

PART 1 - PYTHON PROGRAMMING LANGUAGE

Module 1: Intro to Python, What is Python and history of Python?, Unique features of Python, Python-2 and Python-3 differences, Install Python and Environment Setup, First Python Program, Python Identifiers, Keywords and Indentation, Comments and document interlude in Python, Command line arguments, Getting User Input, Data Types, What are variables?, Python Core objects & Functions, Number & Maths.

Module 2: List, Ranges & Tuples in Python, Introduction, Lists in Python, More About Lists, Understanding Iterators, Generators, Comprehensions and Lambda Expressions, Introduction, Generators and Yield, Next and Ranges, Understanding and using Ranges, More About Ranges, Ordered Sets with tuples

Module 3: Python Dictionaries and Sets, Introduction to the section, More on Dictionaries, Python Sets, with Examples text files, writing

Module 4: Input and Output in Python, Reading and writing Challenge, Writing Binary Files Manually, Text Files, Appending to Files and Using Pickle to Write Binary Files

Module 5: Python built in function, Python user defined functions, Python packages functions, Defining and calling Function, The anonymous Functions Loops and statement in Python, Python Modules & Packages

Module 6: Python Object Oriented, Creating Classes and Objects, Accessing attributes, Built-In Class Attributes, Destroying Objects

Module 7: Python Exceptions Handling, What is Exception?, Handling an exception, try....except...else, try-finally clause, Argument of an Exception, Python Standard Exceptions, Raising an exceptions, User-Defined Exceptions

Module 8: Python Regular Expressions, What are regular expressions?, The match Function, The search Function, Matching vs searching, Search and Replace, Extended Regular Expressions, Wildcard

Module 9: Python Multithreaded Programming, What is multithreading?, Starting a New Thread, The Threading Module, Synchronizing Threads, Multithreaded Priority Queue, Python Spreadsheet Interfaces, Python XML interfaces

Module 10: Using Databases in Python, Python MySQL Database Access, Install the MySQLdb and other Packages, Create Database Connection, CREATE, INSERT, READ, UPDATE and DELETE Operation, DML and DDL Operation with Databases, Performing Transactions, Handling Database Errors, Web Scraping in Python

Module 11: Python For Data Analysis Numpy : Introduction to numpy, Creating arrays, Using arrays and Scalars, Indexing Arrays, Array Transposition, Universal Array Function, Array Processing, Array Input and Output Pandas : What is pandas?, Where it is used?, Series in pandas, Index objects Reindex, Drop Entry, Selecting Entries, Data Alignment, Rank and Sort, Summary Statics, Missing Data, index Hierarchy, Matplotlib: Python For Data Visualization.

PART 2 - DATABASE PROGRAMMING WITH SQL

Introduction : Data vs. Information, History of Database, Major Transformations in Computing | **Entities and Attributes :** Conceptual and Physical Models, Entities, Instances, Attributes, and Identifiers, Entity Relationship Modeling and ERDs | **Relationship Fundamentals :** Relationship Transferability, Relationship Types, Resolving Many-to-Many Relationships, Understanding CRUD Requirements, Anatomy of a SQL Statement | **SELECT and WHERE :** Columns, Characters, and Rows, Limit Rows Selected, Comparison Operators | **WHERE, ORDER BY, GROUP BY, HAVING and Intro to Functions :** Logical Comparisons and Precedence Rules, Sorting Rows, Introduction to Functions | **Single Row Functions :** Character Functions, Number Functions, Date Functions, Conversions Functions, General Functions | **Joins :** Cross Joins and Natural Joins, Join Clauses, Inner versus Outer Joins | **Data Manipulation Language (DML) :** INSERT Statements, Updating Column Values and Deleting Rows | **Data Definition Language (DDL) :** Creating Tables, Using Data Types, Modifying a Table | **Constraints :** Intro to Constraints; NOT NULL and UNIQUE Constraints, PRIMARY KEY, FOREIGN KEY, and CHECK Constraints, Managing Constraints | **Views :** Creating Views, DML Operations and Views, Managing Views.

