file plus a separate reject file.
P05	OOP for pipelines
COVERS	Classes and objects, __init__ and methods, inheritance vs composition, dataclasses, and when a class is better than a function.
ASSIGNMENT	Model an Extractor › Transformer › Loader pipeline as classes. Swap the CSV extractor for a JSON one without touching the loader.

P06	Project structure and environments
COVERS	Modules and packages, virtual environments, pip and requirements.txt, config files and environment variables, .gitignore, a clean repo layout.
ASSIGNMENT	Restructure your P04 and P05 code into a proper package with a config file, requirements file, README and a single run command.
P07	Iterators, generators and decorators
COVERS	Iterators, generators and yield, processing files bigger than memory, chunking, decorators for timing and retries.
ASSIGNMENT	Stream a large log file line by line, count errors per hour, and add a retry decorator to a function that fails randomly.
P08	pandas I
COVERS	DataFrame and Series, reading CSV / Parquet / JSON, selecting and filtering, groupby and aggregation, merge and concat, sorting.
ASSIGNMENT	Answer 10 retail sales questions in pandas and check that the results match your SQL answers from S02 and S03.
P09	pandas II
COVERS	Dates and time series, strategies for missing data, apply vs vectorised operations, pivot_table and melt, dtypes and memory, where pandas stops scaling.
ASSIGNMENT	Clean a messy 500k-row dataset end to end (types, nulls, duplicates, outliers) and write a before-and-after data profile.
P10	APIs and databases
COVERS	REST basics, requests, auth headers, pagination, rate limits, flattening nested JSON, loading into Postgres with SQLAlchemy or psycopg.
ASSIGNMENT	Pull paginated data from a public API, flatten it, and load it into Postgres with an upsert so a rerun never creates duplicates.
P11	Mini ETL project with tests
COVERS	The extract › transform › load flow, writing Parquet, config-driven runs, pytest basics, unit testing transform functions.
ASSIGNMENT	Ship a mini ETL repo: API › clean › Parquet › Postgres, with 5 unit tests, logging, and a README that explains how to run it.
P12	Python coding round for data engineers
COVERS	Patterns DE interviews actually ask: hashmaps, two pointers, sliding window, sorting, string and log parsing, time complexity basics.
ASSIGNMENT	Solve 15 easy-to-medium Python interview questions and note the time complexity of each solution.

TRACK 03 · BOTH PLANS · COMPLETE: WEEKS 13-22 · DE PLAN: WEEKS 1-11

DATA ENGINEERING: LEARNING

15 classes · 1 hr each · ₹3,500 / hr · ₹52,500

The modern stack in the order a real platform gets built: Spark and the lakehouse, then the warehouse and transformation layer, then streaming and cloud, then quality and design. Doubt lectures DL5 to DL8 sit at the end of each block.

D01	How a modern data platform fits together
COVERS	Batch vs streaming, ETL vs ELT, data lake vs warehouse vs lakehouse, file formats (CSV, JSON, Parquet, Avro), table formats (Delta, Iceberg), one reference architecture end to end.
ASSIGNMENT	Draw the data architecture of an app you use daily (sources › ingestion › storage › transform › serving) and justify each choice in a short paragraph.
D02	Spark architecture
COVERS	Why distributed processing, driver / executors / cluster manager, jobs, stages and tasks, lazy evaluation and the DAG, narrow vs wide transformations, the shuffle.
ASSIGNMENT	Run a job on Databricks Free Edition, open the Spark UI, and explain its stages, tasks and the shuffle it triggered.
D03	PySpark DataFrames
COVERS	Reading data with an explicit schema, select / filter / withColumn, joins, groupBy, window functions, UDFs vs built-in functions, writing Parquet.
ASSIGNMENT	Port 8 of your SQL track answers to PySpark and confirm the outputs match.
D04	Spark optimization
COVERS	Partitions, repartition vs coalesce, broadcast joins, data skew and salting, caching, Adaptive Query Execution, the small-files problem.
ASSIGNMENT	Take a slow join job, diagnose it in the Spark UI, fix it with two techniques, and record the runtime before and after.
D05	Databricks and Delta Lake
COVERS	Workspace, compute and jobs, the Delta transaction log, ACID writes, MERGE, time travel, OPTIMIZE / Z-ORDER / VACUUM, Unity Catalog basics.
ASSIGNMENT	Create a Delta table, apply MERGE upserts for 3 batches, query an older version with time travel, then run OPTIMIZE on it.
D06	Medallion architecture and incremental loads
COVERS	Bronze / silver / gold layers, full vs incremental loads, CDC basics, watermarks, Auto Loader, idempotency, schema evolution.
ASSIGNMENT	Design the medallion layers for Project 1: every table, its keys and its load type. Bring it to the M02 session.

