SCALER
DATA SCIENCE

REAL DATA SCIENCE SKILLS THAT DELIVER REAL OUTCOMES!

India's **#1** Outcome-focused, Immersive
Data Science & Machine Learning Program
Developed Based On Insights From
Top 50 Tech Companies

Index

Insights

We designed the **Scaler Data Science curriculum** based on insights from **100 Data Scientists** and **ML Engineers** from **50 top tech companies** and **startups**. Our objective was to identify challenges they face and address them in our course. Here's what we learnt:

95%

struggled initially to understand business problems:

Data Science is a technical role, but **Data Scientists** are tasked with **improving the business**. Data Science programs often focus on the technical aspect, **forgetting** the importance of **understanding the business.**

⭐ **The Scaler Difference:**

✓ Scaler Data Science addresses this with a **Business & Strategy curriculum** designed by experts from **McKinsey** and **BCG.**

✓ It is **integrated seamlessly** into all **projects**, providing **hands-on practice** tackling DS & ML problems in a business scenario.

80%

struggled initially to work with real-world datasets:

Data scientists work with **massive, fragmented, messy,** and **incomplete data**, while students practice on dummy datasets which are small, clean, and organized, **depriving learners** the **opportunity to practice** with **real messy datasets.**

⭐ **The Scaler Difference:**

✓ At Scaler Data Science, you will work on **projects built in partnership** with **top companies.**

✓ You will work on real problem statements and will get your hands dirty by working with **messy and unclean realistic data.**

50%

struggled to understand research papers

Data Science, Machine Learning, and Deep Learning are **rapidly evolving fields**, with **new research** being **published constantly**. Data Scientists who can absorb the latest research insights are of **great value** for **organisations.**

⭐ **The Scaler Difference:**

✓ At Scaler Data Science, you will learn how to **read** the latest **Research Papers**

✓ Expert mentoring to help you **publish new research** in global **peer-reviewed conferences**.
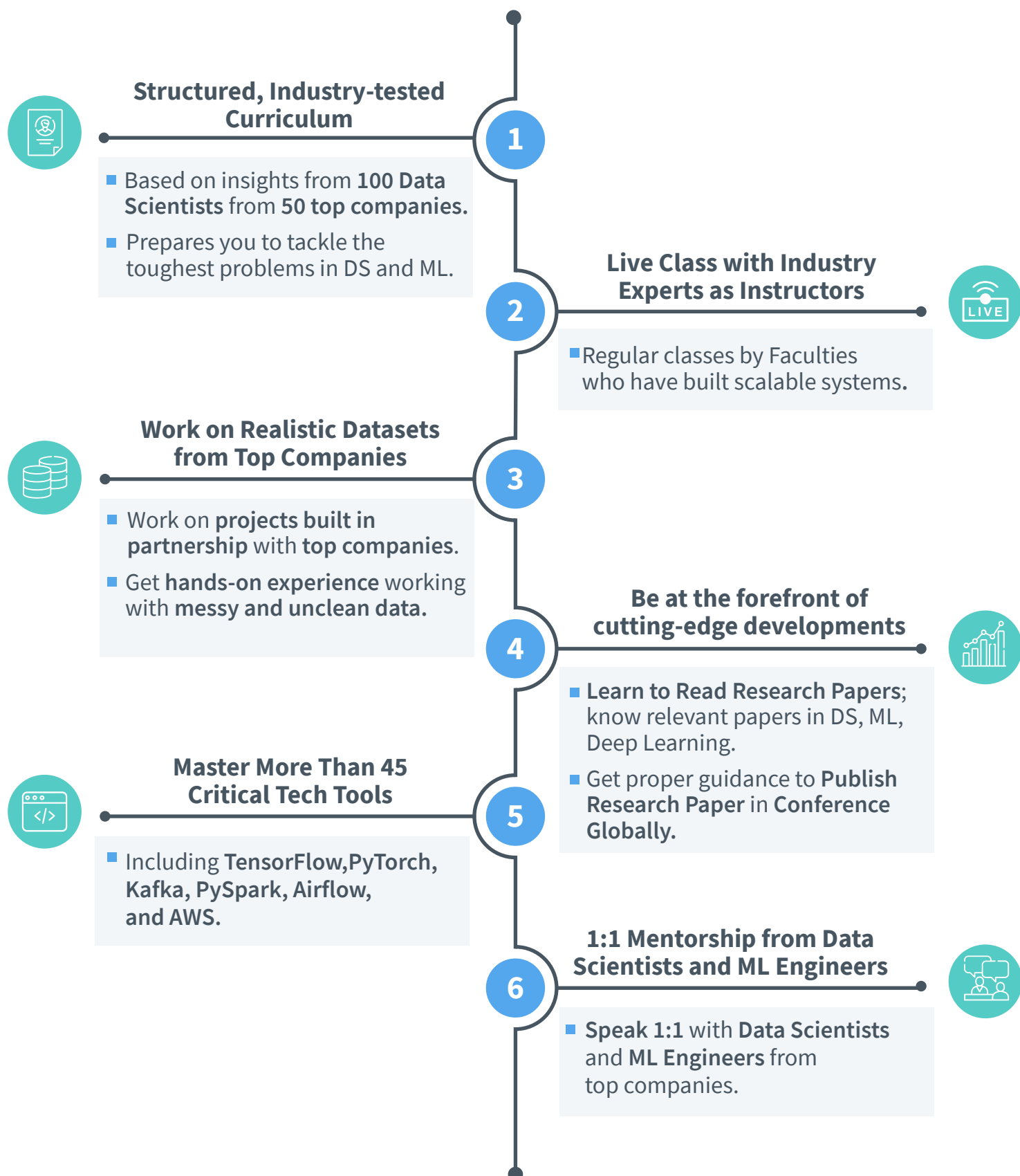
Scaler Data Science program ensures you gain the knowledge & confidence to overcome the toughest challenges a Data Scientist or ML Engineer can face in their careers, from day one!