PART 3 - STATISTICS FOR DATA SCIENCE

Datatypes and its measures : Random Variables and its applications, Introduction to Probability with examples, Sampling Techniques – Why and How, Measures of Central Tendency- Mean, Median, Mode, Measures of Dispersion- Variance, Standard Deviation, Range, Measures of Skewness & Kurtosis, Normality tests for dataset, Basic Graph Representations- Bar Chart, Histogram, Box Plot, Scatterplot, Probability Distributions | **Continuous Probability Distribution :** Normal Distribution, Standard Normal, Distribution (Z Distribution), F-Distribution, Chi-Square Distribution | **Discrete Probability Distribution :** Binomial Distribution, Poisson Distribution, Building Normal Q-Q Plot and its Interpretation, Central Limit Theorem for sampling variations, Confidence Interval – Computation and analysis, Data Cleansing (Dealing with Missing Data, Outlier Detection), Feature Engineering (Label Encoding, One-Hot Encoding), Data Transformation, including merging, ordering, aggregation, Sampling (Balanced, Stratified, ...), Data Partitioning (Create Training + Validation + Test Data Set), Transformations (Normalization, Standardization, Scaling, Pivoting), Binning (Count-Based, Handling Of Missing Values as its own Group), Data Replacement (Cutting, Splitting, Merging, ...), Weighting And Selection (Attribute Weighting, Automatic Optimization), Imputation (Replacement of Missing Observations with Statistical Algorithms)

TOOLS USAGE

REPL Online, Anaconda (Jupyter Notebook / Spyder), R Studio, PyCharm, Tableau, SubLime Text

LIBRARY / PACKAGE

Datetime, Statsmodels, NumPy, Pandas, Seaborn, Matplotlib, NLTK, Word2Vec, SciPy, SciKitLearn, Pandas, SciPy.stats, Statistics, Database - Oracle/MySQL

**Practical Training
Doubt Clearing
Test & Certification
Placement Calls**

PART 4 - RESEARCH METHODOLOGY FOR DATA SCIENCE

Formulating a hypothesis statement (NULL and ALTERNATE) | Type-I and Type-II Errors, P-Value, Level of Significance | Parametric Tests: One Sample/Two Samples T Test, One Sample Z Test, Paired T Test, One-Way ANOVA, Chi-Squared Test | Non-Parametric Tests: One Sample Sign Test, Mann-Whitney Test, Kruskal-Wallis Test

PART 5 - LINEAR ALGEBRA FOR DATA SCIENCE

Motivation - Why to learn Linear Algebra? | Representation of problems in Linear Algebra : Visualizing the problem: Line, System of linear equations, Planes | Matrix : Terms related to Matrix, Basic operations on Matrix, Representing in Matrix form | Solving the problem : Row Echelon form, Inverse of a Matrix | Eigenvalues and Eigenvectors : Finding Eigenvectors, Use of Eigenvectors in Data Science: PCA algorithm | Singular Value Decomposition of a Matrix

PART 6 - MACHINE LEARNING MODULE

Module 1 - Supervised Machine Learning - Part - I (Regression Analysis) : Correlation Analysis, Correlation Coefficient | Introduction of Regression, Principles of regression | Simple Linear Regression Analysis | Splitting of Dataset into Train, Validation and Test data | Understanding Overfitting (Variance) vs Underfitting (Bias) | Generalization Error and Regularization Techniques | Multiple Linear Regression Model | Model Adequacy Checking | Transformation and Weighting to Correct Model Inadequacies | Diagnostic for Leverage and Influence | Generalized and Weighted Least Squares Estimation | Indicator Variables | Multicollinearity, Heteroskedasticity, Autocorrelation | Polynomial Regression Models | Poisson Regression Models | Variable Selection and Model Building

