



2026 | AI-DRIVEN DATA ENGINEERING CAREER PROGRAM

BUILD A JOB-READY DATA ENGINEERING CAREER

**We prepare you not only for your first offer,
but for every project that follows.**

Industry-driven training by working professionals. Learn through real-world pipelines, production-level projects, interview preparation from Week 1 and structured career guidance.

Placement assistance and career preparation. Not a job guarantee.

SQL • Python • PySpark • AWS • Azure
Snowflake • Databricks • Airflow • Kafka • Agentic AI

LEARN | PRACTISE | BUILD | DEPLOY

itmentor.co.in • itmentor.co.in@gmail.com • 9307431763 / 9730096184

— INTRODUCTION

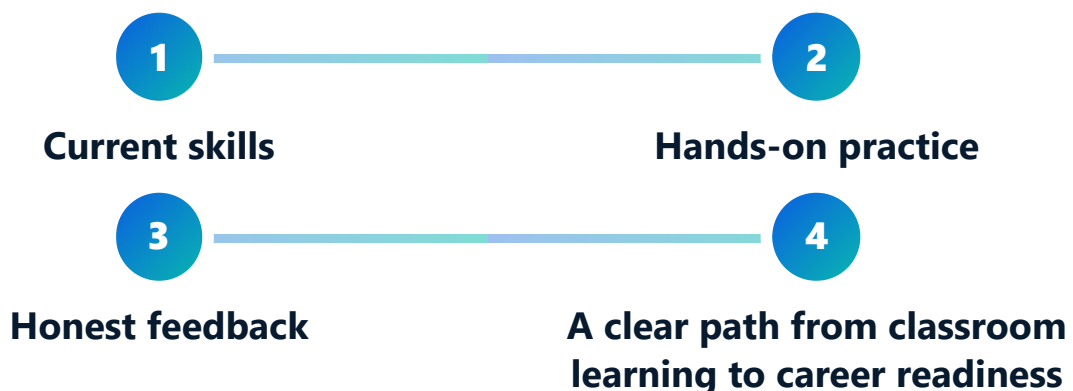
Why we started ITMentor

ITMentor was created to build industry-ready professionals, not just to complete a syllabus. Traditional teaching often ends when the lecture ends, leaving learners unsure whether they can solve real problems, clear interviews or contribute confidently to a live project.

AI is rapidly changing tools, roles and engineering expectations. Drawing on insights from 1,000+ candidate interviews, we know that theory alone is rarely enough. Learners need current skills, hands-on practice, honest feedback and a clear path from classroom learning to career readiness.

Theory alone is rarely enough.

What learners need



ITMENTOR PHILOSOPHY

A job offer is a milestone— not career readiness.

IT careers evolve with projects, client requirements, business priorities and changing technologies. A project may last a few months or several years, but your career must continue beyond it.

That is why preparing only to clear one interview or memorising a fixed set of answers is not enough. You must understand Data Engineering deeply—how data moves, why technologies are selected, how pipelines are built, how failures are handled and how solutions adapt to new requirements.

This level of preparation requires guidance from professionals who actively work in the industry and understand real engineering environments. ITMentor therefore goes beyond syllabus completion. We help learners build the understanding, practical ability and independent learning mindset needed to grow as Data Engineers throughout their careers.

ITMENTOR PHILOSOPHY



**We do not prepare you
only for your first offer.
We prepare you to
remain capable beyond
your first project.**

"Projects change. Technologies evolve. Your ability to understand, adapt and deliver is what sustains your career."

WHY ITMENTOR

Prepared for the industry, not just a syllabus

A practical learning journey built around current technology, real work expectations and career outcomes.

Why choose ITMentor?

01

Mentors from Industry

Learn from working IT professionals with 8+ years of real-world experience.

02

Current Learning

Stay aligned with AI adoption, evolving tools and hiring expectations.

03

Production Practice

Work on realistic pipelines, failures, debugging and deployment decisions.

04

Complete Career Preparation

Build projects, profile, interview confidence and workplace readiness.