Highlights of Scaler Data Science Program

A program that focuses on real outcomes - career-wise and skill-wise!

1 — Structured, Industry-tested Curriculum

- Based on insights from **100 Data Scientists** from **50 top companies**.
- Prepares you to tackle the toughest problems in DS and ML.

2 — Live Class with Industry Experts as Instructors

- Regular classes by Faculties who have built scalable systems.

3 — Work on Realistic Datasets from Top Companies

- Work on **projects built in partnership** with **top companies**.
- Get **hands-on experience** working with **messy and unclean data**.

4 — Be at the forefront of cutting-edge developments

- **Learn to Read Research Papers**; know relevant papers in DS, ML, Deep Learning.
- Get proper guidance to **Publish Research Paper** in **Conference Globally**.

5 — Master More Than 45 Critical Tech Tools

- Including **TensorFlow, PyTorch, Kafka, PySpark, Airflow, and AWS**.

6 — 1:1 Mentorship from Data Scientists and ML Engineers

- **Speak 1:1** with **Data Scientists** and **ML Engineers** from top companies.

Program Overview

Curriculum That's Crafted To Make You a Complete Data Scientist!

Course Outline

Sr. no.	Topic	Duration
1	DSA Intermediate	8 weeks
2	Python + Tools + Web	3 weeks
3	Data Mining & Statistical Analysis	7 weeks
4	Data Science & Machine Learning	9 weeks
5	Deep Learning & Big Data	7 weeks
6	DSA Advanced	15 weeks
	Total Course Duration	**11 Months**

Entry Points for this course:
Based on a 30 Min MCQ Entrance coding Test, you will either enter as:

✅ **Intermediate**
1. DSA Intermediate (2 months)
2. Data Science & ML (7 months)
3. DSA Advance (4 months)

✅ **Advance**
1. Data Science & ML (7 months)
2. DSA Advance (4 months)

Curriculum Deep-dive

01 DSA Intermediate

Functions

- Recursion
- Pointers

Structures, Unions & Dynamic Arrays

- Asymptotic Notations

> ❓ **Why DSA?** Strong DSA skills help you in Problem Solving, using the most effective way. Moreover, it is an essential skill for clearing interviews, even for Data Science and ML roles.

