



Data Science Professional Training Curriculum

STRUCTURE



Data Science Professional Training Curriculum

“Become a Professional in Data Science by joining our comprehensive Training Program at Croma Campus”

About Croma Campus:

Croma Campus Training & Development Private Limited is an education platform since 2010 providing rigorous industry-relevant programs designed and delivered in collaboration with world-class faculty and industry.

- Hands-On Live Projects
- Simulation Test Papers
- Industry Cases Studies
- 61,640+ Satisfied Learners
- 140+ Training Courses
- 100% Certification Passing Rate
- Live Instructor Classroom / Online Training
- 100% Placement Assistance

Course Objectives:

- The objective of this Data Science Professional Course is to deliver all essential skills that can make you future-ready.
- Prepare you for the related data science certification and clear the exam in the first attempt mostly.
- Analyze all types of data either it is structured, unstructured, or complex data types, you could always drive meaningful from it.
- Gain all data science skills from an initial level to the complex stage where you will find yourself more competent and confident as a data science expert.
- Learn the right set of tools, techniques, visualization methods, programming languages, AI, ML, algorithm designing, and other similar skills that will make your resume outshine the crowd.
- In the end, gain all soft skills, intellectual skills, and communication skills to apply for various jobs and get hired quickly.

Course Description:

Data Science is a powerful analytics platform to make discoveries. By using different aspects of computer science, data visualisations, data analytics, statistics, R and Python Programming in data science, you may convert voluminous data into meaningful contents.

Our job-oriented Data Science Training Program at Croma Campus covers all necessary data science aspects in-depth and makes you industry-ready to implement those aspects practically. Also, our data science expert team will guide you to earn certifications and get hired by leading industries worldwide.

Today, there is a dire need of data scientists, data analysts across the world. Data science experts are like gladiators who has the right experience to protect their companies from zillions of things. In other words, data scientists are the walking databases having all vital business insights that are needed by Companies from time to time.

According to latest study by McKinsey Global Institute, there is deficiency of 2 lac data science experts with profound logical abilities. And the demand of skilled experts will continue to increase for next many years to come. People who are currently working in the data science domain are sure to see the certain increment in their current salary packages.

It does not matter where you currently stand in your IT career, it is always good moving ahead with a thoughtful career plan in your mind. So, hurry up and catch the data science express before it is too late with our comprehensive data science training program.

Croma Campus Training Program Deliverables:

- **Session Recordings** - Original Class Room Voice & Video Recording
- **Training Material** - Soft Copy Handbooks
- **Assignments** | Multiple Hands-on Exercises
- **Test Papers** - We provide **Practice Test** as part of our course to help you prepare for the actual certification exam.
- **Live Case Studies**
- **Live Projects** - Hands-on exercises and Project work. You will work on real time industry-oriented projects and assignments for each module to practice.
- **Key focus on Hands-on exercises and Project work.** You will work on real time industry-oriented projects.
- Faculty with more than **10+ Years of Experience** in the Industry.
- **Technical Resume Designing & Job Assistance:** With more than 100+ Clients across the Globe and we help learners to get a good job in their respective field. We also help learners with resume preparation.
- **Interview Q&A**
- **About Croma Campus Training Certificate:** Croma Campus will provide you with an industry-recognized (Certified by **ISO 9001:2015 & E-Cell IIT Jodhpur**) course completion certificate which has lifelong validity.
- **How I Unlock my Croma Campus Certificate:** Attend Complete Batch & Submit at least One Completed Project.

Program Modules:

- **Python Statistics for Data Science**
- **Database – MySQL and SQL Queries**
- **Data Science Professional**
- **Machine Learning**
- **Deep Learning**
- **Power BI Training for Data Visualization**
- **Data Science Master Live Projects**

Data Science Professional Course Content:

