DATAXSCHOOL
START YOUR CAREER

DATAXSCHOOL LEARNING CENTRE

START YOUR CAREER

PROGRAM 03 · AI-ERA DATA INFRASTRUCTURE

Advanced Data Engineering with Gen AI

Build the pipelines that feed modern AI. Twenty-four weeks of production-grade engineering across BigQuery, Spark, Kafka, Flink, Beam, Airflow, NiFi and Kubernetes — mentored by architects who run this stack at enterprise scale every day.

6 MONTHS · 24 WEEKS | 13 MODULES | LIVE INSTRUCTOR-LED | PLACEMENT SUPPORT

2000+	95%	₹6L–40L	4
LEARNERS TRAINED	PLACEMENT RATIO	SALARY RANGE	CAMPUSES

ISO 9001:2015 CERTIFIED · REGD. GOVT. OF INDIA · ADMISSIONS OPEN 2026

ABOUT DATAXSCHOOL

Where Ambition Meets Industry

DataXSchool Learning Centre is a career-focused institution specialising in data, AI and emerging technologies. Our purpose is to close the gap between classroom learning and the reality of a working tech career.

Every programme is built around hands-on labs, real production-grade systems and 1:1 mentorship from senior engineers actively working in the industry. Learners don't just study concepts — they build, break, debug and ship. That is why our graduates walk into interviews able to talk about systems they have actually run, and why hiring partners keep coming back to our campuses.

Our Mission

Equip students with practical skills and real-world experience so they are fully prepared and positioned for successful careers in the software and AI industry.

Our Vision

To be a leading institution that transforms aspiring learners into successful IT professionals, bridging education and employment through innovative, hands-on training.

Our Values

Excellence, integrity and impact. We foster an inclusive community where learners are pushed to build, ship and solve real problems with confidence.

Features of DataXSchool

- Dedicated Placement Support
- Live Practical Training
- Interview Prep & Mock Interviews
- Small Batch Sizes
- Free Communication Skills Lab
- Weekday & Weekend Batches
- Industry Expert Trainers
- Real-Time AI-Era Projects
- Resume & LinkedIn Profile Building
- 1:1 Career Mentorship
- Free Demo & Trial Classes
- ISO 9001:2015 Certified Institution

2000+	95%	₹6L–40L	4
LEARNERS TRAINED	PLACEMENT RATIO	SALARY RANGE	CAMPUSES

PROGRAM OVERVIEW

The Engineers Who Feed the AI Era

Modern data engineers do far more than ETL. They build streaming backbones, model warehouses that serve analytics and LLMs alike, orchestrate reliability at scale, and enforce quality and governance on every byte that moves. This program trains you on the exact stack enterprises are hiring for in 2026.

$87B	18% CAGR	₹6L–40L	CRITICAL
MARKET BY 2027	GLOBAL GROWTH	SALARY RANGE	HIRING PRIORITY

What You'll Be Able To Do

- Design dimensional warehouses on Google BigQuery
- Run distributed workloads on Spark, Flink and Beam
- Build real-time streaming platforms with Apache Kafka
- Orchestrate production pipelines in NiFi and Airflow
- Ship containerised workloads on Docker & Kubernetes
- Enforce data quality & governance end to end

Market Demand Signals



Who This Program Is For

- Freshers & final-year students entering tech
- Working professionals switching into data
- Career-break professionals re-entering the industry
- IT professionals upskilling for AI-era roles
- Analysts moving from reporting into engineering
- Engineering, science, commerce or arts graduates
- Software engineers moving into data platforms
- DBAs and support engineers moving into pipelines

Career Roles You Can Pursue

- Data Engineer
- Big Data Engineer
- ETL Developer
- Analytics Engineer
- Platform Engineer
- Pipeline Architect
- Streaming Engineer
- Cloud Data Engineer
- Data Platform Engineer
- Lakehouse Engineer

24	13	100%	2
WEEKS · 6 MONTHS	MODULES	LIVE INSTRUCTOR-LED	SHIPPED CAPSTONES

YOUR ROADMAP

03 Six Months, One Job Offer

What happens month by month — what you learn, and what you have built by the end of it. Every month closes with something you can show a recruiter.