OUR AIM

By the end of the program, you should have the skills, project evidence, interview preparation and confidence to take your next career step.

ITMentor provides placement assistance and career preparation. Outcomes depend on each learner's effort, eligibility, performance and market conditions.

THE DIFFERENCE

Traditional teaching vs the ITMentor approach

TRADITIONAL TEACHING

Focuses mainly on completing topics

Limited connection to live project work

Progress is rarely measured clearly

Project explanation remains weak

Interview preparation starts too late

Little guidance after course completion

THE ITMENTOR APPROACH

Learning mapped to industry expectations

Practical implementation with every module

Regular progress reviews

Production thinking and troubleshooting

Interview preparation from Week 1

Placement assistance and post-placement guidance

PROGRAM OVERVIEW

A focused 22-week journey

For fresh graduates, IT professionals, ETL/BI developers and career switchers.

5 to 5.5

MONTHS

22 structured weeks

240

HOURS

Guided live and lab hours

20–25

LEARNERS

Learners per batch

Projects

PRODUCTION-LEVEL

With real-world use cases

By the end of the program you will be able to

- Write maintainable, performant SQL and Python transformation code
- Design batch, incremental, API and streaming ingestion pipelines
- Deploy cloud data pipelines on AWS and Azure
- Orchestrate, test, monitor and document pipelines with Airflow and dbt
- Model data for analytics and implement SCD2 history
- Build and tune distributed Spark jobs and read the Spark UI
- Run warehouse and lakehouse workloads on Snowflake and Databricks
- Use AI assistants responsibly and prototype safe, read-only AI tooling

PROGRAM OVERVIEW

Admission process

Four steps from your first conversation to a confirmed seat in the limited batch.

1

Career discussion

Understand your goals and availability.

2

Skill assessment

Identify your current SQL and Python level.

3

Learning-path review

Recommend suitable preparation.

4

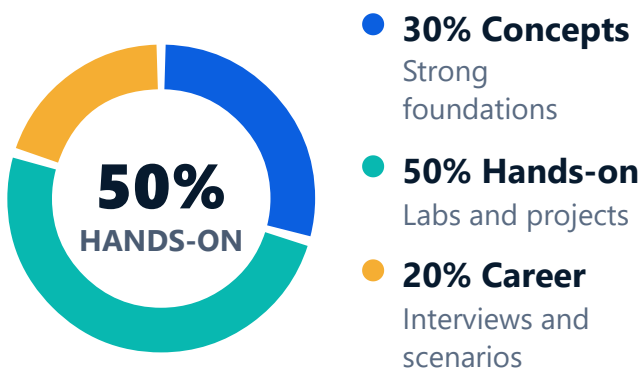
Admission confirmation

Secure your seat in the limited batch.

HOW YOU LEARN

Learning model and weekly rhythm

Learning model



Who can join

- Fresh graduates from technical or analytical backgrounds
- Non-IT graduates willing to work consistently
- IT professionals seeking a Data Engineering transition
- ETL, BI, database, support and backend professionals
- Candidates who completed another course but lack confidence
- Working professionals seeking structured preparation

No previous Data Engineering job title is mandatory. Basic computer understanding, commitment and consistent practice are necessary.

Weekly rhythm

- 1 Weekday live sessions and a weekend lab or review block
- 2 About 10 guided hours a week, plus 5 to 7 hours of self-practice
- 3 Every lecture recorded; trainer connect for doubts after recordings
- 4 Smaller weekly tasks that fit around a full-time job
- 5 Regular progress reviews on assignments and projects
- 6 Weekend interview preparation and mocks are additional and do not reduce the 240-hour core syllabus

Recordings support learning; they do not replace guidance. Watch, practise, ask, resolve doubts, repeat.

AI-DRIVEN DATA ENGINEERING

Why AI-driven Data Engineering matters

Modern Data Engineers are expected to do more than build pipelines. They must understand how to use AI tools responsibly to accelerate engineering work, investigate failures, improve data quality, generate documentation and support better technical decisions—without compromising security, accuracy or engineering fundamentals.