Module 1: Python Statistics for Data Science

- **Introduction To Python:**
 - Installation and Working with Python
 - Understanding Python variables
 - Python basic Operators
 - Understanding the Python blocks.
- **Introduction To Variables:**
 - Variables, expression condition and function
 - Global and Local Variables in Python
 - Packing and Unpacking Arguments
 - Type Casting in Python
 - Byte objects vs. string in Python
 - Variable Scope
- **Python Data Type:**
 - Declaring and using Numeric data types
 - Using string data type and string operations
 - Understanding Non-numeric data types
 - Understanding the concept of Casting and Boolean.
 - Strings
 - List
 - Tuples
 - Dictionary
 - Sets
- **Introduction Keywords and Identifiers and Operators**
 - Python Keyword and Identifiers
 - Python Comments, Multiline Comments.
 - Python Indentation
 - Understating the concepts of Operators

Data Structure

- **List**
 - What is List.
 - List Creation
 - List Length
 - List Append
 - List Insert
 - List Remove
 - List Append & Extend using “+” and Keyword
 - List Delete
 - List related Keyword in Python
 - List Revers
 - List Sorting

- List having Multiple Reference
- String Split to create a List
- List Indexing
- List Slicing
- List count and Looping
- List Comprehension and Nested Comprehension
- **Dictionary**
 - Dict Creation
 - Dict Access (Accessing Dict Values)
 - Dict Get Method
 - Dict Add or Modify Elements
 - Dict Copy
 - Dict From Keys.
 - Dict Items
 - Dict Keys (Updating, Removing and Iterating)
 - Dict Values
 - Dict Comprehension
 - Default Dictionaries
 - Ordered Dictionaries
 - Looping Dictionaries
 - Dict useful methods (Pop, Pop Item, Str , Update etc.)

Sets, Tuples and Looping Programming

- **Sets**
 - What is Set
 - Set Creation
 - Add element to a Set
 - Remove elements from a Set
 - PythonSet Operations
 - Frozen Sets
- **Tuple**
 - What is Tuple
 - Tuple Creation
 - Accessing Elements in Tuple
 - Changinga Tuple
 - TupleDeletion
 - Tuple Count
 - Tuple Index
 - TupleMembership
 - TupleBuilt in Function (Length, Sort)
- **Control Flow**
 - Loops
 - Loops and Control Statements (Continue, Break and Pass).
 - Looping techniques in Python
 - How to use Range function in Loop?
 - Programs for printing Patterns in Python

- How to use if and else with Loop
- Use of Switch Function in Loop
- Elegant way of Python Iteration
- Generator in Python
- How to use nested IF and Else in Python
- How to use nested Loop in Python
- Use If and Else in for and While Loop
- Examples of Looping with Break and Continue Statements
- How to use IN or NOTkeywordin Python Loop.

Exception and File Handling, Module, Function and Packages

- **Python Exception Handling**
 - Python Errors and Built-in-Exceptions
 - Exception handing Try, Except and Finally
 - Catching Exceptions in Python
 - Catching Specific Exception in Python
 - Raising Exception
 - Try and Finally
- **Python File Handling**
 - Opening a File
 - Python File Modes
 - Closing File
 - Writing to a File
 - Reading from a File
 - Renaming and Deleting Files in Python
 - Python Directory and File Management
 - List Directories and Files
 - Making New Directory
 - Changing Directory
- **Python Function, Modules and Packages**
 - Python Syntax
 - Function Call
 - Return Statement
 - Write an Empty Function in Python –pass statement.
 - Lamda/ Anonymous Function
 - *argsand **kwargs
 - Help function in Python
 - Scope and Life Time of Variable in Python Function
 - Nested Loop in Python Function
 - Recursive Function and Its Advantage and Disadvantage
 - Organizing python codes using functions
 - Organizing python projects into modules
 - Importing own module as well as external modules
 - Understanding Packages
 - Programming using functions, modules & external packages

- Map, Filter and Reduce function with Lambda Function
- More example of Python Function

Data Automation (Excel, SQL, PDF etc)