Month 01
Weeks 1–4

Foundations & Data Modeling

Linux · Python · SQL · Git · Erwin · PostgreSQL

YOU'LL BUILD — Retail warehouse in Erwin, live on PostgreSQL.

Month 02
Weeks 5–8

NoSQL & Big Data Processing

MongoDB · Cassandra · Redis · DynamoDB · Spark · PySpark

YOU'LL BUILD — PySpark pipeline over 100M+ records.

Month 03
Weeks 9–12

Real-Time Streaming

Kafka · Schema Registry · Flink · Beam · Dataflow

YOU'LL BUILD — Live clickstream platform with Flink fraud detection.

Month 04
Weeks 13–16

Cloud Warehousing & Orchestration

Google Cloud · BigQuery · Pub/Sub · NiFi · Airflow

YOU'LL BUILD — Cost-tuned BigQuery warehouse, loaded by Airflow.

Month 05
Weeks 17–20

Production Engineering

Docker · Kubernetes · Prometheus · Grafana · ELK · Collibra

YOU'LL BUILD — Monitored, quality-gated platform on Kubernetes.

Month 06
Weeks 21–24

Capstone & Interview Preparation

Capstone · Resume · HR Calls · Technical Calls · Mock Interviews

YOU'LL BUILD — Two shipped capstones, final resume, mock interviews.

Finish

Congratulations — You're Placed

Interview-ready with a shipped portfolio, and in active hiring pipelines with our partner companies.

13 Modules Done · 2 Capstones Shipped · 30+ Tools Mastered · 95% Placement Ratio

02 Foundations, Modeling & Database Engineering

MONTH 01 — Engineering Foundations & Data Modeling

M1 — Data Engineering Foundations, Linux & Git

WEEKS 1–2

- Data engineering lifecycle
- Python for data engineering
- GitHub pull-request workflow
- Linux administration essentials
- Advanced SQL & window functions
- Code review & CI basics
- Shell scripting & cron
- Git branching & merge strategy
- Environment & secrets management

LAB Version-controlled Python ETL job running on a Linux server with scheduled execution

M2 — Data Modeling & Relational Database Systems

WEEKS 3–4

- ER modeling & normalisation
- Slowly changing dimensions
- MySQL administration
- Indexing & query plans
- Dimensional modeling
- Erwin Data Modeler
- PostgreSQL administration
- Partitioning strategies
- Star & snowflake schemas
- Lucidchart architecture diagrams
- Oracle SQL & PL/SQL
- Transactions & isolation levels

LAB Retail warehouse designed in Erwin, documented in Lucidchart, implemented on PostgreSQL

MONTH 02 — NoSQL Systems & Distributed Processing

M3 — NoSQL Database Engineering

WEEKS 5–6

- MongoDB document modeling
- Partition keys & consistency
- Replication & sharding
- Aggregation pipelines
- Redis caching patterns
- CAP trade-offs in practice
- Cassandra wide-column design
- DynamoDB key design & GSIs
- Polyglot persistence patterns

LAB Multi-store backend using MongoDB, Cassandra, Redis and DynamoDB for a single product

M4 — Big Data Processing with Apache Spark & PySpark

WEEKS 7–8

- Spark architecture & execution
- Spark SQL & catalog
- File formats: Parquet, Avro, ORC
- RDDs, DataFrames & Datasets
- Joins, shuffles & partitioning
- Lakehouse table design
- PySpark transformations
- Caching & broadcast tuning
- Spark performance profiling

LAB PySpark batch pipeline processing 100M+ records with tuned partitioning and joins

DETAILED CURRICULUM · MONTHS 3–4 · WEEKS 9–16

03 Streaming, Cloud Warehousing & Orchestration

MONTH 03 — Real-Time & Unified Data Processing

M5 — Real-Time Streaming with Apache Kafka

WEEKS 9–10

- Brokers, topics & partitions
- Producers & delivery semantics
- Consumer groups & offsets
- Kafka Connect source & sink
- Schema Registry & Avro
- Exactly-once processing
- Retention, compaction & tiering
- Cluster sizing & monitoring
- Event-driven architecture

LAB Real-time clickstream platform: Kafka ingestion, schema evolution and sink into the warehouse

