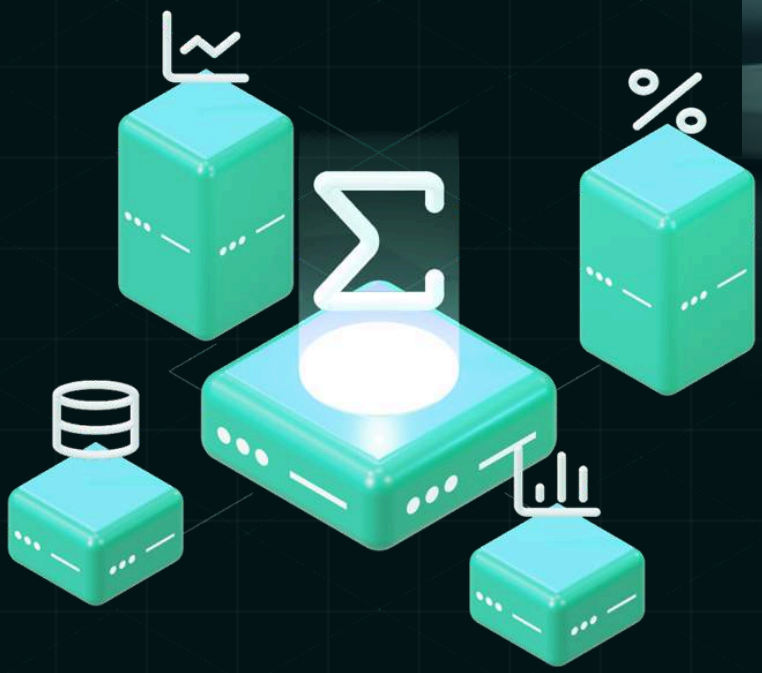
samcommunity.in

✦ DATA SCIENCE & ANALYTICS · INTERMEDIATE

Statistics for Data Science

The maths behind good decisions.