Has strong SQL, Python, PySpark and cloud fundamentals

+

Validates AI-generated outputs before implementation

=

Combines engineering judgement with modern AI capabilities

01

AI-Assisted Development

Use AI tools to support SQL, Python and PySpark development while understanding, reviewing and validating every output.

02

Intelligent Data Quality

Learn how AI-assisted methods can support anomaly identification, data validation and the investigation of unexpected data behaviour.

03

Faster Troubleshooting

Use AI as an engineering assistant to analyse errors, understand logs, identify possible root causes and evaluate solutions.

04

Documentation and Testing

Use AI responsibly to support technical documentation, test-case creation, code explanation and workflow standardisation.

05

Monitoring and Production Support

Understand how AI can assist with incident analysis, pipeline monitoring, performance investigation and production decision-making.

06

Agentic AI Workflows

Understand the fundamentals of AI agents, tool-based workflows, human approval, safety controls and their practical applications in Data Engineering.

AI-DRIVEN DATA ENGINEERING

Traditional engineer vs AI-driven engineer

TOOLS-ONLY APPROACH

- Understands Data Engineering tools
- Writes pipelines manually
- Performs repetitive investigation
- Depends heavily on manual documentation
- May use AI without proper validation

AI-DRIVEN APPROACH

- Has strong SQL, Python, PySpark and cloud fundamentals
- Uses AI to accelerate engineering work
- Validates AI-generated outputs before implementation
- Uses AI for debugging, documentation and testing
- Understands security, governance and responsible AI usage
- Combines engineering judgement with modern AI capabilities

AI does not replace engineering knowledge. It increases the value of engineers who have strong fundamentals, critical thinking and the ability to use AI responsibly.

Students learn where AI can improve engineering productivity, where human judgement is essential and how to verify every AI-assisted output before using it in a production environment.

Confidential company or client data must never be entered into public AI tools.

Where AI fits in the program: integrated across the program and concentrated in Weeks 20–22.

22-WEEK ROADMAP

Week-by-week plan: months 1 and 2

Suggested sequence; concepts, labs and projects overlap. Hours are guided hours (live plus supervised lab).

Month 1 Weeks 1–4 42 hrs	Month 2 Weeks 5–8 41 hrs	Month 3 Weeks 9–13 54 hrs	Month 4 Weeks 14–17 47 hrs	Month 5 Weeks 18–22 56 hrs
WK	FOCUS AND TOPICS	HRS	MILESTONE	
01	Month 1 • Foundations: Linux, Git, SQL		42 guided hours	
01	Data Engineering Foundations, Linux and Git DE fundamentals and lab setup; Linux navigation, permissions, pipes; Git basics	10	Lab ready; first GitHub repo	
02	Linux, Git and SQL Basics Shell scripting; branching and pull requests; SELECT, filters, aggregates, constraints	10	Shell automation script	
03	Core and Intermediate SQL Joins, set operations, subqueries, CTEs, views, conditional aggregation	11	Graded SQL set 1	
04	Advanced SQL Window functions, analytics patterns, plans and indexes, MERGE, transactions	11	Checkpoint 1: SQL assessment	
02	Month 2 • Modelling and Python engineering		41 guided hours	
05	Data Modelling and Data Quality Dimensional modelling, SCD1/2/3, late-arriving data; quality dimensions and reconciliation	11	Analytics modelling lab	
06	Python Foundations Collections, functions, files, exceptions, modules, virtual environments	10	50-problem Python set	
07	Advanced Python and NumPy OOP, generators, decorators, logging, pytest, NumPy vectorisation	10	Tested utility package	
08	Pandas, APIs and ETL Pandas cleaning, groupby, merge; REST APIs, pagination, retries; database loads	10	API ingestion lab	

22 - WEEK ROADMAP

Week-by-week plan: months 3 and 4

Big Data, PySpark, cloud platforms and warehouses.