D07	Snowflake
COVERS	Storage / compute / services layers, virtual warehouses, stages and COPY INTO, Snowpipe, streams and tasks, time travel, keeping credits under control.
ASSIGNMENT	Load files into Snowflake with COPY INTO, set up a stream and task for incremental processing, and note the credits you used.
D08	dbt
COVERS	Project structure, sources and refs, staging › marts, materializations, incremental models, tests, snapshots, generated docs.
ASSIGNMENT	Build staging and mart models on your D07 data with at least 6 tests and one incremental model.
D09	Apache Airflow
COVERS	DAGs, tasks and operators, scheduling and catchup, retries and SLAs, sensors, XComs, backfills, designing tasks that are safe to rerun.
ASSIGNMENT	Write a daily DAG that runs extract › dbt run › dbt test, with retries and a failure alert.
D10	Apache Kafka
COVERS	Event streaming basics, topics and partitions, producers and consumers, consumer groups and offsets, delivery guarantees, message keys and ordering.
ASSIGNMENT	Run Kafka in Docker, write a Python producer for order events and a consumer group of two, and show how partitions get assigned.
D11	Spark Structured Streaming
COVERS	The micro-batch model, reading from Kafka, triggers, watermarks and late data, stateful aggregations, checkpoints, writing to Delta.
ASSIGNMENT	Stream your D10 order events into a Delta table with 5-minute windowed revenue and a 10-minute watermark.
D12	AWS for data engineering
COVERS	S3 layout and storage classes, IAM roles and policies, the Glue catalog and Glue jobs, Athena, Redshift basics, Lambda triggers, watching the bill.
ASSIGNMENT	Land raw files in S3, catalog them with a Glue crawler, query them in Athena, and trigger a Lambda whenever a new file arrives.
D13	Data quality and observability
COVERS	How pipelines fail silently, tests at each layer, data contracts, freshness and volume checks, alerting, lineage, handling bad records.
ASSIGNMENT	Add row count, null, uniqueness and freshness checks to Project 1 that fail the job, and write a short runbook for each alert.

D14	Git, CI/CD and Docker for data
COVERS	Branching and pull requests, code review, CI that runs tests on every PR, dev and prod environments, Docker images for jobs, handling secrets.
ASSIGNMENT	Add a GitHub Actions workflow that lints and tests one project on every PR, and containerize your PII ETL.

D15	Data pipeline system design
COVERS	A framework for design rounds: requirements › volume › architecture › storage › processing › quality › cost › failure modes. Batch vs streaming trade-offs, three worked designs.
ASSIGNMENT	Write full designs for two prompts, a ride-hailing trips pipeline and a clickstream analytics platform, using the framework.

TRACK 04 · BOTH PLANS · PHASES 0 AND 3-7

DATA ENGINEERING: MENTORSHIP

15 sessions · 1 hr each · ₹3,500 / hr · ₹52,500

These run alongside Track 03. Each project starts only after its tools have been taught, every session reviews your work, and each one ends with a deliverable due before the next.

The three portfolio projects

Project	What you build	Stack	Sessions
1. Batch lakehouse	Raw files to bronze, silver and gold Delta tables with incremental loads and quality checks	S3, PySpark, Databricks, Delta Lake, Workflows	M02-M04
2. Analytics engineering	Warehouse models and tested marts, orchestrated end to end	Snowflake, dbt, Airflow	M05-M06
3. Real-time streaming	Live order events processed into windowed metrics that survive a restart	Kafka, Structured Streaming, Delta	M07-M08

M01

Kickoff: skill audit and roadmap

AGENDA

Where you are vs the role you want, target companies and salary band, weekly hours you can commit, your personal 6-month plan, tools setup.

