TB
TRAINING BASKET
QUALIFY FOR TOMORROW

Data science with python

About Training Basket

At Training Basket, we take pride in ourselves for maintaining our position in the technology-training ecosystem as the top quality-training provider in "Specialized Technologies" and thus being able to get the best placement package available for our students in blue chip companies. Training Basket not only teaches the students, but also undertake corporate workshops and conduct seminars in various engineering colleges. We conduct programmes to promote the build-up of adequate and industry ready IT professionals for development and distribution of knowledge and grasp out to the larger association through continuation, at the state, national and international levels.

Training Basket team works at delivering the best technical training and facilities. We provoke new knowledge by alluring in cutting-edge research and to bolster students and expert expansion by offering IT industry high demand training programs.

In addition, global technology leaders like Red Hat and Times of India (TSW) have chosen Training Basket as their Master Certified Training partner to deliver their authentic, certified training, to enterprises and IT professionals across India.
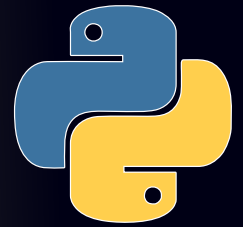
About Data Science with Python

Data Science. Data science is the study of data. It involves developing methods of recording, storing, and analyzing data to effectively extract useful information. The goal of data science is to gain insights and knowledge from any type of data - both structured and unstructured.

Some advantages of data science in business: Mitigating risk and fraud. ... They create statistical, network, path, and big data methodologies for predictive fraud propensity models and use those to create to alerts that help ensure timely responses when unusual data is recognized. Delivering relevant products

Why we use Python for Data Science?

Python provides a more general approach to data science. Python is better for data manipulation and repeated tasks,. Python is a general-use high-level programming language that bills itself as powerful, fast, friendly, open, and easy to learn. Facebook uses Python library Pandas for its data analysis because it sees the benefit of using one programming language across multiple applications.
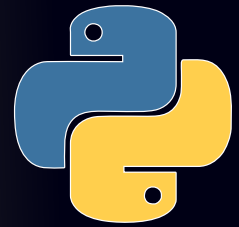
Data Science with Python

Course Duration : 96 Hours

Python Refresher (8 Hours)
- **Python for Data Science: Introduction**
  - Python, Anaconda and relevant packages installations
  - Structure of Python Program (Comments, Indentation)
  - Variables, Keywords and Data types in Python
  - Standard Input and Output
  - Operators

- **Python for Data Science: Data Structures**
  - Numbers
  - Strings
  - Lists
  - Tuples
  - Sets
  - Dictionary

- **Python for Data Science: Control Flow**
  - If, else, elif statements
  - While and For Loop
  - Control Statements (Pass, Break, Continue)

- **Python for Data Science: Functions**
  - Introduction (Built-In and User Defined Functions)
  - Types of user defined functions
  - Lambda functions

- File Handling
- Exception Handling

- **Modules and Packages**
  - Web scraping using BeautifulSoup
  - Database Handling with Python (Sqlite and MySQL)
- Object Oriented Programming: Introduction

1. Python Modules for Data Science (8 Hours)
- **Python for Data Science: Mathematical Computing with Python (numpy)**
  - Numpy Introduction (ndarray)
  - Numerical operations on Numpy
  - Numpy Overview
 - Basic operations, types of

- Initializing arrays (random, ones, zeros, full)
- Accessing elements
- Shape Manipulation
- Transpose
- Slicing
- Examples

- **Python for Data Science: Data Manipulation with Python(pandas)**
  - Understanding Series
  - Understanding DataFrame
  - View and Select Data
  - Missing Values
  - Data Operations
  - Indexing, Selection and Filtering
  - Dropping entries from an axis
  - Concatenation
  - Handling categorical Data (Get Dummies)

- **Python for Data Science: Data Visualization with Python(Matplotlib, Seaborn)**
  - Introduction to Matpotlib
  - Colours, Markers and line styles
  - Customization of Matplotlib
  - Plotting with Pandas
  - Barplots, Histograms plots, Density Plots
  - Introduction to Seaborn, Style Management
  - Plotting with Categorical Data
  - Visualizing Linear Relationships