WK	FOCUS AND TOPICS	HRS	MILESTONE
03	Month 3 • Big Data, PySpark and the start of cloud		54 guided hours
09	Hadoop, Hive and Data Quality HDFS, YARN, MapReduce idea, Hive tables; validation rules, quarantine and idempotency	10	Hadoop-to-cloud note
10	PySpark Foundations Architecture, DataFrames, schemas, file I/O, column expressions	11	Spark labs 1 to 20
11	Advanced PySpark and Spark SQL Aggregations, joins, windows, complex types, UDF trade-offs, partitioned writes	11	Spark labs 21 to 40
12	PySpark Project Integration and AWS Foundations Shuffle, skew, AQE, Spark UI, caching; AWS account safety, IAM, S3	11	CDC and SCD2 lab
13	AWS Data Engineering S3 lake zones, DMS full load and CDC, Glue crawlers and jobs, Data Catalog	11	Raw-to-curated Glue job
04	Month 4 • Cloud and warehouses		47 guided hours
14	Advanced AWS and Azure Foundations Athena, Lambda, Step Functions, EMR and Redshift overview, CloudWatch; Azure identity and ADLS	12	AWS pipeline lab
15	Azure Data Engineering and GCP Exposure Data Factory pipelines, triggers, incremental copy, Databricks link, Key Vault, Azure Monitor; GCP exposure	11	Azure pipeline lab
16	Snowflake Foundations Architecture, warehouses, stages, COPY INTO, VARIANT, roles and grants	12	Loaded JSON event model
17	Advanced Snowflake and Databricks Foundations Streams, tasks, Time Travel, masking, query profile; lakehouse, Delta and Unity Catalog basics	12	Incremental fact in Snowflake

22-WEEK ROADMAP

Week-by-week plan: month 5

Modern platform, orchestration, streaming and career readiness.

WK	FOCUS AND TOPICS	HRS	MILESTONE
05	Month 5 • Modern platform, orchestration and streaming		56 guided hours
18	Databricks and Delta Lake Medallion layers, MERGE, Auto Loader, OPTIMIZE, workflows and jobs	11	Medallion pipeline; first full mock interview
19	Apache Airflow DAG design, TaskFlow, sensors, retries, backfills, alerts, DAG testing	11	Orchestrated multi-step DAG
20	dbt, Monitoring and Agentic AI Integration Models, tests, snapshots, incremental, docs; freshness and alerting; AI-assisted data engineering	12	ELT orchestration lab
21	Kafka, Streaming and Project Integration Topics, partitions, consumers, Structured Streaming, watermarks, checkpoints	11	Streaming recovery lab
22	Docker, CI/CD, Observability, Project Demo and Career Readiness Dockerfiles, compose, GitHub Actions, release flow; pipeline observability; final demo and mocks	11	Checkpoint 5: Pipeline code review

Total guided hours across 22 weeks

42 + 41 + 54 + 47 + 56

240 hours

PROGRAM COMPLETION

Weeks 20–22 integrate AI-driven Data Engineering, the final capstone and career readiness. Learners complete architecture review, implementation, testing, documentation, demo and mock interviews.

REVIEWS AND CHECKPOINTS

- Skill-gap reviews at week 8, week 18 and week 22
- First full mock interview in week 18
- Checkpoints include a SQL assessment and a pipeline code review

The exact weekly sequence may be adjusted based on batch progress and industry updates without reducing the overall learning outcomes.

CURRICULUM

Modules 1 and 2

01

Foundations, Linux and Git

FOUNDATIONAL Weeks 1–2

WHAT YOU LEARN

- Data Engineering fundamentals and lab setup
- Linux navigation, permissions, pipes
- Shell scripting
- Git basics, branching and pull requests

PRACTICAL WORK

- Set up your lab environment
- Create your first GitHub repository
- Write a shell automation script

OUTCOME

- Work comfortably in a Linux terminal
- Use Git the way project teams do

Project link: *Your project repositories, branching model and automation scripts start here.*