Module 2 - Supervised Machine Learning - Part - II (Classification Analysis) : Two-Class Classification : Logistic Regression, Neural Network, Decision Tree, Random Forest, Naïve-Bayes, Support Vector Machine(SVM) | Multiclass Classification(MC) : MC-Logistic Regression, MC-Neural Network, MC-Decision Forest, K-Nearest Neighbor(KNN) | Anomaly Detection | Case Studies I, II, III, IV & V

Module 3 - UnSupervised Machine Learning : Partitioning Clustering | Hierarchical Clustering | Clustering Validation and Evaluation with K-Means Clustering : Assessing clustering tendency, Determining the optimal number of clusters, Clustering validation statistics | DBSCAN: Density-based Clustering | Dimensionality Reduction with Principal Component Analysis(PCA) | Association Rule Learning and Recommendation : Apriori Algorithm, Frequent Pattern Growth | • Case Studies I, II & III

PART 7 - MODEL PERFORMANCE ASSESSMENT IN MACHINE LEARNING

Model Performance Assessment | Confusion Matrix | Precision | Recall | F1-Measure | Accuracy | Error Measures | Mean Squared Error(MSE) | Root Mean Squared Error(RMSE) | Hyper Parameter Optimization/Tuning | Grid Search | Randomized Search | Cross-Fold Validation Techniques | Leave One Out | K-fold | Stratified K-fold | Stratified Shuffle Split | Ensemble Methods in Machine Learning | Bagging and Random Forests | Gradient Boosting | Optimized Distributed Gradient Boosting (XGBoost) | Adaptive Boost (AdaBoost) | Voting Classifier

PART 8 - DATA VISUALIZATION TOOLS & TECHNIQUES

Data Visualization with Matplotlib, Seaborn, Tableau | Relational Plots, Relplot, Scatterplot, Lineplot | Categorical Plots, Catplot, Stripplot, Swarmplot, Boxplot, Violinplot, Barplot, Countplot | Distribution Plots, Jointplot, Pairplot, Distplot | Regression Plots, Lmplot, Regplot, Residplot | Matrix Plots, Heatmap, Clustermap | Tableau Products and Usage, Basic Charts on Tableau, Connecting Tableau with Multiple Sheets and Data Sources, Tableau Filters and Visualization Interactivity, Interaction and Grouping Data, Time Series Chart, Maps and Images in Tableau, Advanced Charts in Tableau and Analytical Techniques, Calculations on Tableau, Tableau Integration with Other Tools

PART 9 - NATURAL LANGUAGE PROGRAMMING WITH MACHINE LEARNING

Natural Language Processing - What is it used for? | NLTK Exploration, Word Tokenization, Different Types of Tokenizers, Bigrams, Trigrams & N-grams, Stemming & Lemmatization, Stopwords Removal, Part of Speech (POS) Tagging, Named Entity Recognition | Bag of Words | TF- IDF Vectorizer | Co-occurrence matrix | Text Similarity/Clustering | Latent Semantic Analysis(LSA) | Topic Modeling | Latent Dirichlet Allocation (LDA) | Text Classification - Sentiment Analysis | Recommender Systems - Collaborative Filtering

PART 10- TIME SERIES DATA ANALYSIS & FORECASTING

Introduction to Time Series Data | Correlation And Autocorrelation | Components of Time Series | Visualization Principles - Scatter Plot , Time Plot and Lag Plot | Auto-Correlation Function (ACF)/ Correlogram | Naive Forecast Methods | Errors in Forecast and Its Metrics | Model Based Approaches | Linear Model, Exponential Model, Quadratic Model | Auto Regression (AR), Moving Average (MA) | Autoregressive Moving Average (ARMA) | Autoregressive Integrated Moving Average (ARIMA) | Additive Seasonality | Multiplicative Seasonality | Random Walk | Smoothing Techniques | Moving Average, Exponential Smoothing | De-Seasoning and De-Trending

ELIGIBILITY
Graduate & Abv

DISCOUNT
Laptop Rs.2000
.....
DISCOUNT
Group Rs.2000

DURATION
6 Months

LAQSHYA
Institute of Skills Training
CALL : 8976789830