40	6	SAM-STATS
HOURS	MODULES	CERTIFICATION

SAMCOMMUNITY ACADEMY

academy@samcommunity.in

samcommunity.in

✦ INSIDE THIS BROCHURE

Contents

Program overview 03

Who it's for 04

Learning outcomes 05

Program runway 06

Detailed curriculum 07

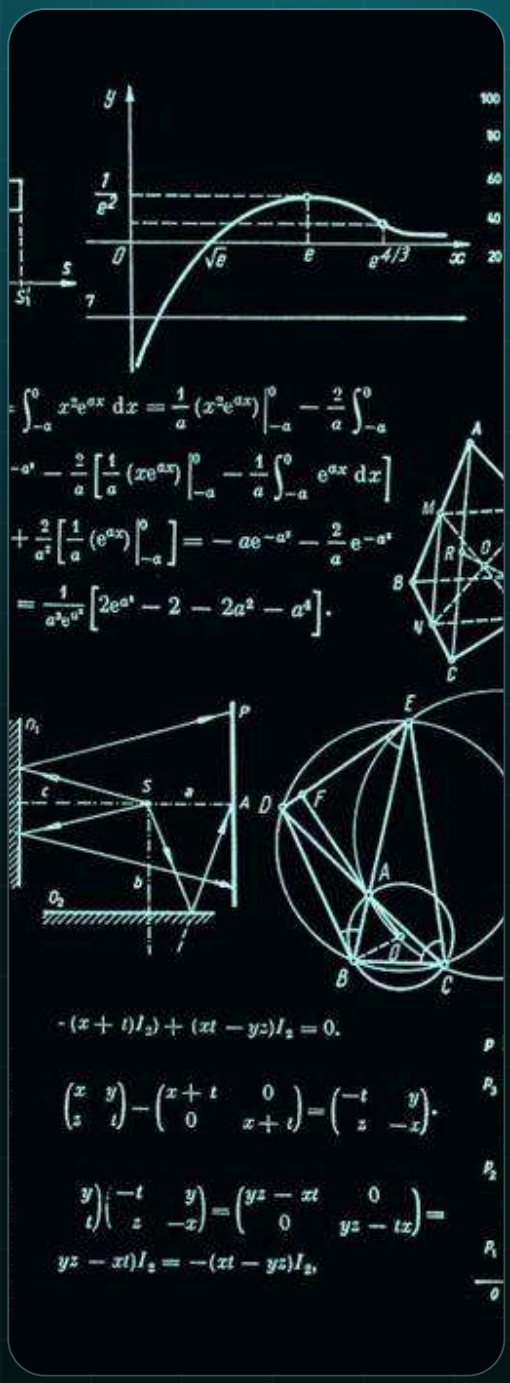
Tools & platforms 09

Projects & capstone 10

Certification & method 11

Learning path 12

FAQs 13



AT A GLANCE

Level	Intermediate
Duration	40 Hours
Modules	6
Certificate	SAM-STATS

✦ PROGRAM OVERVIEW

The maths behind good decisions.



Descriptive and inferential statistics, probability, hypothesis testing and regression for data work.

Good data science rests on sound statistical reasoning: knowing what the data can and cannot support. This intermediate program covers descriptive statistics, probability, distributions, sampling, hypothesis testing, A/B testing and regression, implemented in Python with NumPy, SciPy and statsmodels, so learners can quantify uncertainty and make defensible, evidence-based decisions. Every concept is practised in Python notebooks.

LEVEL	DURATION	CERTIFICATION	DELIVERY
Intermediate	40 Hours	SAM-STATS	Instructor-led · Live virtual · Blended



✦ WHAT YOU'LL PRACTISE

1. Summarise data statistically — 01
2. Apply probability distributions — 02
3. Run hypothesis tests — 03
4. Build and interpret regressions — 04

✦ WHO IT'S FOR

Built for people ready to grow.

01

Business professionals

Data literacy, Excel and SQL foundations

02

Analysts

Power BI, Tableau and Python for analysis

03

Data engineers

Pipelines with Spark, Kafka and Airflow

04

Data scientists & architects

ML, NLP, computer vision and platform design

PREREQUISITES

- Basic Python with NumPy or Pandas 01
- School-level mathematics 02
- Familiarity with data analysis concepts 03

ROLES YOU CAN TARGET

- Data Analyst ↗
- Junior Data Scientist ↗
- Business Analyst ↗
- Product Analyst ↗
- Research Analyst ↗

✦ LEARNING OUTCOMES

What you'll be able to do.

01 Summarise data with descriptive statistics and visual distributions

02 Apply probability rules and common probability distributions

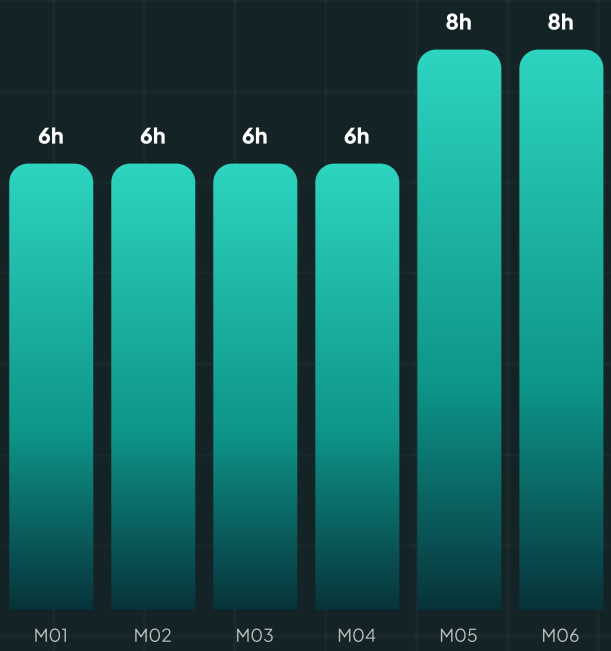
03 Estimate population parameters using confidence intervals

04 Design and interpret hypothesis tests and A/B experiments

05 Build and interpret linear and logistic regression models

06 Communicate statistical findings and uncertainty to stakeholders

HOURS PER MODULE



✦ PROGRAM RUNWAY

Your journey, module by module.

