# TAKE THE LEAP WITH SCALER DATA SCIENCE

An Industry-focussed program that features Business Case Studies & Datasets from Real Companies

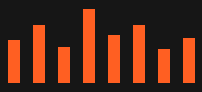
SCALER

Upskilling to
#CreateImpact

# About **Scaler**

Scaler, powered by InterviewBit, is an outcome-focussed leading ed-tech platform for tech professionals. Our industry-vetted approach towards teaching and training young minds helps them upskill and bag the career of their dreams. We are a transformative learning platform devoted towards creating a growth ecosystem to assist software professionals in unlocking their talent and seizing opportunities at every stage of their careers. Learners enroled with us are taught, guided, and mentored by top professionals and experts working at leading organizations including Google, Facebook, Intuit, Microsoft, Amazon, Hotstar, etc. Our learners have witnessed a 5x RoI (Return on Investment) from our program. Our offerings include - Scaler Academy and Scaler Data Science.

# Index



# Industry Insights

About **95%** of professionals wish to understand business problems better

**80%** stumbled as they began working with real-world datasets

The large percentage indicates a clear gap between classroom teachings and industry applications. Enter Scaler.

**Scaler Data Science** is a program curated to help you kick-start your career in Data Science & Machine Learning. We'll make you industry-ready through a rigorous curriculum taught by industry veterans who'll mentor you as you headway towards growth. With our support, you can #CreateImpact in your career and the Data Science world.

Scaler Data Science is designed after procuring industry insights from 100 Data Science & Machine Learning Engineers who come from top 50 tech companies and startups. You'll gain the needed exposure as you work on 80+ Business Case Studies built in partnership with the top companies. Your skills clubbed with our industry expertise will take your Upskilling game to the next level.

# Scaler Program Highlights

Here's what you can expect as you go about upskilling your career with Scaler

**Structured, Industry-vetted Curriculum**


**Live Classes with product industry veterans**

**Hyper-personalised Experience**

**Regular 1:1 Mentorship**

**80+ Business Case Studies with Real-world Datasets**

**20K+ Scaler Learners & Alumni Network**

**A dedicated team of Recruiters**

**600+ Partner Employers & Placement Assistance**

# WHERE WILL YOU FIT IN SCALER DATA SCIENCE?

Our program ensures that you get to grips with the knowledge and confidence to overcome the most formidable challenges a Data Science or Machine Learning engineer can encounter in their journey from day one.

**Entry Points based on your performance in our Coding Entrance Test:**

## Start from Beginner Module

**Recommended for Non-Coders**

A 13 month track with a 1 month beginner module.

## Start from Intermediate Module

**Recommended for 0-2 years of Coding Experience**

A 12 month track for those with minimal exposure to DSA.

## Start from Core Module

**Recommended for 2+ years of Coding Experience**

A 7-10 month Advanced track with Data Science Core.

**Data Science is the nucleus of business, and we spare no effort to put you at the center of it.**

# ADDING SCALER SPICE TO DATA SCIENCE

As a Data Scientist, you'll be using data to solve real-life business problems. Scaler will build your skill to understand business and industry.

- Assess your conceptual understanding as you work on **80 Business Case Studies** with datasets from real companies
- Discuss problems in **Live Case Study Discussion** sessions weekly with a Data Scientist
- Gain a fresh perspective from the submission of your peers on a private **Discussion Forum** and understand a different problem-solving approach
- Stay updated about grades, attendance, streak through the **Live Progress Report Card** on your Dashboard
- Stay on the track with **clear deadlines** - failing which will incur a penalty
- Immediate **evaluation** of your assignments to help assess your progress amongst your peers
- Earn **extra credit** for your dedication & effort to cover up for the unanticipated gaps in your performance
- Learn at your pace with the **flexibility** to pause your course anytime you want