- **Python Object Oriented Programming—Oops**
 - Concept of Class, Object and Instances
 - Constructor, Class attributes and Destructors
 - Real time use of class in live projects
 - Inheritance, Overlapping and Overloading operators
 - Adding and retrieving dynamic attributes of classes
 - Programming using Oops support
- **Python Database Interaction**
 - SQL Database connection using
 - Creating and searching tables
 - Reading and Storing configuration information on database
 - Programming using database connections
- **Reading an excel**
 - Reading an excel file using Python
 - Writing to an excel sheet using Python
 - Python | Reading an excel file
 - Python | Writing an excel file
 - Adjusting Rows and Column using Python
 - Arithmetic Operation in Excel file.
 - Plotting Pie Charts
 - Plotting Area Charts
 - Plotting Bar or Column Charts using Python.
 - Plotting Doughnut Charts using Python.
 - Consolidation of Excel File using Python
 - Split of Excel File Using Python.
 - Play with Workbook, Sheets and Cells in Excel using Python
 - Creating and Removing Sheets
 - Formatting the Excel File Data
 - More example of Python Function
- **Working with PDF and MS Word using Python**
 - Extracting Text from PDFs
 - Creating PDFs
 - Copy Pages
 - Split PDF
 - Combining pages from many PDFs
 - Rotating PDF's Pages
- **Complete Understanding of OS Module of Python**
 - Check Dirs. (exist or not)
 - How to split path and extension?
 - How to get user profile detail?
 - Get the path of Desktop, Documents, Downloads etc.

- Handle the File System Organization using OS
- How to get any files and folder's details using OS?

Data Analysis & Visualization

• Pandas

- Read data from Excel File using Pandas More Plotting, Date Time Indexing and writing to files
- How to get record specific records Using Pandas Adding & Resetting Columns, Mapping with function
- Using the Excel File class to read multiple sheets More Mapping, Filling Nonvalue's
- Exploring the Data Plotting, Correlations, and Histograms
- Getting statistical information about the data Analysis Concepts, Handle the None Values
- Reading files with no header and skipping records Cumulative Sums and Value Counts, Ranking etc
- Reading a subset of columns Data Maintenance, Adding/Removing Cols and Rows
- Applying formulas on the columns Basic Grouping, Concepts of Aggregate Function
- Complete Understanding of Pivot Table Data Slicing using iLoc and Loc property (Setting Indices)
- Understanding the Properties of Pivot Table in Pandas Advanced Reading CSVs/HTML, Binning, Categorical Data
- Exporting the results to Excel Joins:
- Python | Pandas Data Frame Inner Join
- Understanding the properties of Data Frame Left Join (Left Outer Join)
- Indexing and Selecting Data with Pandas Right Join (Right Outer Join)
- Pandas | Merging, Joining and Concatenating Full Join (Full Outer Join)
- Pandas | Find Missing Data and Fill and Drop NA Appending DataFrame and Data
- Pandas | How to Group Data How to apply Lambda / Function on Data Frame
- Other Very Useful concepts of Pandas in Python Data Time Property in Pandas (More and More)

• NumPy

- Introduction to NumPy: Numerical Python
- Importing NumPy and Its Properties
- NumPy Arrays
- Creating an Array from a CSV
- Operations an Array from a CSV
- Operations with NumPy Arrays
- Two-Dimensional Array
- Selecting Elements from 1-D Array
- Selecting Elements from 2-D Array
- Logical Operation with Arrays
- Indexing NumPy elements using conditionals
- NumPy's Mean and Axis

- NumPy's Mode, Median and Sum Function
- NumPy's Sort Function and More
- **Matplotlib**
 - Bar Chart using Python Matplotlib
 - Column Chart using Python Matplotlib
 - Pie Chart using Python Matplotlib
 - Area Chart using Python Matplotlib
 - Scatter Plot Chart using Python Matplotlib
 - Play with Charts Properties Using Matplotlib
 - Export the Chart as Image
 - Understanding plt. subplots () notation
 - Legend Alignment of Chart using Matplotlib
 - Create Charts as Image
 - Other Useful Properties of Charts.
 - Complete Understanding of Histograms
 - Plotting Different Charts, Labels, and Labels Alignment etc.
- **Introduction to Seaborn**
 - Introduction to Seaborn
 - Making a scatter plot with lists
 - Making a count plot with a list
 - Using Pandas with seaborn
 - Tidy vs Untidy data
 - Making a count plot with a Dataframe
 - Adding a third variable with hue
 - Hue and scatter plots
 - Hue and count plots
- **Visualizing Two Quantitative Variables**
 - Introduction to relational plots and subplots
 - Creating subplots with col and row
 - Customizing scatter plots
 - Changing the size of scatter plot points
 - Changing the style of scatter plot points
 - Introduction to line plots
 - Interpreting line plots
 - Visualizing standard deviation with line plots
 - Plotting subgroups in line plots
- **Visualizing a Categorical and a Quantitative Variable**
 - Current plots and bar plots
 - Count plots
 - Bar plot with percentages
 - Customizing bar plots
 - Box plots
 - Create and interpret a box plot
 - Omitting outliers
 - Adjusting the whiskers
 - Point plots