M6 — Stream Processing with Apache Flink & Apache Beam

WEEKS 11–12

- Flink runtime & state backends
- Stateful stream operators
- Event time & watermarks
- Checkpointing & recovery
- Beam programming model
- PCollections & PTransforms
- Windowing & triggers
- Batch + stream unification
- Runner portability (Dataflow)

LAB Fraud-detection stream in Flink plus a portable Beam pipeline deployed on Cloud Dataflow

MONTH 04 — Google Cloud Warehousing & Pipeline Orchestration

M7 — Cloud Data Warehousing on Google BigQuery

WEEKS 13–14

- GCP fundamentals & IAM
- Cloud Storage & Pub/Sub
- BigQuery architecture & slots
- Partitioning & clustering
- Nested & repeated fields
- Streaming inserts & batch loads
- Query cost optimisation
- Materialised views & BI Engine
- BigQuery ML & federated queries

LAB Multi-terabyte BigQuery warehouse with partitioned, clustered tables and cost-tuned queries

M8 — ETL & Orchestration with Apache NiFi and Airflow

WEEKS 15–16

- NiFi flow-based ingestion
- Processors & controller services
- Data provenance & back-pressure
- Airflow DAG authoring
- Operators, sensors & hooks
- TaskFlow API & XComs
- Retries, SLAs & alerting
- Dynamic DAGs & backfills
- Cloud Composer deployment

LAB NiFi ingestion feeding an Airflow-orchestrated transformation and BigQuery load, end to end

04 DevOps, Observability & Data Governance

MONTH 05 Production Engineering & Reliability

M9 Containerisation with Docker & Kubernetes

WEEK 17

- Docker images & layers
- Docker Compose stacks
- Kubernetes pods & deployments
- Services, ConfigMaps & secrets
- Helm charts & namespaces
- Spark & Airflow on Kubernetes
- Autoscaling & resource limits
- CI/CD pipelines with Git
- Rolling releases & rollbacks

LAB Containerised data platform deployed to Kubernetes with automated CI/CD from Git

M10 Monitoring & Logging — Prometheus, Grafana, ELK

WEEK 18

- Prometheus metrics & exporters
- PromQL & alerting rules
- Grafana dashboard design
- Elasticsearch indexing
- Logstash parsing pipelines
- Kibana log analytics
- Pipeline SLAs & freshness
- On-call & incident response
- Root-cause troubleshooting

LAB Full observability stack: metrics, dashboards, log search and alerting for live pipelines

M11 Data Quality, Governance & Gen AI Tooling

WEEKS 19–20

- Great Expectations suites
- Apache Griffin measurement
- Collibra catalog & lineage
- Data contracts & stewardship
- PII handling & compliance
- Jupyter Notebooks for analysis
- TensorFlow / PyTorch deployment
- Feature pipelines for ML
- Gen AI copilots for pipeline code

LAB Quality-gated, catalogued pipeline with automated expectations and lineage documentation

01 REAL-TIME STREAMING LAKEHOUSE

02 BIGQUERY ENTERPRISE WAREHOUSE

03 KUBERNETES PIPELINE PLATFORM

04 GOVERNED DATA QUALITY FRAMEWORK

CAPSTONE PROJECTS — CHOOSE, BUILD AND SHIP TWO, REVIEWED BY YOUR MENTOR

05 Capstone, Interview Prep & Placement

MONTH 06 Build, Prepare, Interview, Get Hired

M12 Capstone Projects & Portfolio Build

WEEKS 21–22

- Project scoping & requirements
- Code review & refactoring
- Demo & presentation skills
- Architecture design review
- Pipeline documentation
- Portfolio write-ups
- Build sprints with mentor
- GitHub repository structure
- Project defence Q&A

OUTPUT Two shipped, documented, mentor-reviewed capstone projects on your public GitHub

M13 Interview Preparation & Placement

WEEKS 23–24

- Resume preparation & ATS tuning
- HR call handling & screening
- Technical call handling
- Mock interviews — technical
- Feedback & improvement loops
- Role-targeted resume versions
- Introducing yourself confidently
- SQL & PySpark rapid drills
- Mock interviews — HR
- Salary negotiation
- LinkedIn & GitHub optimisation
- Answering gap & switch questions
- System design discussion rounds
- Case-study & scenario rounds
- Referrals & hiring drives

OUTPUT Final resume, optimised profiles and repeated recorded mock rounds with written feedback

Resume Preparation

Built line by line with your mentor — project framing, impact statements, keyword and ATS tuning, plus separate versions for data engineer, ETL and analytics engineer roles.