02

SQL and Data Modelling

FOUNDATIONAL Weeks 2–5 · 44 hours

WHAT YOU LEARN

- RDBMS basics, constraints and NULL semantics
- Joins, subqueries, CTEs, views
- Window functions and analytical patterns
- Execution plans, indexes, MERGE and incremental loads
- Dimensional modelling, SCD1/2/3, late-arriving data

PRACTICAL WORK

- 60+ graded query problems on a synthetic e-commerce schema
- Diagnose and fix five slow queries with before/after timings
- Build a source-to-target reconciliation script

OUTCOME

- Write correct, readable and performant SQL
- Model data for analytics and implement SCD2 history

Project link: *Designs the SQL layer and data model of your production-level projects.*

CURRICULUM

Modules 3 and 4

03 Python for Data Engineering

PROGRAMMING Weeks 6–8 · 28 hours

WHAT YOU LEARN

- Language core: collections, functions, files, exceptions
- OOP, generators, decorators, context managers
- Logging, configuration, pytest, retries
- Pandas cleaning, REST APIs, pagination and database loads
- NumPy vectorisation

PRACTICAL WORK

- 50 graded problems from basics to OOP
- Stream-parse a 1 GB log with generators and regex
- Package an ETL utility with config, logging, retries and tests

OUTCOME

- Write tested, reusable pipeline code
- Ingest data from APIs and files reliably

Project link: *Builds the ingestion and validation code in your projects.*

04 Hadoop, Hive and Data Quality

BIG DATA Week 9

WHAT YOU LEARN

- HDFS, YARN and the MapReduce idea
- Hive tables
- Validation rules, quarantine and idempotency
- Reconciliation patterns

PRACTICAL WORK

- Create and query Hive tables
- Write a Hadoop-to-cloud note

OUTCOME

- Understand where big-data concepts came from and how they map to the cloud
- Apply data-quality rules in pipelines

Project link: *Data-validation gates and quarantine logic used in your projects.*

CURRICULUM

Modules 5 and 6

05 PySpark and Spark SQL

BIG DATA Weeks 10–12 · 44 hours

WHAT YOU LEARN

- Driver, executors, stages and tasks; lazy evaluation
- DataFrames, schemas, file I/O (CSV, JSON, Parquet, JDBC)
- Joins, windows, complex types, UDF trade-offs
- Shuffle, skew, AQE, caching and the Spark UI
- Reading from S3 and ADLS; testing with pytest

PRACTICAL WORK

- 40 DataFrame challenges from filters to windows
- Spark UI forensics on a supplied slow job
- Banking CDC + SCD2 pipeline: raw to curated Parquet with reconciliation report

OUTCOME

- Transform large datasets
- Explain exactly why a Spark job is slow and fix it

Project link: *Core transformation layer of your production-level projects.*

06 AWS Data Engineering

CLOUD Weeks 12–14 · 34 hours

WHAT YOU LEARN

- IAM, S3 lake zones, KMS basics
- DMS full load and CDC
- Glue crawlers, jobs and Data Catalog
- Athena, Lambda, Step Functions, EMR and Redshift overview
- CloudWatch alerts and cost control

PRACTICAL WORK

- Secure S3 lake with raw/processed/curated zones
- DMS full load and CDC from RDS to S3
- Glue job with bookmarks, queried in Athena
- CSV vs Parquet scan-cost comparison

OUTCOME

- Build a secure, monitored and cost-aware lake pipeline on AWS

Project link: *The AWS half of the multi-cloud ingestion and storage layer.*

CURRICULUM

Modules 7 and 8

07 Azure Data Engineering

CLOUD Weeks 14–15 · 26 hours

WHAT YOU LEARN

- Subscriptions, Entra ID, RBAC, managed identities
- ADLS Gen2 and Azure SQL
- Data Factory pipelines, triggers, incremental copy
- Key Vault, Azure Monitor, ADF with Databricks
- AWS-to-Azure service mapping

PRACTICAL WORK

- Parameterised ADF pipeline for many tables
- Incremental copy from Azure SQL
- Failure alert via Azure Monitor