DELIVERABLE

Fill the skill audit sheet and share your resume, LinkedIn URL and target company list.

M02

Project 1 kickoff: batch lakehouse

AGENDA

Problem statement and dataset, architecture (S3 › PySpark on Databricks › Delta bronze / silver / gold), repo structure, what 'done' means.

DELIVERABLE

Repo created and bronze ingestion running on sample data.

M03

Project 1 review: bronze and silver

AGENDA

Code review of ingestion and cleaning, schema handling, dedup logic and incremental loads.

DELIVERABLE

Silver layer complete, with incremental loads tested on 3 batches.

M04

Project 1 review: gold and scheduling

AGENDA

Gold tables for business metrics, a Databricks Workflows job, quality checks, README review.

DELIVERABLE

Project 1 finished and pushed with an architecture diagram.

M05

Project 2 kickoff: Snowflake + dbt

AGENDA

Analytics engineering use case, raw › staging › marts design, naming conventions, test plan.

DELIVERABLE

Raw data loaded in Snowflake and staging models built.

M06	Project 2 review: marts, tests and Airflow
AGENDA	Review of dbt marts, tests and snapshots, and the Airflow DAG that runs the whole flow.
DELIVERABLE	Project 2 finished with dbt docs and a working DAG.
M07	Project 3 kickoff: real-time streaming
AGENDA	Event design, Kafka topics and keys, producer code, plan for the streaming job.
DELIVERABLE	Producer publishing events and a consumer reading them.
M08	Project 3 review: streaming into Delta
AGENDA	Structured Streaming job review, watermarks, checkpoints, a failure-and-restart test, the live metrics table.
DELIVERABLE	Project 3 finished with the restart test documented.
M09	Portfolio polish
AGENDA	Code review across all 3 projects, README structure, architecture diagrams, putting real numbers on results, pinned repos.
DELIVERABLE	All 3 repos ready to be opened in an interview.
M10	Resume rewrite
AGENDA	ATS-friendly structure, project and work bullets with scale and impact, matching keywords to your target JDs.
DELIVERABLE	Final resume plus 2 versions tailored to specific JDs.
M11	LinkedIn and Naukri profiles
AGENDA	Headline, About section, featured projects, skills, recruiter visibility settings, a simple posting plan.
DELIVERABLE	Both profiles updated and your first project post published.
M12	Behavioral story bank
AGENDA	STAR stories from your projects and work: ownership, a failure, a conflict, a tight deadline, measurable impact. The 'walk me through your project' answer.
DELIVERABLE	8 written STAR stories and a recorded 2-minute project pitch.
M13	Company targeting and round patterns
AGENDA	How rounds differ at service companies, product companies and GCCs, which companies fit your profile, an application tracker, weekly targets.
DELIVERABLE	A tracker with 40 target companies and roles.

M14	Referrals and outreach
AGENDA	Referral request messages, recruiter outreach, follow-ups, recruiter screens and CTC questions.
DELIVERABLE	20 referral and outreach messages sent and logged.

M15	Offers and your first 90 days
AGENDA	Comparing offers, negotiation, notice period and buyouts, background verification, a 30-60-90 day plan for the new role.
DELIVERABLE	A negotiation script for your offers and a written 90-day plan.

TRACK 05 · BOTH PLANS · SPREAD ACROSS THE PLAN

MOCK INTERVIEWS

10 mocks · 1 hr each · ₹2,500 / hr · ₹25,000

The first three mocks close out the SQL, Python and Spark blocks so gaps get caught early. The other seven run in Phase 7, right before and during your job search. Format: 45 minutes of interview, 15 minutes of feedback, written scorecard. In the DE plan, MI01 and MI02 run in weeks 1 and 2 instead, as the skill check for skipping SQL and Python.