- Customizing points plots
- Point plot with subgroups
- **Customizing Seaborn Plots**
 - Changing plot style and colour
 - Changing style and palette
 - Changing the scale
 - Using a custom palette
 - Adding titles and labels: Part 1
 - Face Grids vs. Axes Subplots
 - Adding a title to a face Grid object
 - Adding title and labels: Part 2
 - Adding a title and axis labels
 - Rotating x-ticks labels
 - Putting it all together
 - Box plot with subgroups

Bar plot with subgroups and subplots

Module 2: Databases – MySQL and SQL Queries

- **Python - MySQL**
 - Introduction to MySQL
 - What is the MySQLdb?
 - How do I Install MySQLdb?
 - Connecting to the MYSQL
 - Selecting a database
 - Adding data to a table
 - Executing multiple queries
 - Exporting and Importing data tables.
- **SQL Functions**
 - Single Row Functions
 - Character Functions, Number Function, Round, Truncate, Mod, Max, Min, Date
- **General Functions**
 - Count, Average, Sum, Now etc.
- **Joining Tables**
 - Obtaining data from Multiple Tables
 - Types of Joins (Inner Join, Left Join, Right Join & Full Join)
 - Sub-Queries Vs. Joins
- **Operators (Data using Group Function)**
 - Distinct, Order by, Group by, Equal to etc.

- **Database Objects (Constraints & Views)**
 - Not Null
 - Unique
 - Primary Key
 - Foreign Key

- **Structural & Functional Database Testing using TOAD Tool**
 - SQL Basic
 - SQL Introduction
 - SQL Syntax
 - SQL Select
 - SQL Distinct
 - SQL Where
 - SQL And & Or
 - SQL Order By
 - SQL Insert
 - SQL Update
 - SQL Delete
 - SQL Advance
 - SQL Like
 - SQL Wildcards
 - SQL In
 - SQL Between
 - SQL Alias
 - SQL Joins
 - SQL Inner Join
 - SQL Left Join
 - SQL Right Join
 - SQL Full Join
 - SQL Union
 - SQL Functions
 - SQL Avg()
 - SQL Count()
 - SQL First()
 - SQL Last()
 - SQL Max()
 - SQL Min()
 - SQL Sum()
 - SQL Group By

Module 3: Data Science Professional Program

- **Introduction to Data Science**
 - What is Analytics & Data Science?
 - Common Terms in Analytics
 - What is data?
 - Classification of data
 - Relevance in industry and need of the hour
 - Types of problems and business objectives in various industries
 - How leading companies are harnessing the power of analytics?
 - Critical success drivers
 - Overview of analytics tools & their popularity
 - Analytics Methodology & problem-solving framework
 - List of steps in Analytics projects
 - Identify the most appropriate solution design for the given problem statement
 - Project plan for Analytics project & key milestones based on effort estimates
 - Build Resource plan for analytics project
 - Why Python for data science?

- **Accessing/Importing and Exporting Data**
 - Importing Data from various sources (Csv, txt, excel, access etc)
 - Database Input (Connecting to database)
 - Viewing Data objects - sub setting, methods
 - Exporting Data to various formats
 - Important python modules: Pandas

- **Data Manipulation: Cleansing - Munging Using Python Modules**
 - Cleansing Data with Python
 - Filling missing values using lambda function and concept of Skewness.
 - Data Manipulation steps (Sorting, filtering, duplicates, merging, appending, sub setting, derived variables, sampling, Data type conversions, renaming, formatting.
 - Normalizing data
 - Feature Engineering
 - Feature Selection
 - Feature scaling using Standard Scaler/Min-Max scaler/Robust Scaler.
 - Label encoding/one hot encoding

- **Data Analysis: Visualization Using Python**
 - Introduction exploratory data analysis
 - Descriptive statistics, Frequency Tables and summarization
 - Univariate Analysis (Distribution of data & Graphical Analysis)
 - Bivariate Analysis (Cross Tabs, Distributions & Relationships, Graphical Analysis)

- Creating Graphs- Bar/pie/line chart/histogram/ boxplot/ scatter/ density etc.)
- Important Packages for Exploratory Analysis (NumPy Arrays, Matplotlib, seaborn, Pandas etc.)