OUTCOME

- Apply the same pipeline thinking on Azure and map services across clouds

Project link: *The Azure half of the multi-cloud pipeline.*

08 GCP Exposure

CLOUD Week 15

WHAT YOU LEARN

- GCP data services overview
- Mapping GCP services to AWS and Azure equivalents

PRACTICAL WORK

- Service-mapping exercise

OUTCOME

- Discuss GCP intelligently in interviews and read GCP-based job descriptions

Project link: *Broadens the multi-cloud service-mapping document.*

CURRICULUM

Modules 9 and 10

09

Snowflake

WAREHOUSE Weeks 16–17 · 26 hours

WHAT YOU LEARN

- Architecture, warehouses, stages, COPY INTO
- VARIANT and semi-structured data
- Streams, tasks, Time Travel, cloning
- Masking, RBAC, query profile and cost governance

PRACTICAL WORK

- Load nested JSON events with VARIANT and FLATTEN
- Stream plus task incremental fact table
- Set a resource monitor and compare warehouse sizes

OUTCOME

- Load, transform, secure and tune a modern cloud warehouse

Project link: *Serving layer of your production-level projects.*

10

Databricks and Delta Lake

WAREHOUSE Weeks 17–18 · 26 hours

WHAT YOU LEARN

- Lakehouse idea, workspace, clusters, Unity Catalog
- Delta ACID, MERGE, time travel, VACUUM
- Medallion bronze/silver/gold, Auto Loader
- Workflows, OPTIMIZE, Z-ORDER and liquid clustering

PRACTICAL WORK

- Medallion pipeline with Auto Loader
- Late-arriving updates via MERGE
- Recover from a bad write using time travel

OUTCOME

- Build a reliable lakehouse with governed tables

Project link: *Curated layers and quality expectations in your projects.*

CURRICULUM

Modules 11 and 12

11

Apache Airflow

ORCHESTRATION Week 19

WHAT YOU LEARN

- DAG design and TaskFlow
- Sensors, retries, backfills, alerts
- DAG testing

PRACTICAL WORK

- Orchestrated multi-step DAG with retries and alerts

OUTCOME

- Orchestrate multi-step pipelines that recover safely

Project link:

Orchestrates the end-to-end project run.

12

dbt and Monitoring

ORCHESTRATION Week 20

WHAT YOU LEARN

- dbt models, tests, snapshots, incremental, docs
- Freshness checks and alerting

PRACTICAL WORK

- ELT orchestration lab with dbt tests and docs

OUTCOME

- Model warehouse data with tests and documentation

Project link:

dbt marts and tests for the Snowflake layer.

CURRICULUM

Modules 13 and 14

13

Kafka and Structured Streaming

STREAMING Week 21

WHAT YOU LEARN

- Topics, partitions, consumers
- Structured Streaming, watermarks, checkpoints

PRACTICAL WORK

- Streaming recovery lab

OUTCOME

- Explain batch vs streaming trade-offs and recover a failed stream

Project link:

Streaming extension of your projects.

14

Docker, CI/CD and Observability

ORCHESTRATION Week 22

WHAT YOU LEARN

- Dockerfiles and compose
- GitHub Actions and release flow
- Pipeline observability

PRACTICAL WORK

- Containerise a pipeline
- CI workflow with tests
- Checkpoint 5: pipeline code review

OUTCOME

- Ship pipeline code with automated tests and monitoring

Project link:

CI/CD and observability for the final project.