6
MODULES · 40 HRS

01	**Descriptive Statistics and EDA** Measures of central tendency, Variance, standard deviation and IQR, Skewness and outliers	6 hrs
02	**Probability Foundations** Probability rules and events, Conditional probability, Bayes' theorem	6 hrs
03	**Probability Distributions** Binomial and Poisson distributions, Normal distribution and z-scores, Exponential and uniform distributions	6 hrs
04	**Sampling and Estimation** Sampling methods and bias, Sampling distributions, Standard error	6 hrs
05	**Hypothesis Testing and A/B Testing** Null and alternative hypotheses, p-values and significance levels, t-tests, chi-square and ANOVA	8 hrs
06	**Regression Analysis** Simple and multiple linear regression, Assumptions and diagnostics, Interpreting coefficients	8 hrs
★	**Capstone · Statistical Decision Report** Assessment, capstone review and certification	

✦ DETAILED CURRICULUM

The syllabus.

Modules 01–03

MODULE 01

Descriptive Statistics and EDA

6 hours

Hands-on lab
Profile a customer dataset and summarise key distributions.

TOPICS COVERED

- Measures of central tendency
- Variance, standard deviation and IQR
- Skewness and outliers
- Histograms, box plots and density plots
- Correlation and covariance

MODULE 02

Probability Foundations

6 hours

Hands-on lab
Simulate business scenarios to estimate probabilities and expected values.

TOPICS COVERED

- Probability rules and events
- Conditional probability
- Bayes' theorem
- Random variables and expectation
- Simulation with NumPy

NumPy

MODULE 03

Probability Distributions

6 hours

Hands-on lab
Model order arrivals and defect counts using appropriate distributions.

TOPICS COVERED

- Binomial and Poisson distributions
- Normal distribution and z-scores
- Exponential and uniform distributions
- Central limit theorem
- Working with SciPy distributions

LEARN BY DOING

6 guided labs, one in every module.



✦ DETAILED CURRICULUM

Syllabus, continued.

Modules 04–06

MODULE 04

Sampling and Estimation

6 hours

Hands-on lab
Estimate average order value with confidence intervals and bootstrapping.

TOPICS COVERED

- Sampling methods and bias
- Sampling distributions
- Standard error
- Confidence intervals
- Bootstrapping and resampling

MODULE 05

Hypothesis Testing and A/B Testing

8 hours

Hands-on lab
Design and analyse an A/B test on website conversion rates.

TOPICS COVERED

- Null and alternative hypotheses
- p-values and significance levels
- t-tests, chi-square and ANOVA
- Type I and II errors and power
- A/B test design and sample size
- Multiple testing pitfalls

MODULE 06

Regression Analysis

8 hours

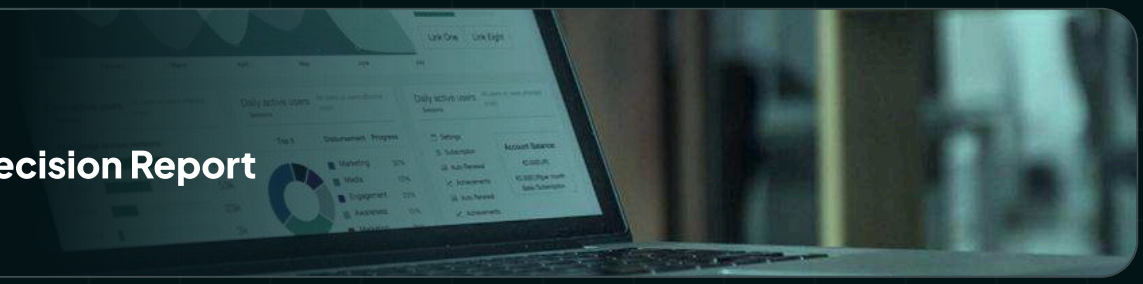
Hands-on lab
Build a regression model explaining sales drivers and interpret the coefficients.