HR Call Handling

Structured practice for the screening call: self-introduction, notice period, expectations, career-gap and switch questions, and how to hold your salary position politely.

Technical Call Handling

How to think aloud on a live call — scoping the question, SQL and PySpark under time pressure, pipeline design discussion, and recovering when you don't know an answer.

Mock Interviews

Repeated full-loop mocks with industry mentors simulating real company rounds, each followed by written feedback and a targeted improvement plan before the next round.

COURSE MATERIAL

06 The Complete Tool Stack

Grouped by the job each tool does in a real pipeline — so you learn where every technology sits and why it is there.

01 Ingest › 02 Process › 03 Store › 04 Warehouse › 05 Operate

01	**Model & Design**	DATA MODELING	Erwin Data Modeler · Lucidchart
02	**Ingest & Integrate**	ETL TOOLS	Apache NiFi · Apache Airflow
03	**Stream in Real Time**	PIPELINE ORCHESTRATION	Apache Kafka · Kafka Connect · Schema Registry
04	**Process at Scale**	PROCESSING FRAMEWORKS	Apache Spark · PySpark · Apache Flink · Apache Beam
05	**Store — Relational**	DATABASE SYSTEMS	MySQL · PostgreSQL · Oracle
06	**Store — NoSQL**	DATABASE SYSTEMS	MongoDB · Cassandra · Redis · DynamoDB
07	**Warehouse & Cloud**	WAREHOUSING · CLOUD	Google BigQuery · Google Cloud Platform
08	**Ship & Automate**	CONTAINERS · VERSION CONTROL	Docker · Kubernetes · Git
09	**Observe & Debug**	MONITORING & LOGGING	Prometheus · Grafana · Elasticsearch · Logstash · Kibana
10	**Govern & Trust**	QUALITY & GOVERNANCE	Great Expectations · Apache Griffin · Collibra
11	**Hand Off to ML**	ML & DATA SCIENCE	Jupyter Notebooks · TensorFlow · PyTorch

30+	11	100%	2026
TOOLS & TECHNOLOGIES	PIPELINE STAGES	HANDS-ON LABS	INDUSTRY STANDARD

MEET YOUR MENTORS

Trained By Industry, Not Textbooks

Your mentors are not full-time educators. They are practising architects who design and operate enterprise data platforms every day — and bring that current tooling, judgement and network into every session.



FOUNDER & CHIEF MENTOR

Srikanta Patra

PRINCIPAL DATA ARCHITECT · FOUNDER, DATAXSCHOOL

Srikanta Patra is a seasoned Principal Data Architect with 15+ years of industry experience across Data Analytics, Data Science, Database Engineering and Cloud Data Platforms. Over his career he has held key technical leadership roles at Sequoia, Chingari, Junglee Games, BCD Travel, Atlas Systems, Synchronoss Technologies and Wipro. As Founder & Chief Mentor of DataXSchool he has mentored thousands of students and professionals, combining real-world production experience with a practical, industry-first teaching approach.

EXPERIENCE	EXPERTISE	DOMAINS
15+ Years	Data Engineering & Cloud Platforms	Gaming · Travel · Telecom · Healthcare

FOUNDER AND MENTOR

WORKING ON BELOW COMPANIES

SEQUOIA · JUNGLEE GAMES · BCD travel · Atlas Systems · synchronoss Technologies · wipro



PROGRAM MENTOR

Jitendra Mohanty

DATA ARCHITECT · TECH MAHINDRA · BENGALURU

Jitendra Mohanty is a seasoned Data Architect with 15+ years of experience in Data Engineering, Cloud Architecture, AWS, Kubernetes, Big Data, Data Warehousing and DevOps. He is currently working as a Data Architect at Tech Mahindra, Bengaluru, delivering enterprise solutions for global clients. Previously he served as a Cloud Architect at Boeing and has held key engineering roles at Neustar, Continental and DMX Technologies Group. At DataXSchool he mentors aspiring Data Engineers with practical, real-world industry expertise.

