Learnbay

Data Science Foundation Program

- 450+ Hiring Partners
- Hybrid Model for Project Sessions
- 175% Average Salary Hike

www.learnbay.co

CONTEXT

2cr
worth scholarships awarded

600+
professionals secured jobs after a career break

35k+
Trusted Learners

About The Program

The Data Science Foundation Program is tailored for individuals looking to excel in the dynamic field of data science. This comprehensive program encompasses essential topics such as data analysis and machine learning. Participants gain valuable industry insights and hands-on experience. Our mission is to offer accessible education that empowers individuals to thrive in the evolving world of data science.

We exist to provide accessible, reasonable, and industry-relevant education that empowers India's workforce to grow and develop.

4.79/5

4.66/5

4.8/5

Great Place To Work. Certified

Program Details

COURSE PREREQUISITE

Prior knowledge of **programming/coding** is not mandatory. Just the urge to learn programming and basic ideas about advanced math is enough.

PROGRAM ELIGIBILITY

- **Working professionals** having more than 6 months of experience in any domain (Technical/Non-Technical)

KEY FEATURES

- Dedicated Placement Cell | 100% Guaranteed Interview calls
- Globally Recognised **Certification from IBM & Microsoft**

JOB ROLES TO TARGET

Get equipped with the industry relevant skills and aim for job roles like Financial Data Analyst, Risk Analyst, Insurance Data Analyst, Fraud Detection Analyst etc.

Click below

Check Eligibility

Alumni Spotlight

Shravanthi A

Data Scientist

Learnbay has helped me a lot to learn data science applications in the e-commerce industry. The live class concept was really helpful in receiving proper DS training. Thanks to all my mentors and the placement team.

Mechanical Domain ➡ **Data Scientist @** Ford

230% Salary Hike

Preksha Mishra

Lead Data Scientist

The course structure is excellent with emphasis on concept building and tools & software at the same time. The support team is excellent and supportive and quite agile to respond to doubts.

Telecom Domain ➡ **Data Scientist @** HCL

140% Salary Hike

Karan Chawala

xceedance

⬇

EXL

Data Scientist

Jaya Sinha

ADITYA BIRLA GROUP

⬇

Infosys

Senior Analyst

Shubham Dev

publicis sapient

⬇

brillio

Lead Data Analyst

Alumni Spotlight

Mohd. Israr
Data Scientist

Thanks to the Learnbay data science course & excellent guidance, I was able to ace the TCS interview and secure a job with a 210% pay raise. The real-world time projects helped me develop my concepts as a data scientist.

Mechanical Domain ➡ **Data Scientist @** TCS

210% Salary Hike

Saurabh Kumar
Data Scientist

When I joined Learnbay I did not have any knowledge apart from the very basics. I gradually build my concept via various trainers and get trained in data science with strong knowledge/concepts.

Mathematics Professor ➡ **Data Scientist @** Teleperformance

135% Salary Hike

Aravind

TheMathCompany ⬇ CATERPILLAR

Senior Data Scientist

Ritesh Kumar

unacademy ⬇ Capgemini

Data Scientist

Ramki

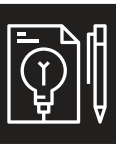
cognizant ⬇ ANBSYSTEMS energizing solutions

Data Analyst

Learnbay's ProjectLab

Choose Learnbay for your career journey because we're more than just a training provider. Our Project Innovation Lab lets you apply your skills in real-world scenarios. Get dual certifications for a competitive edge. Specialize in your desired domain. Discover how Learnbay can boost your career growth. Don't settle for less – choose Learnbay, your path to success!

1 Project Innovation Lab



Work in an industry like environment and gain practical hands-on experience of data scientist with dedicated mentors from industry.

2 Dedicated Placement Cell

Experience 100% job assistance with guaranteed interview call from leading MNCs and startups globally.

3 Degree & Certification

Gain top-notch skills for a successful career through our degree and certification program

1 Project Innovation Lab

Learnbay's Project Innovation Lab replicates industry like environment for real time projects. With our **ProjectLab,** you gain real proof of hands-on experience by having your project certified by the industry.

In our ProjectLab, you work like a data scientist with dedicated project mentors from industry and get certified on capstone project.

450+
Hiring Partners

PUNE

1-1 Doubt Session

HYDERABAD

Capstone Project Certificate from IBM

35k
Trusted Learners

DELHI

BENGALURU

Project Innovation Labs Across India

2 Career Service

Get 1 Year of Job and Placement support

Unleash your career potential with 1 year of unlimited job access, interview support, and profile review.

