

AI-First Data Engineer

Build reliable data pipelines and the data foundation for AI

DURATION

16 weeks

FORMAT

**Live online classes with
recordings**

TIME NEEDED

10–12 hours a week

FEE

\$1,500 · ₹1,30,000 approx.

What you'll be able to do

- Write advanced SQL and production-quality Python for data
- Design data models and warehouses (star schema, lakehouse)
- Build batch and streaming pipelines with Spark, Airflow, dbt and Kafka
- Run data on the cloud with Snowflake, BigQuery or Databricks
- Build AI data pipelines: embeddings, vector databases and RAG
- Ship a portfolio-grade end-to-end data platform

Curriculum

01 SQL for data engineers

Joins, window functions and CTEs · Query performance and indexes · SQL interview patterns

02 Python for data engineering

Python for files, APIs and automation · pandas and Polars · Testing and packaging code

03 Data modelling and warehousing

OLTP vs OLAP · Star and snowflake schemas · Slowly changing dimensions

04 ETL and ELT pipelines

Batch ingestion from APIs and databases · Transformations with dbt · Incremental loads and idempotency

05 Orchestration with Airflow

DAGs, scheduling and retries · Backfills and dependencies · Monitoring and alerting

06 Big data with Spark

Spark and PySpark basics · Partitioning and performance · Lakehouse with Delta Lake

07 Streaming data

Kafka fundamentals · Real-time pipelines · Exactly-once and late data

08 Cloud data platforms

AWS, GCP and Azure data services · Snowflake, BigQuery and Databricks · Cost, security and access control

09 Data quality and governance

Data tests and contracts · Lineage and observability · Privacy and compliance

10 AI for data engineers

Embeddings and vector databases · RAG and LLM data pipelines · AI-assisted coding with Claude Code

11 Capstone, interviews and job assistance

End-to-end data platform capstone · SQL, Python and system design mock interviews · Resume, portfolio and referrals

Projects

- Batch ELT pipeline with Airflow and dbt into a cloud warehouse
- Real-time streaming pipeline with Kafka and Spark
- RAG data pipeline with embeddings and a vector database

Tools you'll use

Code and query

SQL

Python

pandas

Jupyter

Git

Pipelines

Apache Airflow

dbt

Apache Spark

Apache Kafka

Docker

Cloud and warehouse

Snowflake

BigQuery

Databricks

AWS

PostgreSQL

AI

Claude

Claude Code

ChatGPT

GitHub Copilot

Pinecone

Replit

Your instructor



Varun Kumar

Founder and lead instructor, SkillNavora. Product manager and author of the 570-page SkillNavora PM Master Handbook.

Career support

Job assistance

- Resume and LinkedIn makeover with 1:1 reviews
- Portfolio of 25+ case studies, capstone and your live product
- Referrals through our network when you're job-ready
- Job-search strategy: target roles, companies and outreach
- Offer evaluation and salary negotiation guidance

Interview preparation

- Weekly mock tests on product sense, metrics and strategy
- 1:1 mock interviews with detailed feedback
- Question bank: product design, RCA, guesstimates, metrics, behavioural
- AI PM and system design rounds for PMs
- Practice with AI between sessions

Fees and scholarships

\$1,500

₹1,30,000 approx. · Includes live classes, real-world projects, capstone, mock interviews and job assistance.

- **Up to 100%: Merit scholarship.** For eligible candidates who do well in the SkillNavora scholarship test and interview. Limited seats every cohort.
- **Students: Students and freshers.** Special pricing for college students and recent graduates, with a valid student or graduation ID.
- **EMI: Flexible payment plans.** Pay the fee in easy monthly instalments, with plans for every cohort.

Book a free counselling call

WhatsApp or call +919211315006 · support@skillnavora.in · www.skillnavora.in/apply

Placement support helps you prepare and apply; it is not a job guarantee. Details may change; the website has the latest information.