2. **Data Science (16 Hours)**
   - **Introduction to Data Science**
     - What is Data Science
     - Data Science Buzzwords
     - Difference between Analysis and Analytics

     - **Difference between applying:**
       - Traditional Data Science, Big Data, BI and ML

   - **Data Analytics**
     - Data Analytics Process
   - Exploratory Data Analysis(EDA)
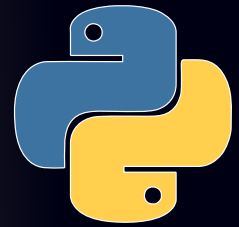
- **Statistics**
  - Introduction to Statistics

  - **Categories of Statistics**
    - Data Collection
    - Descriptive
    - Inferential
  - Population and Sample
  - Statistical Analysis Process
  - Data Distributions
  - Mean, Median, Mode
  - Variance and Standard Deviation
  - Covariance and correlation
 - Hypothesis Testing
- Data Wrangling

3. **Machine Learning (56 Hours)**
  - Introduction to Machine Learning
    - What is ML, AI, DL?
    - What is the difference between AI, ML and DL
    - Applications of Machine Learning
    - Types of Machine Learning
      - Supervised
      - Unsupervised
      - Reinforcement(Not in this scope)
    - Flow of Operation
    - Review of machine learning algorithms
    - Scikit learn
    - Introduction to SciKit Learn (sklearn)
    - Sample Dataset in SciKit Learn
    - Holdout Validation, K-fold cross Validation
    - Cross Validation using SciKit Learn
    - Train Test using SciKit Learn
    - Data preprocessing
    - Evaluation and improvement techniques
      - Accuracy measurement
      - Confusion Matrix

  - **Linear Regression:**
    - Introduction to Linear Regression
    - Understanding the real meaning of Linear Regression
    - Multiple Linear Regression and Non-linear Regression
    - Under fitting, Over fitting, Bias and Variance
    - Cost Function (Sum of Square Error)
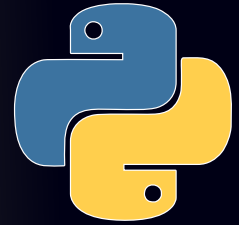
- Multiple Linear Regression using Gradient Descent based approach
- Coefficient of Determination (R^2)

- **Logistic Regression:**
  - Logistic regression vs Linear Regression
  - Can we use Regression Mechanism for Classification?
  - Logistic Regression – Deriving the Formula
  - Logistic Regression for Multi-class Classification
  - Logistic Regression Decision Boundary
  - Logistic regression on the titanic dataset
  - Visualizing a logistic regression model

- **Decision Tree:**
  - Introduction to Decision tree
  - Details of tree induction
  - GINI index computation
  - Entropy and information gain
  - Pruning
  - Metrics for performance Evaluation
  - Iris Decision Tree Example

- **Support Vector Machine (SVM):**
  - Linear Classifiers
  - Margin of SVM's
  - SVM optimization
  - SVM for Data which is not linear separable
  - Learning non-linear patterns
  - Kernel Trick
  - SVM Parameter Tuning
  - Linear SVM using Python
  - SVM with RBF kernel with Python

- **Random Forest:**
  - Comparison between Random Forest and Decision Tree
  - Introduction to Ensemble learning
  - Why Ensemble learning
  - Independently constructed ensembles for classification: Majority voting
  - Independently constructed ensembles for classification: Bagging
  - Independently constructed ensembles for classification: Random forests
  - Sequentially constructed ensembles for classification boosting
  - Sequentially constructed ensembles for classification boosting example
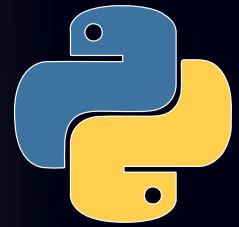
- **K Nearest Neighbours:**
  - Introduction to KNN algorithm
  - Decision boundary KNN Vs Decision tree
  - What is the best K
  - KNN Problems
  - Feature selection using KNNs
  - KNN using sklearn
  - Digits classification using KNN in Python
  - IRIS dataset classification using KNN