02 Preparation

Setup & Utilities

- Linux Terminal, Jupyter
- Git Essentials

Structures, Unions & Dynamic Arrays

- Python, OOP, Exception handling
- Important Modules (time, random, os, openCV, etc.)
- File I/O, Connect to DB, Large files

> ⭐ **Scaler Advantage:** Prior knowledge of Python is not required. Essential tools like Git will be covered so that you don't struggle on the job.

03 Data Mining / Data Analysis

Databases and SQL

- Basics - MySQL, SQL Queries, CRUD, ACID
- Intermediate - Index, Partition, Query optimization, CTE, Concurrency, Sharding
- NoSQL, MongoDB
- DB Schema Design

Data Governance

- Data Dictionary

> ⭐ **Scaler Advantage:** The Program also covers advanced topics such as Data Governance and Data Dictionary which will help you stand out in your career.

04 Maths & Statistical Analysis

Statistics & Probability

- Statistical measures, Standard deviation, Probability, Bayes
- Normal Distribution, Density, Binomial Experiment, CLT
- Poisson, Markov
- Sampling, Confidence, Hypothesis Testing
- Anova, Chi-square

Linear Algebra

- Linear algebra, Eigenvectors
- Coordinate Geometry

Calculus

- Derivatives, Optimization techniques

> ⭐ **Scaler Advantage:** A strong foundation in Mathematics is a gift that keeps on giving. We will cover these topics in a fun and engaging way.

05 Data Analysis in Python

Python

- ✅ Numpy - Vector, Matrix, Array, Slicing, Coercion, Random, Distribution, Time
- ✅ Pandas - Dataframe, Pivot, Merge, Missing values, Apply, Download
- ✅ Matplotlib, Seaborn

> ⭐ **Scaler Advantage: Numpy** and **Pandas** are going to be your air and water as a Data Scientist and ML Engineer. We will teach you how to use these tools to solve difficult **business** problems.

06 Data Science & Machine Learning

EDA (Exploratory Data Analysis)

- ✅ Random Sampling, Replacement, Simulation
- ✅ Summarization, Central Tendency, Dispersion, Skewness, Kurtosis, Law of large numbers
- ✅ R-square, Multicollinearity, Parsimonious, Autocorrelation, Heteroscedasticity
- ✅ Significance, Confusion Matrix, Sensitivity, Specificity, Odds Ratio
- ✅ Outlier Detection, Normality Test (QQ, KS, SW), Correlation Analysis
- ✅ Feature Engineering, Normalization, Standardization
- ✅ Parametric and Non-Parametric Tests

Supervised Models

- ✅ KNN
- ✅ Linear Regression

- ✅ Logistic Regression
- ✅ Ensemble: Bagging, Boosting, Decision Trees, Random Forest
- ✅ Regularization, Overfit
- ✅ SVM

Unsupervised Models

- ✅ Clustering (KMeans, Hierarchy), Anomaly Detection

Predictive Modeling

- ✅ Time Series - MLE, LSM, ARIMA (optional... MAM, JASA, ADF, Correlogram, AIC, LSE)
- ✅ Forecasting - Naive, Moving Avg, Growth, Random Walk, Drift forecast

Intermediate

- ✅ PCA, Dimensionality Reduction, Factor analysis, t-SNE
- ✅ Validation Analysis, Residual Analysis
- ✅ Recommender system (Collaborative/Content filtering, Propensity analysis). Cold start problem

Natural Language Processing

- ✅ Text Analytics, Text Mining, Tf-Idf, POS Tagging, a bag of words, import prereq
- ✅ Naive Bayes, Spam Detection, Sentiment Analysis

> ⭐ **Scaler Advantage:** We will cover the latest models in Data Science and Machine Learning. You will work on **projects built in partnership** with top companies. You will get your hands dirty by working with **messy and unclean real-world data.**

