Data Science

KK NAGAR : No:10, PT Rajan Salai, KK Nagar, Chennai-600078
OMR : No 92, Rajiv Gandhi Salai (OMR), Navalur, adj. to AGS Cinemas, Chennai-600130

Website: www.slainstitute.com | www.softlogicsys.in | www.slajobs.com
Mob: +91 88707 67784 | +91 8608 700340 | +91 8681 884318

SLA is one of its kind Training cum Placement Company servicing 30+ countries, empanelling 100+ IT experienced mentors, doing 70+ trending technology training programs with more than 350+ IT MNC hiring partners.

SLA services

- **Individuals who are new to IT Industry to upskill and get placed in IT companies.**
- **IT professionals who want to upgrade technological skills to get qualified for the next salary hike and professional growth.**
- **Corporate Technology Training needs**

As a value addition, SLA, not only equip individuals in Technology skills but also offers free training in Aptitude, Interview skills and Basic Coding skills for anyone to successfully land in IT Industry. Training is given by IT experienced mentors who enable the individuals to do their **own Real Time Projects** by the end of the Program thereby giving **Hands-on training services**.

SLA doesn't involve in any backdoor job process for placement support instead support our fellow students with genuine end to end placement support through unlimited interviews with our IT hiring partners.

If you want to just have an IT certification, you can do your course anywhere. If you aspire to get into an IT Job or upgrade skills and settle in a better IT Job, then you should choose SLA. With SLA, you get life long free placement support and your IT dream career will come into reality for sure.

Please go through the long list of our student reviews / offer letters / testimonials @ www.joinsla.com to get to know more about us.

FUNDAMENTALS OF DATA SCIENCE – SYLLABUS

(Duration- 3 weeks)

[Prerequisites: Core python, Functional python – mandatory]

Chapter 1: Python Data science introduction
Chapter 2: Hands on with Pandas – Data Analysis library [Data Processing
Chapter 3: Working with Pandas Data
Chapter 4: Numpy – Mathematical Computation
Chapter 5: Hands on with Matplotlib library - [Basic Data Visualization
Chapter 6: Hands on with Data Distributions (using numpy, pandas, seaborn)
Chapter 7: Advanced Data Visualization using SEABORN

Chapter 1: Python Data science introduction

1. What is Data science
2. Introduction to python data science
3. Installation of Pandas,numpy,scipy,sklearn,seaborn,nltk
4. Basic terminologies of DS
    a) Data science
    b) Data scientist
    c) Data set
    d) Data mining
    e) Data visualization
    f) Data modeling
    g) Data wrangling
    h) Big data
    i) Machine learning
    j) Algorithms
    k) Deep learning

Chapter 2: Hands on with Pandas – Data Analysis library [Data Processing]

1. Why Pandas?
2. Features of Pandas
3. Data structures in Pandas
    a) Series
    b) DataFrame
    c) Panel
    d) Panel4D
4. Series creation
    a) Using ndarray
    b) Using dict
    c) Using scalar values
    d) Using list
5. Accessing elements of Series
    a) Using indexing
    b) Using slicing
    c) Using ranging
    d) Using iloc method
    e) Using loc method
6. Vectorizing operations
    a) Vector operations using same index values
    b) Vector operations using different index values
7. DataFrame creation
    a) Using list
    b) Using dict
    c) Using ndarray
    d) Using series

e) Using DataFrame

8. Viewing DataFrame elements
    a. Using describe function
    b. Using column name
    c. Using iloc method
    d. Using iat method
    e. Using head()
    f. Using tail()
    g. Using index method

Chapter 3: Working with Pandas Data

1. Handling missing values
    a) Using Dropna()
    b) Using Fillna()
    c) Using add between 2 vector series
2. Data operations with customized functions
    a) Using groupby()
    b) Using sorting
    c) Using merge
    d) Using duplicate
    e) Using concatenation

3. Statistical functions in data operations

a) Max()
b) Min()
c) Mean()
d) Std()

4. SQL operations in pandas
    a) Creating table using sqlite3
    b) Executing sql queries
    c) Inserting values
    d) Fetching records
    e) Creating recordset
    f) Display resultset
    g) Converting resultset into DataFrame
5. Data Processing
    a) Processing CSV data
    b) Processing JSON data
    c) Processing XLS data
    d) Date and time in data
    e) Reading html contents

Chapter 4: Numpy – Mathematical Computation

1. Why numpy?
2. Powerful properties of numpy
3. Types of arrays
   a. One dimensional
   b. Two dimensional
   c. Three dimensional
4. Attributes of ndarray
   a) Using .ndim
   b) Using .shape
   c) Using .size
   d) Using .dtype
5. Basic operations
   a. (+, -, *, /, %, //, &, |, ~, <, <=, >, >=, ==, !=)
   b. Accessing array elements using axis values
   c. Indexing with Boolean array
6. Creating functions for arrays
   a) Using arange()
   b) Using linspace()
   c) Using ones()
   d) Using zeros()
   e) Using diag()
   f) Using random.rand()
   g) Using random.randn()
   h) Using random.seed()

7. Copy and view
    a) Deep copy
    b) Shallow copy
    c) Simple assignment
8. Universal functions
    a) Sqrt
    b) Cos
    c) Floor
    d) Exp
9. Shape manipulation
    a) Using flatten
    b) Using reshape
    c) Using resize
    d) Using split
    e) Using stack
10. Broadcasting
    a) Using tile()
    b) Using ones()
    c) Using newaxis()

Chapter 5: Hands on with Matplotlib library - [Basic Data Visualization]

1. Chart properties
    a) Creating a chart
    b) Labeling the axes
    c) Formatting line style and color
    d) Saving the chart in a file
2. Styling the chart
    a) Adding annotations
    b) Adding legends
    c) Presentation style
3. Types of presentation styles
    a) Scatter plots
    b) Heat maps
    c) Bubble chart
    d) Bar chart
    e) Pie chart
    f) XKCD style
    g) 3D chart
    h) Box and whisker plots
    i) Time series plot
    j) Graph data / line graph
    k) Geographical data

Chapter 6: Hands on with Data Distributions (using numpy, pandas, seaborn)

1. Why and How Data to be distributed?
    a) Calculating mean
    b) Calculating median
    c) Calculating mode
    d) Measuring variance
2. Types of distribution
    a) Uniform distribution
    b) Normal / Gaussian distribution
    c) Exponential PDF
    d) Binomial distribution PMF
    e) Poisson distribution PMF
    f) Bernoulli distribution
    g) P value
    h) Correlation
    i) Chi-square test
    j) Linear regression

Chapter 7: Advanced Data Visualization using SEABORN

1. Visualization techniques used
    a) Histogram
    b) Histogram with grid
    c) Distplot
    d) Pairplot
    e) Scatterplot
    f) Lmplot
    g) box plot