# **CURRICULUM** THAT COMPLEMENTS YOUR CAREER

| TOPIC | TOTAL LECTURES | TOTAL DURATION |
|---|---|---|
| Beginner Programming | 12 | 4 Weeks (1 Month) |
| Intermediate Programming | 12 | 4 Weeks (1 Month) |
| Common Core for DSML | 42 | 14 Weeks (3.5 Months) |

## SPECIALISATION

| TOPIC | TOTAL LECTURES | TOTAL DURATION |
|---|---|---|
| Data Science & Analytics | 43 | 14 Weeks (3.5 Months) |
| (OR) | | |
| Data Science & Machine Learning | 62 | 21 Weeks (5 Months) |
| Advanced Programming (OPTIONAL) | 35 | 12 Weeks (3 Months) |

Placement Assistance module begins a month prior to the module.

The highlight of Scaler Data Science is the placement assistance by our Careers Team. We'll help you with job opportunities & you can grasp the know-how of DSA Advanced alongside.

### DSA PROGRAMMING

How do you think great engineers are 100x more efficient than mediocre ones? No, they don't solve faster but solve problems in the simplest way. Being well-versed with DSA amps up your problem-solving skills, which is essential to ace interviews - for both Data Science & Machine Learning roles.

Want to take a deeper dive into our curriculum?

**Check it out on Page 15 - 24.**

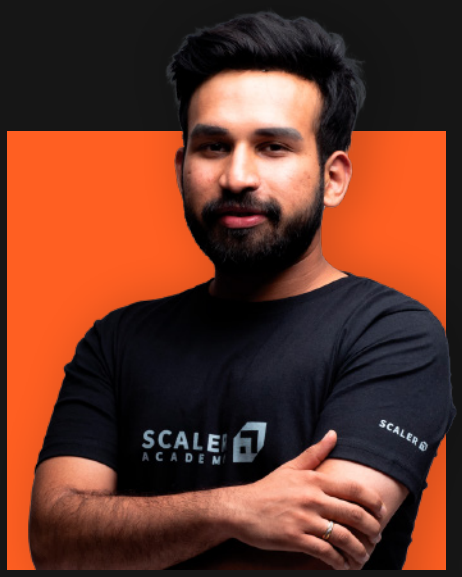
# LEARN FROM EXPERTS
# WHO'VE BEEN THERE, DONE THAT!

Wear your learner's hat as our Industry Experts walk you through every concept with a fresh perspective. Have a look at our teaching army who'll impart industry wisdom so that you gain real-world exposure.


## Mudit Goel

Ex - LinkedIn, INTUIT

At LinkedIn and Intuit, Mudit was granted patents by the US Government. He led the Data Science team at D2L (ranked among the most innovative companies in Data Science). Mudit founded Coding Elements, which was selected by the Govt. of India to teach coding to 2 Million students. He currently leads the Data Science and ML program at Scaler.




## Anant Mittal

UNIVERSITY OF MARYLAND · INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY HYDERABAD

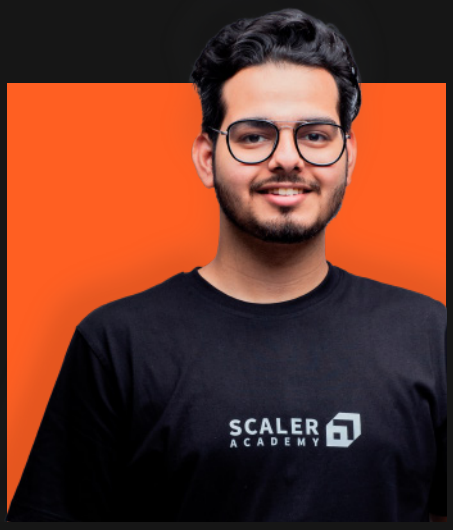
As a researcher at the University of Maryland, he worked on cutting-edge systems to find biomarkers of task activities in the brain. He designed and developed COVID-19 & hygiene-related analytics solutions such as temperature screening and violations related to preventive measures.