07 Deep Learning

- ✅ Neural Networks, MLP, Backprop

Overhead

- Keras
- TensorFlow
- PyTorch
- Project discussions
- Reading Research Papers

Computer Vision

- CNN, Object Detection
- Object Segmentation, Localization, YOLO, SSD
- Autoencoders
- GAN
- Data Augmentation
- Siamese Net
- Architectures - VGG, GoogleNet, ResNet, DenseNet
- Transfer Learning

Natural Language Processing

- Similiarity, Sequential Models, Embeddings, Transformers, word2vec/GloVe
- Time Series - LSTM, GRU
- Attention Networks / BERT

Other

- Reinforcement Learning, Q-Learning, Autonomous Players
- ML System Design

> ⭐ **Scaler Advantage:** We will teach you how to read the latest **Research Papers** and we will discuss the relevant papers in class. This is an important skill to stay updated in the fast-changing field of ML and Deep Learning.

08 Data Engineering / Data Infrastructure / MLOps

Big Data / Distributed System

- ✅ Warehouse - S3, Hadoop HDFS, ElasticSearch/Solr, HBase, MongoDB, Redshift
- ✅ Ingestion - Kafka, Nifi
- ✅ Analysis - PySpark, Airflow, Hive
- ✅ DevOps - Databricks, Snowflake, KubeFlow, MLFlow, KNIME
- ✅ DevOps - Docker, Kubernetes, Ansible/Puppet, Kerberos, AWS, GCP, Azure
- ✅ System Design - Data Pipeline

> ⭐ **Scaler Advantage:** Scaler Program also covers advanced topics such as Data Governance and Data Dictionary which will help you stand out in your career.

09 Web and API

Basics

- ✅ Web APIs, Downloading large files programmatically

Intermediate

- ✅ Web Scraping & Crawler
- ✅ Web Automation

Advanced

- ✅ Flask
- ✅ Cloud Server Deployment

> ⭐ **Scaler Advantage:** Being able to build a crawler to scrape websites and deploying your **web apps** will help you ace your career as a Data Scientist and ML Engineer.

10 Business Experience

Business Skills

- ✅ Case studies, MoM, Growth, LTV, NPS, DAU, MAU, Profit, Revenue

> ⭐ **Scaler Advantage:** Data Scientists help with **business** decisions and need to solve **open-ended** business problems. They need to know what data to even ask for. As such, understanding the business is a critical skill. It will help you in difficult business situations when you have to **creatively** build a new **strategy** or diagnose issues in the existing **operations** of the company.

11 DSA Advanced

- ✅ Time Complexity, Arrays, Strings, Binary Search, 2 Pointers, Recursion, Hashing, Sorting, Bit manipulation
- ✅ Stacks, Queues, Linked Lists, Trees, Tries, Heap
- ✅ Greedy, DP, Graphs

The Scaler Data Science curriculum covers all the essential topics and tools you need to make real-world impact as a Data Scientist!

Meet Your Instructors

A comprehensive curriculum by itself doesn't make you a complete data scientist - real skills require real-world experience. **Learn from Experts and Thought-leaders who have been there, done that!**

Anshuman Singh
Co-Founder SCALER ACADEMY, Ex - facebook

He is the co-founder of Scaler Academy and two-time ACM ICPC world finalist. He was one of the founding team members of Facebook Messenger and worked directly with Mark Zuckerberg on product development.

Kshitij Mishra
Ex - snapdeal

As part of a two-member SDE team, he managed the entire codebase of Snapdeal 'Seller Search' services (when Snapdeal was doing 7M orders/month) and also helped Snapdeal move from the highest available AWS model to the lowest within 4 months.

Pragy Agarwal
Ex - media.net

As a Platform Engineer at Media.Net, he scaled low latency systems with millions of queries per second. He has 16+ years of coding experience across a wide variety of programming languages & paradigms. He He has a keen eye for 'good coding practices'.