CURRICULUM

Module 15

15 AI for Modern Data Engineering

AI Integrated across the program; concentrated in Weeks 20–22

WHAT YOU LEARN

- Introduction to Generative AI for Data Engineers
- Effective Prompt Engineering for Technical Work
- AI-Assisted SQL Development and Optimisation
- AI-Assisted Python and PySpark Development
- Debugging Pipelines with AI Support
- Generating Test Cases and Technical Documentation
- AI-Assisted Data Quality and Anomaly Investigation
- Using AI for Log Analysis and Root-Cause Investigation
- Introduction to Agentic AI Workflows
- Human-in-the-Loop Review and Approval
- Responsible AI Usage
- Data Privacy, Security and Governance
- Verifying AI-Generated Code and Technical Answers
- Practical AI Use Cases in Production Data Engineering

PRACTICAL WORK

- Review and correct AI-suggested SQL and PySpark code
- Use AI to analyse pipeline logs and propose root causes, then verify them
- Generate tests and documentation for a pipeline and review them

OUTCOME

- Use AI as a professional engineering assistant, not a substitute for knowledge
- Verify every AI-assisted output before it reaches production

Project link: *AI-assisted development, testing, documentation and troubleshooting applied to your production-level projects.*

CURRICULUM

Modules 16 and 17

16 Agentic AI for Data Engineering

AI Weeks 20–22

WHAT YOU LEARN

- Using AI assistants responsibly
- Read-only AI tooling for metadata Q&A and log triage
- RAG basics, guardrails and responsible use

PRACTICAL WORK

- Prototype a safe, read-only AI helper for pipeline metadata

OUTCOME

- Use AI to speed up engineering work without compromising safety

Project link: *AI-support layer (read-only) around your projects.*

17 Production Projects Integration

CAREER Throughout; final weeks 20–22

WHAT YOU LEARN

- Requirements to deployment, monitoring and recovery
- Architecture diagram and design decisions
- Runbook, cost notes and documentation

PRACTICAL WORK

- Working pipelines from source to dashboard
- 15-minute demo and architecture viva

OUTCOME

- Build, debug and explain production-level projects confidently

CURRICULUM

Module 18

18 Interview and Career Preparation

CAREER Week 1 onwards

WHAT YOU LEARN

- Resume, LinkedIn, GitHub and job-portal profiles
- Technical, scenario and system-design rounds
- HR and behavioural rounds
- Mock interviews with written feedback

PRACTICAL WORK

- Saturday interview preparation
- Alternate-Sunday mock interviews
- Project storytelling in 2 and 10 minutes

OUTCOME

- Handle technical and HR interviews with confidence

Project link: *You practise explaining your own project.*

How progress is measured

- Graded assignments and lab sets in every module (for example a graded SQL set and a 50-problem Python set)
- Skill-gap reviews at week 8, week 18 and week 22
- Final architecture review, implementation, testing, documentation and demo
- Checkpoints through the program, including a SQL assessment and a pipeline code review
- First full mock interview in week 18

PROJECTS

Production-level project work

Integrated production-level projects and real-world use cases, built from requirements through deployment, monitoring and recovery.

Production Data Platform capstone

Build a secure, restartable and monitored pipeline using SQL Server and APIs as sources, AWS or Azure for ingestion and storage, PySpark for transformation, Snowflake for serving, dbt for modelling, Airflow for orchestration, and automated data-quality checks, alerts and CI/CD.

WHAT THE LEARNER BUILDS

- Source-to-target architecture and data model
- Batch, incremental and API ingestion
- PySpark transformations and curated layers
- Snowflake models and analytics-ready outputs

YOU DELIVER

- Working pipeline from source to dashboard
- Architecture diagram and design decisions
- Test and quality evidence; failure-recovery demo

PRODUCTION EXPECTATIONS

- Idempotency, retries and safe reprocessing
- Data quality, reconciliation and alerting
- Logging, monitoring, cost and security controls
- Tests, documentation, Git workflow and CI/CD

YOU ALSO SUBMIT

- Cost notes and runbook
- GitHub repository with README
- 15-minute demo and architecture viva

PROJECTS

Production project architecture

ORCHESTRATION Airflow / Step Functions · retries · backfills · alerts

1

SOURCES

SQL Server · REST APIs · Files / events

2

CLOUD RAW

DMS / ADF / Python ingest · S3 or ADLS

3

PROCESS

PySpark on Glue / Databricks · DQ gates

4

CURATED

Delta / Parquet · Silver and gold layers

5

SERVE

Snowflake · dbt marts · tests + docs

6

CONSUME

BI dashboard · Analysts

QUALITY AND MONITORINGreconciliation ·
freshness · alerts ·
runbook**ENGINEERING PRACTICE**Git · Docker · GitHub
Actions CI**AI SUPPORT (READ-ONLY)**metadata Q&A · log
triage

PROJECTS

Project lifecycle

From data sources to the final demonstration, every project follows the same thirteen stages.