TOPICS COVERED

- Simple and multiple linear regression
- Assumptions and diagnostics
- Interpreting coefficients
- Logistic regression
- Model fit metrics
- Regression with statsmodels

CAPSTONE

Statistical Decision Report



✦ YOUR STACK

Tools you'll build with.

Real platforms and frameworks — the same stack modern teams ship with, practised in cloud-hosted labs.

Python
Automation & scripting

Jupyter
Notebooks & experiments

NumPy
Numerical computing



LAB ACCESS

24×7 cloud labs for the length of the program

✦ HANDS-ON WORK

Projects that prove you can do it.

01

Customer Data Profile

Describe a real dataset with summary statistics, distributions and correlations, highlighting outliers and data quality concerns.

02

A/B Test Analysis

Design an experiment, calculate sample size, run significance tests and recommend a decision with stated uncertainty.

03

Sales Driver Regression

Fit and diagnose a multiple regression model explaining sales, interpreting coefficients for a business audience.

CAPSTONE PROJECT

Statistical Decision Report

Investigate a business question end to end: explore the data, choose appropriate tests or models, quantify uncertainty with confidence intervals, check assumptions and present a clear, decision-focused report with limitations.

Real-world scenario | Mentor review | Portfolio ready



CERTIFICATION

SAM-STATS

Awarded on successful completion of the program assessments.

- MCQ assessments
- Practical lab evaluation
- Project & capstone review
- Final presentation



✦ HOW YOU LEARN

The SAM Learning Framework™

PHASE 01

S

Scenario

Real-world attack and use-case scenarios open every module.

PHASE 02

A

Apply

Hands-on lab work, CTF challenges and live projects.

PHASE 03

M

Master

Assessment, peer review and certification prep.

PHASE 04

M

Mentor

Post-program career guidance and support.

DELIVERY MODELS

- Instructor-Led (ILT)
- Virtual Instructor-Led
- Self-Paced Online
- E-Learning Licensing
- Blended Learning
- Corporate Onsite
- Bootcamp Intensive
- Hackathon / CTF
- Mentorship

academy@samcommunity.in

samcommunity.in

✦ LEARNING PATH · DATA SCIENCE & ANALYTICS

Where this program fits.