EXPERIENCE	CURRENT ROLE	DOMAINS
15+ Years	Data Architect, Tech Mahindra	Data Engineering · Cloud · DevOps

PROGRAM MENTOR WORKS ON

Continental · neustar · BOEING

15+	1:1	100%	∞
YEARS INDUSTRY EXPERIENCE	PERSONAL MENTORING	LIVE SESSIONS, NO RECORDINGS	DOUBT CLEARING & CODE REVIEW

INSIDE OUR CAMPUSES

07

Built For Real Practice

Modern, fully-equipped campuses across India — dedicated lab seats, instructor-led classrooms and collaborative spaces. Every learner gets a machine, a mentor and the room to build. Visit any campus and sit in on a live session before you decide.

Bengaluru

Brigade IRV Centre, 9th Floor, Nallurhalli, Whitefield 560066

FLAGSHIP CAMPUS







INSTRUCTOR-LED SESSIONS & COLLABORATIVE WORKSPACES — LIVE BATCHES

Bhubaneswar

Satyakali Residency, N-6/452, Nayapalli, Jayadev Vihar 751015

ODISHA





CLASSROOM & PRACTICE LAB — DEDICATED SEATS, SMALL BATCHES

Berhampur

1st Floor, Jail Road, Lanjipali, Odisha 760008

ODISHA





RECEPTION & LEARNING LAB — 1:1 MENTORSHIP

CAREER SERVICES

08 From Day One to Day You're Hired

Our placement engine is built into the curriculum from week one. You don't graduate and then job-hunt — you graduate already in active interview pipelines, with a portfolio of shipped pipelines behind you.

95%	50+	10+	Free
PLACEMENT RATIO	CAREER WEBINARS	INDUSTRY MENTORS	MOCK INTERVIEWS

1:1 Career Mentoring

Every learner is matched with a senior industry mentor for ongoing guidance — career strategy, technical doubts and salary negotiation.

Assured Interviews

On graduating into our Placement Pool you get guaranteed interview opportunities with top MNCs and high-growth startups.

Resume & LinkedIn

Built and reviewed by industry hiring managers. We rebuild your resume, optimise LinkedIn and clean up your Naukri presence.

Mock Interview Prep

Repeated mock rounds simulating real company loops — SQL, PySpark, system design and behavioural — with detailed feedback.

Hiring Job Fairs

Quarterly campus-style fairs bring recruiters directly to our learners across all four campuses.

Hackathons & Projects

Team hackathons and real client briefs build portfolio depth — the kind of work that actually moves the needle in interviews.

Our Student Placement Companies

Organisations where DataXSchool learners have been placed across programs and campuses.

