

# Data Analytics & Science Syllabus

## A] Python for Data Science

### Anaconda Distribution installation

- Installation
- Jupyter Notebook introduction
- Jupyter Commands
- Magic commands

### An Introduction to Python:-

- What is Python and history of Python?
- Features of Python.
- Installation and Working with Python.
- Understanding Python variables.
- Python basic Operators.
- Python Identifiers, Keywords and Indentation.
- Understanding python blocks.
- Command line arguments.
- Getting User Input.
- Python Data Types.
- What are variables?

### Program Flow Control :-

- Conditional blocks using if, else and else if.
- For loops in python.
- For loop using ranges, string, list and dictionaries.
- Use of while loops in python.
- Loop manipulation using pass, continue, break and else.
- Programming using Python conditional and loops block.

## **List, Ranges, Dictionaries, Tuples and Sets in Python:-**

- Introduction.
- Lists in Python
- Tuples in Python
- Python Dictionaries
- Dictionary manipulation.
- Ordered Sets with tuples
- Sets
- Python Sets Examples

## **File Input and Output in Python:-**

- File opening modes in Python.
- Reading files in python.
- Writing files in python.
- Understanding built-in functions.

## **B] Data Science**

### **NumPy**

- Introduction to NumPy Array
- Installation of NumPy module
- Why NumPy? Advantages over Python List?
- The Basics of NumPy
- Defining MultiDimensional Array
- Printing of NumPy Array
- Accessing NumPy array using fancy, index, slicing methods
- Array Manipulations
- Array Math
- Logical Operations using NumPy Array
- NumPy functions
- NumPy Math functions
- Random number generation
- Dealing with NaN values within NumPy Array
- Date time handling using NumPy array
- Broadcasting/Reshape/Resizing

- Aggregate function
- Help on NumPy
- Save and load array to and from file

## Pandas

- Introduction to Pandas
- Installation of Pandas
- Reading different file formats with different parameters
- CSV
- Excel
- Fixed width file
- JSON
- Series & operations
- Data Frames
- Creating customized index
- Hierarchical indexing
- Adding header to the rows
- Data Frame functions
- Creating data frame using Series, Data Dictionaries, files, List, Nddarray
- Accessing data frames
- Accessing rows - slicing, fancy, index
- Accessing columns by name, index
- Loc, iloc
- Adding column/s
- Adding row/s
- Manipulating values within data frame
- Dropping column/s
- Conversion of column data type
- Aggregate functions
- Statistical functions on Data Frames
- Data Visualizations
- Different graphs like
- Histogram
- Bar chart
- Pie Chart
- Multi line chart

- Area chart

## MatPlotLib

- What is Matplotlib?
- Installation, configuration
- How to plot
- Histogram
- Bar Chart
- Pie chart
- Line chart
- Multi line chart
- Area chart
- Box plot
- Heatmap
- Scatter plot

## C) Machine Learning

- **Introduction to machine learning**
- **Introduction to scikit-learn, Keras, etc**
- **Regression**
  - Introduction classification problems, Identification of a regression problem, dependent and independent variables.
  - How to train the model in a regression problem.
  - How to evaluate the model for a regression problem.
  - How to optimize the efficiency of the regression model.
- **Classification**
  - Introduction to classification problems, Identification of a classification problem, dependent and independent variables.
  - How to train the model in a classification problem.
  - How to evaluate the model for a classification problem
- **Clustering**
  - Introduction to clustering problems, Identification of a clustering problem, dependent and independent variables.
  - How to train the model in a clustering problem.

- How to evaluate the model for a clustering problem.
- How to optimize the efficiency of the clustering model
- **Performance Metrics**
  - Classification reports - To evaluate the model on various metrics like
  - recall, precision, f-support, etc.
  - Confusion matrix - To evaluate the true positive/negative, false positive/negative outcomes in the model.
  - $r^2$ , adjusted  $r^2$ , mean squared error, etc.
- **Supervised Machine Learning**
- **Linear Regression**
- **Unsupervised Machine Learning**
- **K-Means Clustering**

## D) Django

- Introduction to Django
- Difference Between an App and Project
- Understanding Migrations and Their Purpose
- Admin Part: Creating a Superuser in Python
- Views in Django URL Routing
- Render and Relative Import
- Overview of Settings File in Django
- Configuring Templates in Django
- Understanding Models
- Linking Models to Admin
- Creating ModelForms
- Form Validation in Django
- Context in Django
- Forms in a View
- Custom Forms
- Setting up Email in Projects
- Static Files and Serving in Django
- Bootstrap Grid System
- Introduction to CSS and Blocks
- URL Names as Links
- Adding Authentication in Django Project
- Authentication Links in Navigation Bar

- Adding Login Form in Bootstrap
- Query Set Basics
- Rendering Images and Videos on Our Site

## **E] Structured Query Language:-**

- MYSQL Introduction
- Data Types
- DDL, DML, TCL
- Constraints
- DISTINCT Clause
- WHERE Clause
- MYSQL Conditions (AND, OR, BOOLEAN, LIKE, IN)
- MYSQL Functions (MIN, MAX, AVG, SUM, COUNT)
- ORDER BY Clause
- GROUP BY Clause
- Relationships in SQL
- Joins in SQL. Mini Project - 2

## **Python Database Connectivity (PDBC)**

- SQL Database connection using python.
- Install the MySQL dB and other Packages
- DML and DDL Operations with Databases.
- Performing Transactions.
- Handling Database Errors.
- Disconnecting Database.

## **SQLAlchemy - Object Relational Mapper**

- ORM Introduction
- SQLAlchemy Overview.
- SQLAlchemy over PDBC..
- Classical Way of Mapping
- Declarative Way of Mapping
- DML and DDL Operations with Database.
- Queries in SQLAlchemy.
- Applying Filters

## **Project**

- Core Project
- Live Project

## **Certificate**

SJA Academy