# MEET YOUR INSTRUCTORS


## Srikanth Varma

Ex - Lead Machine Learning Scientist, amazon

Srikanth enjoys teaching and loves solving problems that matter, by building products and services from the ground up. He is a lifelong learner, tinkerer and a team builder, who has worked with Amazon, Yahoo, and co-founded Applied Roots.




## Mohit Uniyal

Co-founder, CODING MINUTES, Mentor - Google Code-in, TensorFlow

Mohit is the co-founder of 2 ed-tech startups dedicated to competitive coding preparation. He began his mentoring journey as a Mentor@TensorFlow at Google Code-in. An ex-instructor and Product Engineer at Coding Blocks, he has extensive teaching experience. Currently, he is a Data Scientist at Coding Minutes.

## Harshit Tyagi

Ex-Data Engineer, elucidata, Instructor - in Learning

Harshit is a Lead Instructor (Data Science & Machine Learning) at Scaler Data Science. He is the creator of bestselling Data Science & Engineering courses at LinkedIn Learning, OpenClassrooms, Manning, and O'Reilly Media. An ex-Data Engineer at Elucidata, he created biomedical apps for Research Scientists at Yale, UCLA, and MIT.

# MEET YOUR INSTRUCTORS

## Suraaj Hasija

Ex Sr. Data Scientist , MasterCard

He leads Data Science Business Cases at Scaler Data Science. Working as a Data Scientist 2 at MasterCard, he has transformed industry insights into actionable success for the business. Responsible for building International Market Pricing Strategy, he leads some of the exemplary large-scale Data Science projects within MasterCard.

## Prateek Narang

Ex-Software Engineer, Google

An ex-Googler, Prateek is the co-founder of 2 ed-tech startups dedicated to competitive coding preparation. He has completed his MS in Machine Learning from IIT Delhi. Also, he is a popular Udemy Instructor, teaching coding to 75K+ students. Currently, Prateek is an Instructor and the Engineering Lead at Scaler.

## Anshuman Singh

Co-Founder, SCALER ACADEMY, Ex - facebook

He is the co-founder of Scaler Academy and a 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.

# PERSONALISE YOUR EXPERIENCE WITH 1:1 MENTORSHIP

Get clarity on your career path and tackle every stage of your Upskilling journey with 1:1 guidance twice a month. Our mentors will help you with your queries, give interview insights, provide placement assistance and make sure you're on the right track.

**Sahil Chelaramani**
Data Scientist,
Microsoft

He has worked on Bing Search and Azure Global Development teams. He has experience in building large Deep Learning projects and robust Data Science systems.

**Girijesh Prasad**
Senior Manager, Data Science, Morgan Stanley

He has extensive experience in delivering end-to-end Data Science solutions - from infrastructure to models. He can also share his experience about management and business.

**Rajeev Baditha**
Data Scientist,
Walmart

He has theoretical as well as hands-on knowledge of Data Science, and has worked at Walmart and Fractal Analytics. He has a Master's degree from Indian Statistical Institute.

**Hitesh Hinduja**
Senior Manager, Artificial Intelligence, OLA ELECTRIC

He is passionate about cutting-edge research. He also leads a team of 20 to deliver the best electric vehicles, while leveraging end-to-end Machine Learning pipelines.

# ADVISORY BOARD:
## CORNERSTONE OF SCALER DATA SCIENCE

How exciting it'd be for you to work on business problems that come directly from your dream companies? It'll give you a good analysis of the (Data Science) pitch before you go about hitting a home run in your career.

Our Advisory Committee consists of DS professionals who give us a good peek into the industry insights to help us craft about 80 business case studies. Take a look at our industry advisors who help us (and you!) to be better than before.

### Pawan Kumar
Head of Data Science, Uber
Ex - LinkedIn

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.

### 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.

# AMP UP YOUR CAREER

With the right support & career guidance, your Skills can have a real-world impact. We'll push your limits & prepare you to be the player of your (Career) Innings.