- **Introduction to Statistics**
 - Descriptive Statistics
 - Sample vs Population Statistics
 - Random variables
 - Probability distribution functions
 - Expected value
 - Normal distribution
 - Gaussian distribution
 - Z-score
 - Central limit theorem
 - Spread and Dispersion
 - Inferential Statistics-Sampling
 - Hypothesis testing
 - Z-stats vs T-stats
 - Type 1 & Type 2 error
 - Confidence Interval
 - ANOVA Test
 - Chi Square Test
 - T-test 1-Tail 2-Tail Test
 - Correlation and Co-variance

- **Introduction to Predictive Modelling**
 - Concept of model in analytics and how it is used?
 - Common terminology used in Analytics & Modelling process
 - Popular Modelling algorithms
 - Types of Business problems - Mapping of Techniques
 - Different Phases of Predictive Modelling

- **Data Exploration for Modelling**
 - Need for structured exploratory data
 - EDA framework for exploring the data and identifying any problems with the data (Data Audit Report)
 - Identify missing data
 - Identify outliers' data
 - Imbalanced Data Techniques

- **Data Pre-Processing & Data Mining**
 - Data Preparation
 - Feature Engineering

- Feature Scaling
- Datasets
- Dimensionality Reduction
- Anomaly Detection
- Parameter Estimation
- Data and Knowledge
- Selected Applications in Data Mining

Module 4: Machine Learning

- **Introduction to Machine Learning**
 - Machine Learning
 - Machine Learning Algorithms
 - Algorithmic models of Learning
 - Applications of Machine Learning
 - Large Scale Machine Learning
 - Computational Learning theory
 - Reinforcement Learning
- **Techniques of Machine Learning**
 - Supervised Learning
 - Unsupervised Learning
 - Semi-supervised and Reinforcement Learning
 - Bias and variance Trade-off
 - Representation Learning
- **Regression**
 - Regression and its Types
 - Logistic Regression
 - Linear Regression
 - Polynomial Regression
- **Classification**
 - Meaning and Types of Classification
 - Nearest Neighbor Classifiers
 - K-nearest Neighbors
 - Probability and Bayes Theorem
 - Support Vector Machines
 - Naive Bayes
 - Decision Tree Classifier
 - Random Forest Classifier

- **Unsupervised Learning: Clustering**
 - About Clustering
 - Clustering Algorithms
 - K-means Clustering
 - Hierarchical Clustering
 - Distribution Clustering

- **Model optimization and Boosting**
 - Ensemble approach
 - K-fold cross validation
 - Grid search cross validation
 - Ada boost and XG Boost

Module 5: Deep Learning

- **Introduction to Deep Learning**
 - What are the Limitations of Machine Learning?
 - What is Deep Learning?
 - Advantage of Deep Learning over Machine learning
 - Reasons to go for Deep Learning
 - Real-Life use cases of Deep Learning

- **Deep Learning Networks**
 - What is Deep Learning Networks?
 - Why Deep Learning Networks?
 - How Deep Learning Works?
 - Feature Extraction
 - Working of Deep Network
 - Training using Backpropagation
 - Variants of Gradient Descent
 - Types of Deep Networks
 - Feed forward neural networks (FNN)
 - Convolutional neural networks (CNN)
 - Recurrent Neural networks (RNN)
 - Generative Adversal Neural Networks (GAN)
 - Restrict Boltzman Machine (RBM)