1 Mock Interview with Industry Experts

Master the art of data science and stay ahead of the curve with mockups and industry insights

Resume Building Session

Craft a powerful resume showcasing your expertise in software development to stand out from the competition

4 Guaranteed Interview Calls

Receive 4 interview calls from a diverse pool of interested employers/recruiters.

3

Certificate

World's leading certifcations

IBM Course Certificate

- Obtain an internationally recognized certificate through training.
- Validate your Data Science skills with IBM Certificate
- Enhance your IT profile with IBM's certification

Others Vs Learnbay

Benefits	Learnbay	Others
Guaranteed Interview Calls	✓	✗
Industry capstone project certificate from IBM	✓	✗
Domain specialized programs for professionals	✓	✗
100% live interactive sessions with industry experts	✓	✗
On-demand video call with industry experts	✓	✗
Personalised Resume Review Session	✓	✗

Program Fee & Financing

Scholarship

Scholarships are awarded based on profile review. Eligible candidates can avail upto 25% scholarship on desired courses. Click the button below to apply.

Click below

Check Scholarship Eligibility

Financing as low as

Rs. 4,917/month

No Cost EMI

Razorpay

LIQUILOANS

PropelId

ShopSe

Program Fee

Rs. 75,000/- +18% GST

Learning Path

L1 **Cohort Orientation + Special Programming Classes**

L2 **Python Programming (Basic + Advance)**

Python, Anaconda, Github, Pandas

L3 **Statistics and Machine Learning**

Matplotlib, Scikit-Learn, Seaborn

L4 **Data Science Tools**

SQL, MongoDB, Tableau, PowerBI, Big Data & Spark Analytics, Time Series

L5 **AI Tools**

Deep Learning, NLP

L6 **Deplyment**

AWS+Azure

L7 **AI Generative Tools and Future Trends**

ChatGPT, Midjourney, DALL·E

Program Syllabus

TERM 1

Python Programming

Module 1 (50 hours)

Programming Basics & Environment Setup

- Installing Anaconda, Anaconda Basics and Introduction
- Get familiar with version control, Git and GitHub.
- Basic Github Commands.
- Introduction to Jupyter Notebook environment. Basics Jupyter notebook Commands.
- Programming language basics

Strings, Decisions & Loop Control

- Working With Numbers, Booleans and Strings, String types and formatting, String operations
- Simple if Statement, if-else Statement if-elif Statement.
- Introduction to while Loops, for Loops, Using continue and break

Class Hands-on:

- 6 programs/coding exercise on string, loop and conditions in classroom

Functions And Modules

- Introduction To Functions
- Defining & Calling Functions
- Functions With Multiple Arguments

Python Programming Overview

- Python Overview
- Python 2.7 vs Python 3
- Writing your First Python Program
- Lines and Indentation, Python Identifiers
- Various Operators and Operators Precedence
- Getting input from User, Comments, Multi line Comments

Python Data Types

- List, Tuples, Dictionaries
- Python Lists, Tuples, Dictionaries Accessing Values, Basic Operations
- Indexing, Slicing, and Matrixes
- Built-in Functions & Methods
- Exercises on List, Tuples And Dictionary

Functions And Modules

- Anonymous Functions - Lambda
- Using Built-In Modules, User-Defined Modules, Module Namespaces,
- Iterators And Generators

Class Hands-on:

8+ Programs to be covered in class of functions, Lambda, modules, Generators and Packages.

Program Syllabus

TERM 1

Python Programming

Module 1 (50 hours)

File I/O An d Exceptional Handling and Regular Expression

- Opening and Closing Files
- open Function, file Object Attributes
- close() Method, Read, write, seek.
- Exception Handling, try-finally Clause
- Raising an Exceptions, User-Defined Exceptions
- Regular Expression- Search and Replace
- Regular Expression Modifiers
- Regular Expression Patterns

Class hands-on :

- 10+ Programs to be covered in class from File IO, Reg-ex and exception handling.

Data Analysis Using Pandas

- Pandas : Introduction to Pandas
- Importing data into Python
- Pandas Data Frames, Indexing Data Frames ,Basic Operations With Data frame, Renaming Columns, Subsetting and filtering a data frame.