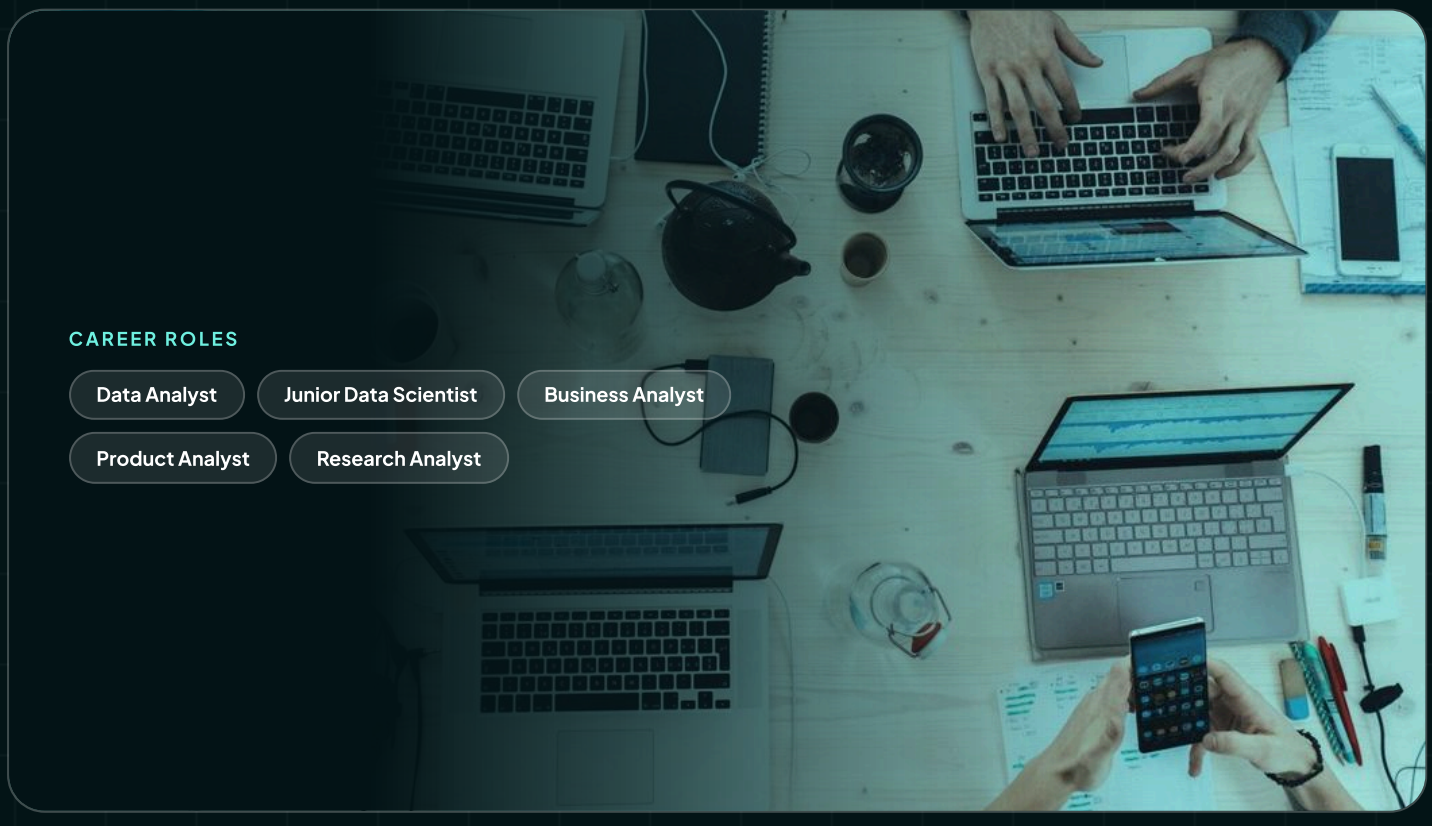
Beginner	Intermediate	Advanced	Expert
3 programs	You are here	4 programs	2 programs

BEFORE THIS

Data Literacy for Business Professionals	16 Hours
Excel for Data Analysis	24 Hours
SQL for Data Analysts	40 Hours

UP NEXT

Advanced Data Engineering (Spark/Kafka/Airflow)	96 Hours
Machine Learning for Data Scientists	80 Hours
NLP for Data Scientists	64 Hours
Computer Vision	64 Hours



CAREER ROLES

- Data Analyst
- Junior Data Scientist
- Business Analyst
- Product Analyst
- Research Analyst

✦ FAQS

Questions, answered.

Q01

Do I need prior experience?

You should know basic Python and school-level mathematics. No advanced maths background is required, as concepts are explained intuitively before formulas, with practical Python examples supporting every topic covered in the program.

Q02

How are classes scheduled and delivered?

Sessions are instructor-led and run as live virtual or blended batches, with weekday and weekend options where available. Sessions mix intuitive explanations, worked examples and notebook exercises, with notes shared for revision.

Q03

Will I receive a certificate?

Yes. Learners who complete the required labs, projects and final assessment receive the Samcommunity Academy SAM-STATS certificate. It recognises the practical skills you demonstrated during the program rather than attendance alone.