- Access job opportunities from **600+ Partner Employers**
- Optimise your resume, LinkedIn Profile & get the needed **Placement assistance**
- Practice mock **interviews** with Industry Ninjas
- Connect & collaborate with **20K+ Scaler** learers and alumni

**Our Alumni are at reputed Tech Companies & promising Startups**

# FEE STRUCTURE

## Total Tuition Fee

(Inclusive of GST)

## INR 2.99L

Our payment plans include multiple easy financing options that make Scaler accessible to everyone. With our affordable solutions, your investment can be as low as **INR 9,811 per month** - it looks like your monthly grocery bill, doesn't it?

We also offer a 2 weeks trial period and a money-back guarantee if you choose to drop out of the course.

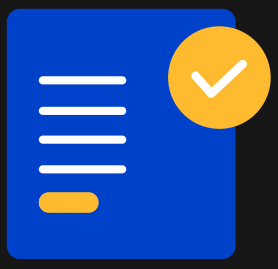
## 3 Easy Steps to Enrol

**Easy Registration**

Make a new account on www.scaler.com or login using social media / InterviewBit account

**Quick Evaluation**

Take a simple 30 min MCQ test focused on aptitude and basic coding to find the right course for you



**Start Upskilling**

Start your learning journey and get set to scale greater professional heights in the Data Science arena

# **CURRICULUM** THAT COMPLEMENTS YOUR CAREER

The Curriculum is crafted to make you a dynamic Data Scientist. We want to prepare you to tackle the toughest challenges a Data Scientist or ML Engineer can face in their journey. It's time you cover-drive the ball (your skills) that it hits the boundary of your dream career!

SCALER ADVANTAGE | Note: Listed below is the Advanced section i.e. the Data Science Core

## 01 Python for Data Science

### TOPICS COVERED

Numpy, Pandas, Data Visualization using Matplotlib and Seaborn, Regular Expressions/Pattern Matching.

SCALER ADVANTAGE

Prior know-how of Python is not a mandate. You'll cover essential tools like Git.
Solve complex business problems using Numpy & Pandas.

# 02 Probability & Statistics

## TOPICS COVERED

**Probability Theory and Descriptive Statistics :**
Combinatorics, Mean, Median, Mode, Percentile, IQR, Outlier, Marginal Probability, Joint Probability, Conditional Probability, Bayes Theorem

**Probability Distributions :**
PDF, PMF, CDF, PPF, Uniform Distribution, Bernoulli Distribution, Multinomial Distribution, Gaussian/Normal Distribution, Poisson Gaussian, Exponential Distribution, Geometric Distribution, Log-normal distribution, Pareto/Power Law Distribution

**Inferential Statistics :**
Confidence Interval Estimates, Central Limit Theorem

SCALER ADVANTAGE

From Emergency Call Centre to Casino of Las Vegas - Experience Probability & Statistics with a fresh perspective.

# 03 Maths Refresher for Data Science

## TOPICS COVERED

Coordinate Geometry, Linear Algebra, Linear Programming, Optimisation Basics, Estimation Problems

SCALER ADVANTAGE

Solidify your fundamentals & fall in love with Mathematics as you solve engaging problems - from Drone Delivery to Soccer Matches

# 04 Data Acquisition & Unstructured Data

## TOPICS COVERED

**Databases and SQL :** Relational, Non-relational, Entity-Relationship diagrams, SQL Commands, Aggregate Functions, Joins, SubQueries, Normalisation, Scaling patterns, ACID, Dask SQL, Cloud SQL (Athena/BigQuery), Web API, Data Cleaning.