- **Deep Learning with Keras**
 - Define Keras

- How to compose Models in Keras?
 - Sequential Composition
 - Functional Composition
 - Predefined Neural Network Layers
 - What is Batch Normalization?
 - Saving and Loading a model with Keras
 - Customizing the Training Process
 - Intuitively building networks with Keras
-
- **Convolutional Neural Networks (CNN)**
 - Introduction to Convolutional Neural Networks
 - CNN Applications
 - Architecture of a Convolutional Neural Network
 - Convolution and Pooling layers in a CNN
 - Understanding and Visualizing CNN
 - Transfer Learning and Fine-tuning Convolutional Neural Networks
-
- **Recurrent Neural Network (RNN)**
 - Intro to RNN Model
 - Application use cases of RNN
 - Modelling sequences
 - Training RNNs with Backpropagation
 - Long Short-Term Memory (LSTM)
 - Recursive Neural Tensor Network Theory
 - Recurrent Neural Network Model
 - Time Series Forecasting
-
- **Natural Language Processing**
 - NLP with python
 - Bags of words
 - Stemming
 - Tokenization
 - Lemmatization
 - TF-IDF
 - Sentiment Analysis
-
- **Overview of Tensor Flow**
 - What is Tensor Flow?
 - Tensor Flow code-basics
 - Graph Visualization
 - Constants, Placeholders, Variables
 - Tensor flow Basic Operations

- Linear Regression with Tensor Flow
 - Logistic Regression with Tensor Flow
 - K Nearest Neighbor algorithm with Tensor Flow
 - K-Means classifier with Tensor Flow
 - Random Forest classifier with Tensor Flow
-
- **Neural Networks Using Tensor Flow**
 - Quick recap of Neural Networks
 - Activation Functions, hidden layers, hidden units
 - Illustrate & Training a Perceptron
 - Important Parameters of Perceptron
 - Understand limitations of A Single Layer Perceptron
 - Illustrate Multi-Layer Perceptron
 - Back-propagation – Learning Algorithm
 - Understand Back-propagation – Using Neural Network Example
 - TensorBoard

Module 6: Power BI

- **Introduction to Power BI**
 - Overview of BI concepts
 - Why we need BI?
 - Introduction to SSBI
 - SSBI Tools
 - Why Power BI?
 - What is Power BI?
 - Building Blocks of Power BI
 - Getting started with Power BI Desktop
 - Get Power BI Tools
 - Introduction to Tools and Terminology
 - Dashboard in Minutes
 - Refreshing Power BI Service Data
 - Interacting with your Dashboards
 - Sharing Dashboards and Reports

- **Power BI Desktop**
 - Power BI Desktop
 - Power BI Dashboards
 - Power BI Q & A
 - Extracting data from various sources
 - Workspaces in Power BI
 - Data Transformation
 - Measures and Calculated Columns

- Query Editor

- **Modelling with Power BI**
 - Introduction to Modelling
 - Modelling Data
 - Manage Data Relationship
 - Optimize Data Models
 - Cardinality and Cross Filtering
 - Default Summarization & Sort by
 - Creating Calculated Columns
 - Creating Measures & Quick Measures

- **Data Analysis Expressions (DAX)**
 - What is DAX?
 - Data Types in DAX
 - Calculation Types
 - Syntax, Functions, Context Options
 - DAX Functions
 - Date and Time
 - Time Intelligence
 - Information
 - Logical
 - Mathematical
 - Statistical
 - Text and Aggregate
 - Measures in DAX

- **Publishing and Sharing**
 - Introduction and Sharing Options Overview
 - Publish from Power BI Desktop and Publish to Web
 - Share Dashboard with Power BI Service
 - Workspaces and Apps (Power BI Pro) and Content Packs (Power BI Pro)
 - Print or Save as PDF and Row Level Security (Power BI Pro)
 - Export Data from a Visualization and Publishing for Mobile Apps
 - Export to PowerPoint and Sharing Options Summary

- **Refreshing Datasets**
 - Understanding Data Refresh
 - Personal Gateway (Power BI Pro and 64-bit Windows)
 - Replacing a Dataset and Troubleshooting Refreshing

Module 7: Data Science Professionals - Live Projects

- Managing credit card Risks
- Bank Loan default classification
- YouTube Viewers prediction
- Super store Analytics (E-commerce)
- Buying and selling cars prediction (like OLX process)
- Advanced House price prediction
- Analytics on HR decisions
- Survival of the fittest
- Twitter Analysis
- Flight price prediction