- **K Means:**
  - Applications of Clustering
  - Understanding Euclidean Distance
  - Basics of Clustering
  - Elbow Method
  - Hierarchical clustering
  - K-means Algorithm example

- **Dimensionality Reduction using PCA:**
  - What is PCA?
  - Understanding Matrix Transformations and Definition of Eigen Vectors
  - How is PCA Computed?
  - PCA Examples

- **Introduction to Natural Language Processing :**
  - Installing NLTK
  - Tokenize words
  - Tokenizing sentences
  - Stop words with NLTK
  - Stemming words with NLTK
  - Speech tagging
  - Sentiment analysis with NLTK

4. **Python for Data Science: Introduction to Big Data and Apache Spark (8 Hours)**
  - **Big Data and Spark with Python:**
    - Big Data Overview
    - Hadoop Core components
    - Python Integration with HDFS using Hadoop Streaming
    - Map Reduce
    - Spark Overview
    - Python Integration with spark using pyspark
    - Big Data Case Studies

JOB BASKET

Q: Why should I choose Training Basket over other training providers?

Ans. Training Basket provides a unique amalgamation of quality, convenience, flexibility and cost. Training Basket has some of the best trainers in the industry. Our trainers excel not just in depth and width of knowledge, but also in their patience and ability to explain difficult concepts in simple terms.

Training Basket has made serious investments with long-term vision for ensuring good environmental factor for studies by first buying their own suites in the prestigious iThum Towers in NOIDA, Sec-62. This lets us configure our labs and classroom suited best to our student's comfort and focused studies.

Q: What is the criterion for availing the Training Basket job assistance program?

Ans: All Training Basket students who have successfully completed their training in any of our courses are directly eligible for placement assistance.

Q: Which are the companies that Training Basket has placed students in the past?

Ans: We have exclusive tie ups with MNC's like Ericsson, Cisco, Cognizant, Tech Mahindra, MEON, Bingo, Genpact etc.

Q: Do I need a prior industry experience in getting an interview opportunity?

Ans: There is no need to have prior opportunity for getting an interview call. The successful completion of any industry level technology training at Training Basket is like an industry experience. This training makes you confident to clear interviews and we also conduct in-house mock interviews on our online assessment platform where we assess our student's kills by testing their code online or industry specific assessment before sending them for interviews.

Q: How does Training Basket assist in placement?

Ans: You will be guided on creating an attractive template based resume. You will get opportunity to attend free personality development program and mock interviews conducted by our SME's to boost your confidence for real interviews. Plus you will be given our level assessment platform where we assess our student's kills by testing their code online or industry specific assessment.

Q: If I don't clear in first attempt, will I get another chance?

Ans: Yes, for sure. Your resume will be active on our job portal and will be visible to all our associates and clients. Training Basket will continue to send your resume to future job requirements matching your profile till you land a job.

Q: Does Training Basket Guarantee job through it's job assistance program ?

Ans: Training Basket does not guarantee job placement but it will continue to assist you on best efforts basis to place you in it's affiliated companies' network.

COMPANIES WHERE OUR STUDENTS ARE PLACED

Tech Mahindra

DELL

Source Soft Solutions
We only believe in the best

Chi Networks

ADS-eVER

MEON
DELIVERING GREAT SURFACES

BioMax™

NECTAR PHYTOCHEMS

BINGO
Change • Innovate • Lead

Dimensions
Nine Dimensions Infotech

WTS Webtech Solutions

HCL

IBM

Infosys

accenture

HUAWEI

Reliance
Jio Infocomm

pinga™ solutions

DATA BRIDGE
MARKET RESEARCH

CoreIP
Technology Pvt Ltd

iSOURCe

ns network solutions

NUCLEUS SOFTWARE

Shailers
Solutions...for your need

hp

SAMSUNG

indiatech
Solutions