**Unstructured Data :** OpenCV, Image processing, Smoothening, Morphological Operations, NLTK, Text processing.

**SCALER ADVANTAGE**

Learn how to extract data from various sources (cloud or local) & work with data in different formats (Tables, Files, Images, Audio, Video, Text)
Build a crawler to scrape websites & deploy web apps

# 05 Applied Data Science

## TOPICS COVERED

Hypothesis Testing, Parametric vs non-parametric tests, Z-test, Chi-square test, Skewness, Kurtosis, Normality test, Experiment Design, ANOVA, Simulations, Power of Test, A/B testing, Diff n Diff Analysis, Multi-arm bandit testing, EDA, Multi-dimensional/multi-variate analysis, Covariance, Correlation, Pearson, Spearman Rank, Feature Engineering, Column normalisation, Standardisation, Missing Values, Outlier treatment.

**SCALER ADVANTAGE**

Say Hello to the latest & advanced techniques used by unicorns to conduct experiments & analysis of results.

## 06 Data Visualisation with Tableau

### TOPICS COVERED

Managing Data Sources and Visualizations, Analyzing Data Using Statistical Tools, Creating Basic Charts, Dashboards & Actions

**SCALER ADVANTAGE**

Get the hang of the predominant industry tool- Tableau - for visualising, dashboarding & reporting to ace your role as a Data Analyst, Data Scientist or ML Engineer.

## 07 Product, Strategy, Business

### TOPICS COVERED

Metric Design, Decode Product and Strategy Rounds, Domain Knowledge - Banking, Finance, Marketing, Social Media, Operations, Healthcare, Experiment Design Advanced

**SCALER ADVANTAGE**

Outshine your Problem-Solving as you learn to break down business situations, design correct metrics & deal with uncertainty.

# Specialisation 1
## Data Science & Analytics

TOPICS COVERED

## 01 ANALYTICS

**Advanced Tableau :** Mapping Geographic Data, Using stories to build dashboards, Working with Time and Dates, Creating Conditional Calculations Using Logical Functions, Creating Level of Detail (LOD) Expressions, Summarizing Data Using Table Calculations, Managing Text Strings.

**Advanced SQL :** Nested Queries, String Functions & Pattern Matching, Mathematical Functions, Date-Time Functions.

**Excel :** Basic EXCEL Formulas & Functions, Text, Time and Dates Data Formats, Pivot Table, Statistics in EXCEL, Data Handling in EXCEL, Charts in Excel.

**R- Programming :** Creating objects, Matrices, Loops, Functions in R, Data Manipulation in R, Creating frequency tables and crosstables, Building charts, Performing univariate analysis, Working with dates and time in R, Dplyr, GGplot.

**Power BI :** Data & Relationship View, Connecting Different Data sources, Data Modelling, Time Series, Aggregation and Filters, Maps, Plotting Techniques, Tables and Relationships, Creating an Interactive Business Intelligence Report.

## 02 MACHINE LEARNING

**Supervised Machine Learning :** KNN, Linear Regression, Logistic Regression, Decision Trees, Feature importance, ML Metrics.

**Unsupervised Machine Learning :** Dimensionality Reduction & Visualization, Anomaly Detection, K-Means, PCA, t-SNE.

**Miscellaneous Machine Learning Topics :** Text and Image vectorization using Deep-Learning, Interpretable ML, ML Life Cycle, NLP Concepts.

**Time Series Forecasting :** Resampling, Autocorrelation, Forecasting, Seasonal Naive, Double/Triple Exponential (Holt) Residual Analysis, Stationarity tests, Autoregressive methods, moving averages, ARIMA, SARIMA.

## 03 BUSINESS AWARENESS

Design of Survey, Metric Design, Big Data Frameworks, Business Case-studies Risk, Product, Banking, Finance, E-Commerce, Social Media, Marketing, Transporation, Healthcare, Operations.

**SCALER ADVANTAGE**

From Python & R to Tableau & Power BI - you'll get to grips with everything that is industry-relevant.
Become a power-user of tools like MS Excel & master advanced SQL.
Dive deeper into complex business cases & gain insights into domain-specific techniques.