CLOVER INFOTECH · wipro TECHNOLOGY · ahana SYSTEMS · accenture

INTUTATIVE TECHNOLOGY · HELLO AI · Deloitte. · TATA ELEXI

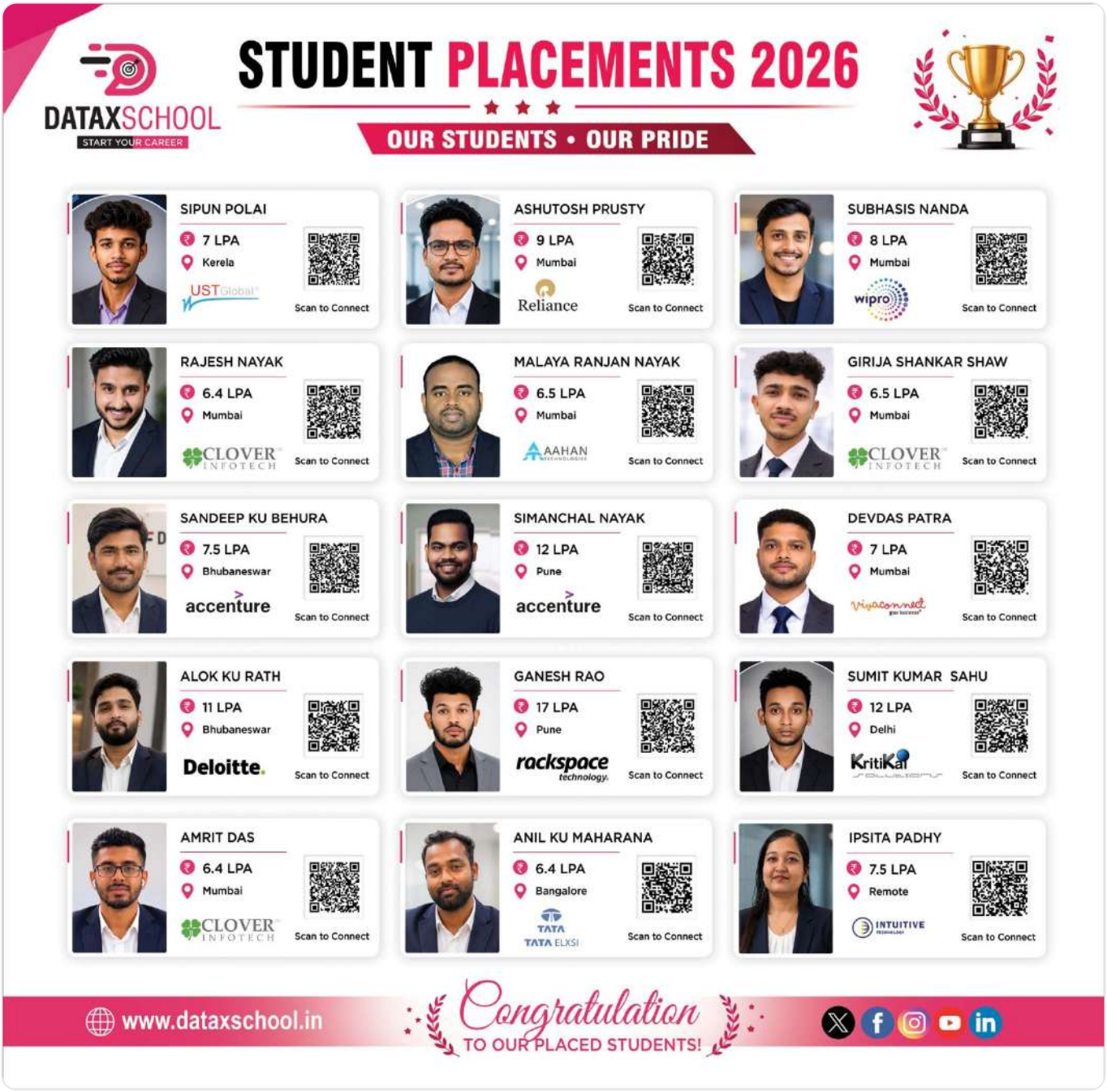
NTT DATA · Tech Mahindra · TCS TATA CONSULTANCY SERVICES · Infosys

IBM · LTIMindtree · rackspace · MAVINIER

OUR STUDENTS · OUR PRIDE

09 Outstanding Placements

2000+	95%	₹6L–40L	Pan-India
LEARNERS	PLACEMENT	SALARY RANGE	REACH



A SELECTION OF RECENT DATAXSCHOOL PLACEMENTS ACROSS PROGRAMS — VERIFIED OFFERS, REAL LEARNERS.

JOIN US

10 How to Apply

Who Can Apply?

- Freshers & final-year students
- Working professionals switching into tech
- Career-break professionals re-entering tech
- Engineering, science, commerce or arts graduates

1 Submit Application

Tell us your background, goals and target role. Takes about five minutes.

2 Screening & Interview

A short aptitude check and a 1:1 conversation with our admissions panel.

3 Admission & Onboarding

Offer letter, enrolment, and immediate access to pre-course material and the portal.

Visit Any of Our Campuses

BENGALURU

Brigade IRV Centre, 9th Floor, Nallurhalli, Whitefield, Bangalore – 560066

+91 90408 93720

BHUBANESWAR

Satyakali Residency, N-6/452, in front of Bank of India, near The Crown Hotel, Nayapalli, Jayadev Vihar, Odisha 751015

+91 78490 07838

BERHAMPUR

1st Floor, Jail Road, Lanjipali, Berhampur, Odisha 760008

+91 84579 54836

FREE TRIAL · ADMISSIONS OPEN 2026

Book a Free Demo Class

Sit in on a live session before you decide. Zero commitment.

+91 84579 54836

info@dataxschool.in · **www.dataxschool.in**