Mock	Round	When	What it covers
MI01	SQL round	After the SQL track DE plan: week 1	4 live-coding questions on joins, aggregations and window functions, including one gaps-and-islands problem.
MI02	Python coding round	After the Python track DE plan: week 2	3 problems on data parsing and DSA basics, written and run live, with complexity discussed.
MI03	PySpark and Spark internals	After Phase 3	PySpark coding plus questions on stages, shuffles, partitions, skew and joins.
MI04	Advanced SQL and data modeling	Phase 7	Hard SQL patterns followed by a schema design question: grain, facts, dimensions, SCDs.
MI05	Databricks and Delta Lake deep dive	Phase 7	Delta internals, MERGE, medallion design, optimization, and walking through your Project 1.
MI06	Snowflake, dbt and Airflow scenarios	Phase 7	Scenario questions: late files, failed DAGs, backfills, incremental models, warehouse costs.
MI07	Streaming and AWS scenarios	Phase 7	Kafka and Structured Streaming questions plus AWS service choices, IAM and cost trade-offs.
MI08	Data pipeline system design	Phase 7	One full design round, whiteboard style, with follow-ups on scale, failure and cost.
MI09	Project deep dive and managerial round	Phase 7	Cross-questioning on your projects, behavioral questions, and the hiring manager conversation.
MI10	Full-loop simulation	Final week	A mixed technical and HR round run like a real final loop, followed by a written readiness report.

Scorecard after every mock

Problem solving, code quality, communication and technical depth, each rated with comments, plus the three fixes to work on before the next mock.

INCLUDED FREE

DOUBT LECTURES AND INTERVIEW SUPPORT

8 doubt lectures (4 in the DE plan) · live support for 10 real interviews · no extra cost

Additional doubt lectures

Code	Lecture	When	Plan	Focus
DL1	SQL doubts I	After S08	Complete only	Joins, subqueries, CTEs, windows, anything open from S01 to S08.
DL2	SQL doubts II	After S16	Complete only	Modeling, SCDs, tuning, interview patterns, before MI01.
DL3	Python doubts I	After P06	Complete only	Core Python, functions, files, OOP and project structure.
DL4	Python doubts II	After P12	Complete only	pandas, APIs, the mini ETL and coding patterns, before MI02.
DL5	DE doubts I	After D06	Both plans	Spark, Databricks, Delta and medallion design, while Project 1 is live.
DL6	DE doubts II	After D09	Both plans	Snowflake, dbt and Airflow, while Project 2 is live.
DL7	DE doubts III	After D12	Both plans	Kafka, streaming and AWS, while Project 3 is live.
DL8	DE doubts IV	After D15	Both plans	Quality, CI/CD, system design and full-stack revision before the mock block.

Live interview support (up to 10 real interviews)

Stage	What happens
Before the interview	Prep call with Durgesh on the JD, the company's round pattern, the likely questions, and which of your projects and stories to lead with.
After the interview	Debrief on what was asked and how you answered, the gaps it exposed, and a short fix plan before your next interview.

BEFORE YOU START

TOOLS, FEES AND NEXT STEPS

What you will install · what you pay · how to begin

Tools you will use

Area	Tools
SQL	PostgreSQL (or MySQL), DBeaver
Python	Python 3.11+, VS Code, Jupyter, pandas, pytest, SQLAlchemy, requests
Processing and lakehouse	Apache Spark / PySpark, Databricks Free Edition, Delta Lake
Warehouse and transformation	Snowflake (trial account), dbt Core
Orchestration	Apache Airflow (Docker)
Streaming	Apache Kafka (Docker), Spark Structured Streaming
Cloud	AWS: S3, IAM, Glue, Athena, Redshift, Lambda
Engineering	Git, GitHub, GitHub Actions, Docker

COMPLETE PLAN

₹2,00,000

SQL + Python + Data Engineering + 10 mocks. Up to 10% off: lowest ₹1,80,000.

DE PLAN

₹1,30,000

For those who already know SQL and Python. Up to 10% off: lowest ₹1,17,000.

Good to know

This is a skills-and-proof program. It prepares you for data engineering interviews and gives you work to show for it, but it does not guarantee a placement, a salary or an interview result. Cloud usage beyond free tiers and trial credits is paid by the learner.

How to start

Book a short call with Durgesh on prepnplaced.com. You pick the plan together, settle the discount if one applies, confirm your weekly slots, and your first session (M01, the skill audit) goes on the calendar.