1

Data Sources

2

Data Ingestion

3

Raw Storage

4

Data Validation

5

**PySpark
Transformation**

6

Data Quality

7

**Cloud
Processing**

8

**Warehouse
Loading**

9

dbt Modelling

10

**Airflow
Orchestration**

11

Monitoring

12

Documentation

13

**Final
Demonstration**

CAREER PREPARATION

Career readiness and placement assistance

Interview preparation starts in Week 1. Become ready to explain what you built, and why.

PLACEMENT ASSISTANCE

All eligible candidates receive placement assistance until they secure their first job offer, in line with ITMentor's placement-assistance terms. This is assistance with preparation, profile building and interview readiness. It is not a guarantee of a job, a package or a joining date.

Outcomes depend on each learner's effort, eligibility, performance and market conditions.

Your career-preparation path



Roles you can aim for

Data Engineer

Cloud Data Engineer

ETL Developer

Big Data Developer

Analytics Engineer

Data Platform Engineer

Role fit depends on your background, skills and the market.

CAREER PREPARATION

Interview and profile preparation

Career guidance

- Role mapping: which job titles fit your background and the skills to emphasise
- Skill-gap review at three points: week 8, week 18 and week 22
- Guidance on choosing between data engineering, analytics engineering and platform roles
- Honest advice on market expectations, salary research and negotiation basics
- Learning plan for after the program: what to deepen first

Resume and profile

- ATS-friendly resume built around your real projects and honest, measurable outcomes
- Project bullets written in problem, action and result form, never with invented numbers
- LinkedIn profile review and keyword optimisation
- GitHub portfolio: clean repositories, READMEs, architecture diagrams, commit history
- Project storytelling practice: explain it in 2 minutes and in 10

Mock interviews

- Technical rounds: SQL, Python, PySpark and cloud
- Scenario and system-design rounds: batch vs streaming, lakehouse design, failure recovery
- Project deep-dive rounds on your own portfolio
- Behavioural and HR round practice
- Written feedback after every round, with a fix list

Weekend interview track

- Interview preparation every Saturday
- Mock interviews on alternate Sundays
- Technical and scenario-based question practice
- Project explanation and HR round preparation
- Feedback and an improvement plan

Ready to build your Data Engineering career?

Turn learning into project evidence, interview confidence and real career momentum.



LEARN | BUILD | DEPLOY | INTERVIEW | GROW

1 INDUSTRY-RELEVANT SKILLS

SQL, Python, PySpark, cloud, warehouse and orchestration.

2 PROOF THROUGH PROJECTS

Production-level project work with code, tests and deployment evidence.

3 INTERVIEW READINESS

Mock interviews, profile building and role-focused preparation.

4 CAREER SUPPORT

Placement assistance for eligible candidates until the first offer.

Start your journey with ITMentor

CALL

9307431763
9730096184

Call 9307431763

Call 9730096184

WHATSAPP

9307431763
9730096184

WhatsApp 9307431763

WhatsApp 9730096184

EMAIL

itmentor.co.in@gmail.com

WEBSITE

itmentor.co.in

ADDRESS · CLASSROOM + LIVE ONLINE

Ulkanagari, Shrikrishna Colony,
Chhatrapati Sambhajinagar, Maharashtra – 431001

Connect With ITMentor



YouTube
ITMentor



Instagram
@itmentor_



LinkedIn
ITMentor



Facebook
ITMentor

Please speak with our career team for the latest details. ITMentor provides placement assistance and career preparation, not an unconditional job guarantee. Outcomes depend on the learner's effort, eligibility, performance and market conditions.