Data Analysis Using Numpy

- Introduction to Numpy. Array Creation, Printing Arrays, Basic Operation - Indexing, Slicing and Iterating, Shape Manipulation - Changing shape, stacking and splitting of array
- Vector stacking, Broadcasting with Numpy, Numpy for Statistical Operation

Assignment 1 (Week 2):

10 Coding exercises on Python Basics - Variables, Operators, Strings, Loops, Control Statement

Assignment 2 (Week 3):

10 Python programs and practice set on List, Tuples, Dictionaries & Matrices operations

Assignment 3 (Week 4):

10 Coding exercises on Functions, Lambda, Input-Output, File and Regular Expression

Program Syllabus

TERM 1

Python Programming

Module 1 (50 hours)

Data Visualization using Matplotlib

- **Matplotlib:** Introduction, plot(), Controlling Line Properties, Subplot with Functional Method, Multiple Plot, Working with Multiple Figures, Histograms

Data Visualization using Seaborn

- **Seaborn:** Intro to Seaborn And Visualizing statistical relationships , Import and Prepare data. Plotting with categorical data and Visualizing linear relationships.
- Seaborn Exercise

CASE STUDY

3 Case Study on Numpy, Pandas, Matplotlib

1 Case Study on Pandas And Seaborn

Assessment Test in Python :

- 2 hour of Assesment Test in Python (Coding & Objective Questions)

Real time Use cases in Python to be Covered in Class with 5 assignments

Program Syllabus

TERM 2

Statistics

Module 1 (30 hours)

Fundamentals of Math and Probability

- Probability distributed function & cumulative distribution function. Conditional Probability, Baye's Theorem
- Problem solving for probability assignments
- Random Experiments, Mutually Exclusive Events, Joint Events, Dependent & Independent Events

Introduction to Statistics, Statistical Thinking

- Variable and its types
- Quantitative, Categorical, Discrete, Continuous,
- *all with examples

Five Point Summary and Box Plot

- Outliers, Causes of Outliers, How to treat Outliers, I-QR Method and Z-Score Method

Inferential Statistics

- Central Limit Theorem
- Point estimate and Interval estimate
- Creating confidence interval for population parameter

All about Population & Sample

- Population vs Sample, Sample Size
- Simple Random Sampling, Systematic Sampling, Cluster Sampling, Stratified Sampling, Convenience Sampling, Quota Sampling, Snowball Sampling and Judgement Sampling

Descriptive Statistics

- Measures of Central Tendency – Mean, Median and Mode
- Measures of Dispersion – Standard Deviation, Variance, Range, IQR (Inter-Quartile Range)
- Measure of Symmetricity/ Shape – Skewness and Kurtosis

Inferential Statistics

- Characteristics of Z-distribution and T-Distribution.
- Type of test and rejection region.
- Type of errors in Hypothesis Testing

Program Syllabus

TERM 2

Statistics

Module 1 (30 hours)

Hypothesis Testing

- Type of test and Rejection Region
- Type o errors-Type 1 Errors, Type 2 Errors. P value method, Z score Method. The Chi-Square Test of Independence.
- Regression. Factorial Analysis of Variance. Pearson Correlation Coefficients in Depth. Statistical Significance
- Null and Alternative Hypothesis One-tailed and Two-tailed Tests, Critical Value, Rejection region, Inference based on Critical Value
- **Binomial Distribution:** Assumptions of Binomial Distribution, Normal Distribution, Properties of Normal Distribution, Z table, Empirical Rule of Normal Distribution & Central Limit Theorem and its Applications

Linear Algebra

- Dot Product, Projecting Point on Axis.
- Matrices in Python, Element Indexing, Square Matrix, Triangular Matrix, Diagonal Matrix, Identity Matrix, Addition of Matrices, Scalar Multiplication, Matrix Multiplication, Matrix Transpose, Determinant, Trace
- T-Test, Analysis of variance (ANOVA), and Analysis of Covariance (ANCOVA) Regression analysis in ANOVA

Class Hands-on:

- Problem solving for C.L.T Problem solving Hypothesis Testing Problem solving for T-test, Z-score test Case study and model run for ANOVA, ANCOVA

Data Processing & Exploratory Data Analysis

- What is Data Wrangling
- Data Pre-processing and cleaning?
- How to Restructure the data?
- What is Data Integration and Transformation

Program Syllabus

TERM 2

Statistics

Module 1 (30 hours)

EDA

- Finding and Dealing with Missing Values.
- What are Outliers?
- Using Z-scores to Find Outliers.
- Bivariate Analysis, Scatter Plots and Heatmaps.
- Introduction to Multivariate Analysis

Note: Problem-Solving Techniques and Case Studies using Statistics will be covered in class from week 2