Q04

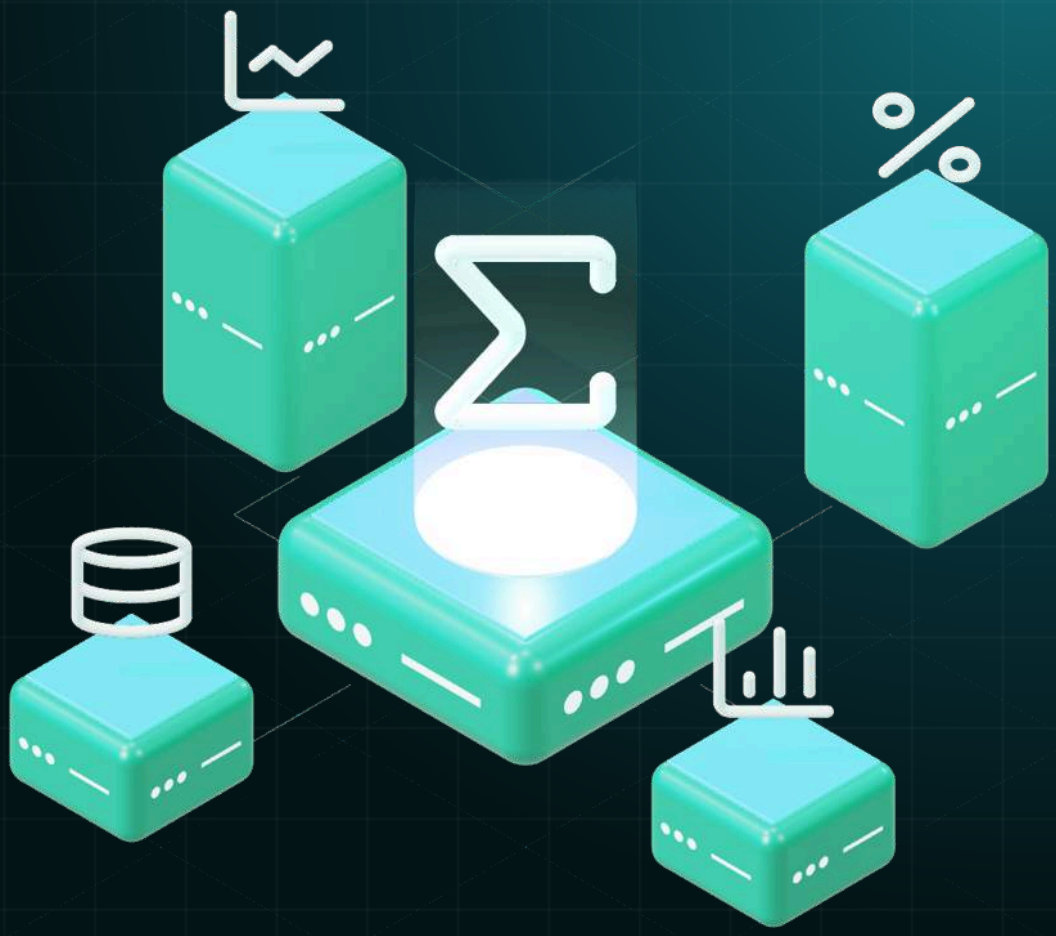
How do I access labs and tools?

Labs use Jupyter notebooks or Google Colab with NumPy, Pandas, SciPy and statsmodels. Sample datasets and starter notebooks are provided, along with installation guidance before the first session.

Q05

Is career support included?

You get guidance on portfolio projects, resume positioning and interview preparation for relevant roles. Statistics projects strengthen portfolios for analyst and data science roles. We do not guarantee jobs, placements or salaries.



✦ LET'S TALK

Start Statistics for Data Science with Samcommunity.

INDIVIDUALS

academy@samcommunity.in

Batches, enrolment & career guidance

CORPORATE & INSTITUTIONS

corporate@samcommunity.in

Team batches & custom programs

Xylyn Samcommunity Private Limited · DPIIT-recognized startup · Surat, Gujarat, India