# Specialisation 2
Data Science & Machine Learning Ops

### TOPICS COVERED

## 01 ESSENTIAL MATHS FOR MACHINE LEARNING

**Linear Algebra :** Mapping Geographic Data, Using stories to build dashboards, Working with Time and Dates, Creating Conditional Calculations Using Logical Functions, Creating Level of Detail (LOD) Expressions, Summarizing Data Using Table Calculations, Managing Text Strings.

**Coordinate Geometry :** Line, Plane, HyperPlane, Half space, Classification using plane.

**Calculus :** Functions, Limits, Derivatives, Partial derivatives, Saddle points.

## 02 SUPERVISED LEARNING

Linear Regression, Gradient Descent, Multicollinearity, VIF, R-square, Heteroscedasticity, Polynomial Regression, Bias-Variance trade-off, Regularisation, Logistic Regression, Squashing function, AUC, ROC, Precision-Recall Curve, Confusion matrix, Specificity, KNN, Decision Trees, Ensemble learning, Bagging, Boosting, SHAP Values, Support Vector Machine, Bayesian Machine Learning, Sklearn.

## 03 UNSUPERVISED LEARNING

KMeans, Customer Segmentation, Hierarchical clustering, DBSCAN, Anomaly Detection, Local Outlier Factor, Isolation Forest, Dimensionality Reduction, PCA, t-SNE, GMM, Information Theory, Expectation Maximisation.

## 04 RECOMMENDER SYSTEMS

Collaborative/Content filtering, Propensity analysis, Cold start problem

## 05 PREDICTIVE MODELING & TIME SERIES FORECASTING

EDA, Resampling, Autocorrelation, Forecasting, Seasonal Naive, Double/Triple Exponential (Holt) Residual Analysis, Stationarity tests, Autoregressive methods, moving averages, ARIMA, SARIMA, Garch.

## 06 DEEP LEARNING

**Neural Networks :** MLP, Backpropagation, Hyperparameter Tuning, Practical Aspects of DL, Keras, Tensorflow, Pytorch.

**Computer Vision :** Convolutional Neural Nets, Data Augmentation, Transfer Learning, CNN Visualisation, Popular CNN Architecture - Alexnet, VGG, ResNet, Inception, DenseNet, EfficientNet, MobileNet, Object Segmentation, Localisation and Detection, Generative Models, VAEs, GANs, Attention Models, Siamese Networks, Advanced CV.

**Natural Language Processing :** Text Processing and representation - Tokenization, Stemming, Lemmatization, Vector space modelling, Cosine Similarity, Euclidean Distance, POS tagging, Dependency parsing, Topic Modeling, Language Modeling Embeddings, Recurrent Neural Nets, Information Extraction, Entity Recognition, Transformers, BERT, Building Chatbots.

**Reinforcement Learning and Forecasting :** Reinforcement Learning, Q-learning, Autonomous players, RNNs and LSTMs for forecasting.

## 07 ML OPS AND DATA ENGINEERING

**ML Ops :** Project scoping, Experiment tracking using MLFlow/W&B, API (Flask/FastAPI), Streamlit, Testing, Versioning, Docker, CI/CD pipelines, AWS lambda, Monitoring using AWS Kibana, Drift

**ETL Pipeline**

**Big Data and Distributed Systems :** Hadoop, Map-reduce, Spark, PySpark, SparkSQL, Spark ML, Orchestration using Airflow, Serving using Flask.

**Data Governance :** Data Quality, Data Dictionary, Data Access Management, Data storage and recovery procedure, Handling stereotypes and Social biases in data.

**Cloud Computing :** Need for Cloud, What is Cloud Computing,ML on cloud, Key cloud services providers AWS, Microsoft Azure and Google Cloud Platform(GCP),Key Cloud Services, Challenges in Cloud Adoption, Sagemaker on AWS, Vertex AI on GCP,Data Science studio on Azure, Case Studies.

SCALER
**ADVANTAGE**

Get hands-on with real-world's unclean data as you work on projects built in partnerships with top companies.

# Thank you

SCALER

WWW.SCALER.COM

#CreateImpact