Statistics Assignments : Total 4 practice set and Assignments from Statistics

Program Syllabus

TERM 2

Machine Learning

Module 2 (40 hours)

Machine Learning Introduction

- Definition, Examples, Importance of Machine Learning
- Definition of ML Elements: Algorithm, Model, Predictor Variable, Response Variable, Training - Test Split, Steps in Machine Learning,
- ML Models Type: Supervised Learning, Unsupervised Learning and Reinforcement Learning

Data Preprocessing

- Types of Missing values (MCAR, MAR, MNAR) , Methods to handle missing values
- Outliers, Methods to handle outliers: IQR Method, Z Method
- Feature Scaling: Definition , Methods: Absolute Maximum Scaling, Min-Max Scaler , Normalization, Standardization, Robust Scaling

Data Preprocessing

- Encoding the data: Definition, Methods: OneHot Encoding, Mean Encoding, Label Encoding, Target Guided Ordinal Encoding

Logistic Regression Model

- Definition. Why is it called the “Regression model”?
- Sigmoid Function, Transformation & Graph of Sigmoid Function

Evaluation Metrics for Classification model

- Confusion Matrix, Accuracy, Misclassification, TPR, FPR, TNR, Precision, Recall, F1 Score, ROC Curve, and AUC. Using Python library Sklearn to create the Logistic Regression Model and evaluate the model created

K Nearest Neighbours Model

- Definition, Steps in KNN Model, Types of Distance: Manhattan Distance, Euclidean Distance, ‘Lazy Learner Model’.
- Confusion Matrix of Multi Class Classification
- Using Python library Sklearn to create the K Nearest Neighbours Model and evaluate the model

Program Syllabus

TERM 2

Machine Learning

Module 2 (40 hours)

Decision Tree Model

- Definition, Basic Terminologies, Tree Splitting Constraints, Splitting Algorithms:
- CART, C4.5, ID3, CHAID
- Splitting Methods:
- GINI, Entropy, Chi-Square, and Reduction in Variance
- Using Python library Sklearn to create the Decision Tree Model and evaluate the model created

Hyperparameter Tuning

- GridSearchCV, Variable Importance.
- Using Python library Sklearn to create the Random Forest Model and evaluate the model created.
- Use cases

Random Forest Model

- Ensemble Techniques: Bagging/bootstrapping & Boosting.
- Definition of Random Forest, OOB Score
- K-Fold Cross-Validation

Naive Baye's Model

- Definition, Advantages, Baye's Theorem Applicability, Disadvantages of Naive Baye's Model, Laplace's Correction, Types of Classifiers: Gaussian, Multinomial and Bernoulli
- Using Python library Sklearn to create the Naive Baye's Model and evaluate the model created

CASE STUDY

- **Business Case Study for Kart Model**
- **Business Case Study for Random Forest**
- **Business Case Study for SVM**
- **To classify an email as spam or not spam using logistic Regression.**
- **Application of Linear Regression for Housing Price Prediction**

Program Syllabus

TERM 2

Machine Learning

Module 2 (40 hours)

K Means and Hierarchical Clustering

- Definition of Clustering, Use cases of Clustering
- K Means Clustering Algorithm, Assumptions of K Means Clustering
- Sum of Squares Curve or Elbow Curve

Hierarchical Clustering

- Dendrogram, Agglomerative Clustering, Divisive Clustering, Comparison of K Means Clustering and Hierarchical Clustering
- Using Python library Sklearn to create and evaluate the clustering model

Principal Component Analysis(PCA)

- Definition, Curse of Dimensionality, Dimensionality Reduction Technique, When to use PCA,
- Use Cases
- Steps in PCA, EigenValues and EigenVectors, Scree Plot.
- Using Python library Sklearn to create Principal Components

Support Vector Machine(SVM)

- Model: Definition, Use Cases, Kernel Function, Aim of Support Vectors, Hyperplane, Gamma Value, Regularization Parameter
- Using Python library Sklearn to create and evaluate the SVM Model

Summary of all Machine Learning Models and Discussion about the Capstone Project

Note : All Machine Learning Algorithms are covered in depth with real time case studies for each algorithm. Once 60% of ML is completed, Capstone Project will be released for the batch.