Anant Mittal

UNIVERSITY OF MARYLAND , INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY HYDERABAD

As a researcher at the University of Maryland, he worked on cutting-edge systems to find biomarkers of task activities in the brain. He designed and developed COVID19 & hygiene-related analytics solutions such as temperature screening and violations related to preventive measures.

Mudit Goel

Ex - Linked in, Intuit.

Mudit co-founded Coding Elements, a coding school, impacting the lives of 2 Million+ students across 10 countries. Coding elements was selected by the Government of India to provide coding courses for students across the nation. Mudit currently leads the Data Science and Machine Learning program at Scaler.

Advisory Committee

Ramit Sawhney

Georgia Tech | TOWER RESEARCH CAPITAL

A seasoned software engineer, Ramit is a globally published and recognised researcher at IIIT-Delhi, Georgia Institute of Technology, the AI Institute at University of Southern Carolina, and Open Source Maintainer at AnitaB.org

He has been empanelled as the research advisor for the Scaler Data Science Program

Pawan Kumar

Site Lead, Uber
Ex - Linked in

Pawan is an experienced Data Scientist with a strong product sense and an innate ability to communicate complex insights clearly. He has been leading and spearheading the Data Science divisions at Uber and LinkedIn.

He has been empanelled as an advisor for the Scaler Data Science Program

Mentors

Our mentors ensure you have the right guidance, every step of the way!

We offer 1:1 Mentorship sessions twice a month, wherein mentors will answer your queries, doubts, help you with interview tips, and even mock interview sessions.

Birendra Yadav
Sr. Data Scientist, bright
Ex - Goldman Sachs

He can help you improve your algorithmic thinking and its application. He can help and aid you in preparing for data science interviews.

Dhumil Agarwal
SDE 2 - ML - Panasonic
Ex - Myntra , BITS Pilani

He is proficient in Java, Kafka, Python, Big Data, Data Engineering, Git, Design Patterns, Azure, SQL, Back-end Development.

Aryan Jadon
Data Science Asso. - AUTODESK.
Ex - Moody's

He is a developer with strong foundations in finance and programming having hands-on experience in backend development, data science, and data analytics.

Dheeraj Khanna
ML Engineer - WSe GROUP

He has a blend of Full-Stack engineering skills and Business skills with hands-on experience in building, implementation, and the delivery of end-to-end data science products and solutions.
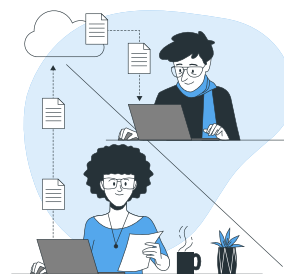
Career Support & Alumni Network

We offer career support and guidance, to ensure the skills you have learned are having a real-world impact!

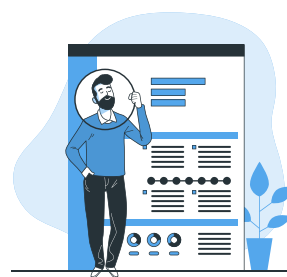Access job opportunities from **600+** partner employers

Exchange job opportunities with **20k+** Scaler student community

Practice mock interviews with people from the industry

Optimize your resume & LinkedIn profile by the experts

Our alumni work at reputed technology companies and promising startups

Tuition Fee

⭐ **Total Fee Rs 2.5L (inclusive of GST)**

✓ You get scholarship upto **Rs 25,000** and there are super affordable EMI and finance options for you to avail as well!

✓ With scholarship & EMI options, your payment can be as low as Rs 7,200 per month - that's like your monthly grocery bill!

✓ 2 weeks trial period - money-back guarantee if you want to exit the course!

Apply In 3 Easy Steps!

1

Easy Registration

In <2 min, make a new account or login using social media / InterviewBit account

Click here to create your account

2

Quick Evaluation

Simple 30 min MCQ test focused on aptitude and basic coding to find the right course for you!

Take The Test Now

3

Enroll in your course

Signup with easy EMI Options and avail scholarships to start your learning journey!

Signup for the Course