Program Syllabus

TERM 2

CASE STUDY

Module 2 (40 hours)

- Recommendation Engine for e-commerce/retail chain
- Twitter data analysis using NLP

Program Syllabus

TERM 3

SQL

Module 1 (14 hours)

SQL and RDBMS

- RDBMS And SQL Operations.
- Single Table Queries - SELECT, WHERE,
- ORDER BY, Distinct, And, OR
- Multiple Table Queries: INNER, SELF,
- CROSS, and OUTER, Join, Left Join, Right
- Join, Full Join, Union

NoSQL, HBase & MongoDB

- NoSQL Databases
- Introduction to HBase
- HBase Architecture, HBase
- Components, Storage Model of HBase
- HBase vs RDBMS
- Introduction to Mongo DB, CRUD
- Advantages of MongoDB over RDBMS

Programming with SQL

- Mathematical Functions
- Variables
- Conditional Logic
- Loops
- Custom Functions
- Grouping and Ordering

Advance SQL

- Advance SQL Operations
- Data Aggregations and summarizing the data
- Ranking Functions: Top-N Analysis
- Advanced SQL Queries for Analytics

JSON Data & CRUD

- Basics and CRUD Operation
- Databases, Collection & Documents
- Shell & MongoDB drivers
- What is JSON Data
- Create, Read, Update, Delete
- Finding, Deleting, Updating, Inserting Elements
- Working with Arrays
- Understanding Schemas and Relations

Programming with SQL

- Partitioning
- Filtering Data
- Subqueries

Program Syllabus

TERM 3

SQL

Module 1 (14 hours)

Assignments

- Working with multiple tables
- Practice Joins, Grouping and Subqueries
- Using GROUP BY and HAVING Clauses
- Practice Aggregation Queries

Program Syllabus

TERM 3

Tableau

Module 3 (14 hours)

Introduction to Tableau

- Connecting to data source
- Creating dashboard pages
- How to create calculated columns
- Different charts

Dashboard and Stories

- Working in Views with Dashboards and Stories
- Working with Sheets
- Fitting Sheets
- Legends and Quick Filters
- Tiled and Floating Layouts, Floating Objects

Hands-on Assignments

- Connecting data source and data cleansing
- Working with various charts
- Deployment of Predictive model in visualization

Visual Analytics

- Getting Started With Visual Analytics
- Sorting and grouping
- Working with sets, set action
- Filters: Ways to filter, Interactive Filters
- Forecasting and Clustering

Tableau (Advance)

- Mapping
- Coordinate points
- Plotting Latitude and Longitude
- Custom Geocoding
- Polygon Maps
- WMS and Background Image

Tools covered

Program Syllabus

TERM 3

Power BI

Module 4 (14 hours)

Getting Started With Power BI

- Installing Power BI Desktop and Connecting to Data
- Overview of the Workflow in Power BI Desktop
- Introducing the Different Views of the Data Mode
- Query Editor Interface
- Working on Data Model

Programming with Power BI

- Working with Time Series
- Understanding aggregation and granularity
- Filters and Slicers in Power BI Maps
- Scatterplots and BI Reports
- Connecting Dataset with Power BI Creating a Customer Segmentation Dashboard Analyzing the Customer Segmentation Dashboard

Assignments

- Create Bar charts
- Create Pie charts
- Create Tree maps
- Create Donut Charts
- Create Waterfall Diagrams
- Creating Table Calculations for Gender

Tools covered

Power BI

The **IBM** exam will be conducted for all the modules after completion of the course

Real-time Projects

J.P.Morgan

12 hours

Learn and develop classification techniques for the digital transformation of banking

JPMorgan offers tax-friendly insurance choices. You can help them forecast insurance premiums. Targeted marketing using your Random Forest Algorithm skills can help obtain better premium values.

python pandas scikit-learn

NETFLIX

17 hours

Building a content recommendation model on the basis of regional viewer categorization

Netflix is a global entertainment video streaming site. They offer content in various regional languages. Build a local recommendation engine for Netflix customers residing in south Bangalore on their weekend and weekday activities, utilizing NLP.

NumPy SQL jupyter

OLA

18 hours

Reduction of waiting time via a highly precise forecasting model

Make a demand forecasting model based on specific time period rider demands. Such a model will help both riders and cab drivers to ensure the least possible waiting time. You can include measures like latitude and longitude identification.

python TensorFlow APACHE Spark

أرامكو السعودية saudi aramco

14 hours

Understanding in-depth about logging while drilling (LWD) technique

Saudi Aramco company is working on the development of high-efficiency drilling models. Use the bright sides of big data analytics to identify the most cost- effective and highly productive drilling sites.

hadoop matplotlib aws

Contact Us

Click here to whatsapp

or call us at

+91 77956 87988

www.learnbay.co