Note :

I will start with set of objective questions which give me idea of participants and also it will encourage participants to learn

Every fundamental of learning is backed by objective questions and Hands on

I will last my course with set of multiple choice questions which demonstrate the improvements in participants

This is 140 Hours (Approx)

In this course we are going to learn following Python Libraries

- copy
- NumPy
- SciPy
- Pandas
- Matplotlib
- Seaborn
- StatsModels
- Scikitlearn
- NLTK

Python Basic :

Python will be introduced to participants. Participants will learn about Python data type, typecasting etc..

- Introduction to Python
- Running Python from eclipse
- IDLE
- Data type in Python
  - Integer
  - Float
  - String
  - Complex
  - None
- Typecasting data
- Operators in python

Collections in python:

This section will provide through knowledge of Python collection.   Python collections are backbone to write   efficient code.

List
- Introduction to List
- List operations
  - Creation
  - Modification
  - Concatenation
  - Summation
  - Length
- List Comprehensions

Dictionary
- Introduction to Dictionary
- Dictionaries operations
  - Creation
  - Modification
  - Concatenation
  - Length

Set
- Introduction to Set
- Set Operations
  - Creation
  - Union
  - Intersection
  - Subtraction
  - Symmetric Difference
  - Concatenation
  - Length

Tuples
- Introduction to Tuples
- Immutability
- Tuples Operations
  - Length
  - Index

Conditionals and Looping :

This section will describe about implementing conditionals in Python. Conditionals and loops, helps software developer, to implement complex logic.

- Introduction to conditionals
  - if
  - if and else

- if elif and else
  - Introduction to looping
  - Different type of loop in Python
    - for loop
    - while loop
  - Loop control statement
    - break
    - continue
  - pass

Strings:

Participants will become efficient to solve problem related to String Operations.

- Introduction to Strings
- String Concatenation
- String Search
- Substitution
- String Normalization
- String Joining
- Splitting Strings
- Some more Operations.

Functions in Python:

Function help programmer to organize code in efficient manner. It helps to write efficient clean code. Code re-usability is increased.

- Inbuilt Functions
- User Defined Functions and its definition
- A simple function
- return statement
- Local and global variables
- Default argument
- Variable length arguments
- Variable length key-value arguments
- Anonymous functions or Lambda Functions

Module and Packages

Modular code improve code clarity. Code expressiveness. This section will enable participants to write modular code.

- Introduction to Modules and Packages
- Modules in Python
- Packages in python
- Writing a Module
- Importing Module
- Selective Import
- Writing packages
- importing packages

Some useful packages in Python

Package os and shutil of Python help programmers to perform automation job. Like creating files and directories using python code.

os
- Joining path
- Creating new directory
- Absolute and relative path
- File size
- Getting content of a folder

shutil
- Copying files and directories
- Deleting files and directories
- Renaming files and directories

Class and Objects :

Python is multi-paradigm language. A Python programmer can write python code using Procedural , Object oriented programming and Functional Programming Paradigm. This section will enable participants to write code using Object oriented programming paradigm.

- Introduction to Programming Paradigms
- Introductions to Object Oriented Programming
- Class in Python
- Constructors
- Objects
- Instance Variables and Class Variables

- Different type of methods
- Instance Methods
- Class Method
- Static Method
- Magic Methods

Exception handling in python

Exceptional situation is inevitable in Nature. Programming is also not exception. Learning Exception handling in Python will enlighten participants to handle exception in code.

- Introduction to Exception
- Exception handling
  - try
  - except
  - else
  - finally
- Inbuilt Exceptions
- User defined exceptions

NumPy

NumPy is very old data science library, which internals have been written using C for mathematical computation efficiency. Direct integration of Python to C Code made NumPy faster. This deals with linear algebra problems.

- Introduction to NumPy.
- NumPy data type
- NumPy array and its operations
- NumPy ndarray
- Indexing slicing and Stacking of array and ndarray
- Manipulating array shapes
- Splitting arrays
- Data aggregation in array
- Understanding Summary statistics
- Broadcasting in NumPy Array
- Matrices and its operation in NumPy
- Linear systems and NumPy
- File IO in NumPy

Pandas

Data preprocessing is an inevitable part of data analysis. Pandas do it well. Pandas is easy to learn. This part of the course makes participants efficient to perform Data Preprocessing and Exploratory data analysis.

- Data Preprocessing
- Use of Data Preprocessing
- Introduction to Pandas.
- Pandas data structure
  - Series
  - DataFrame
- Series
  - Introduction to series
  - Series Architecture
  - Series Index
  - Positional and Labeled index
  - Series Indexing
  - Series Slicing
  - Series Grouping
  - Series Aggregation
  - Summary Statistics on Series data
  - Time Series data consideration in series
  - Series Concatenation
- DataFrame
  - Introduction to DataFrame
  - DataFrame Creation
  - Indexing
  - Slicing
  - Filtering
  - Adding column
  - Deleting column
  - Making some column index
  - Index to column
  - Data filtering
  - Summary statistic of data
  - Understanding Split apply combine logic
  - Data grouping
  - Data aggregation
  - Data joining
  - DataFrame Concatenation
  - Time Series data

- Missing values imputation

Data Visualization in Python :

For Data visualization we are going to use two libraries . The library Seaborn is defined over Matplotlib and remove some boiler plate code and provide more appealing charts. Let me start with course topics of Matplotlib, and thereafter I will enlist the topics of Seaborn.

Matplotlib and Seaborn :

- Introduction to Matplotlib.
- Line plot
- Multiple line plot
- Scatter plot
- Bubble Plot
- Plotting a mathematical functions
- Horizontal Bar plot
- Vertical Bar plot
- Stacked Bar plot
- Side by side Bar plot
- Histograms
- Box plot
- Who compute Bar plot statistics
- Multiple box plot
- Pi plot
- Different aspects of coloring
- Violin plots
- Text plot
- Understanding Concept of Figure and Axes
- Facet in Figure
- Word Cloud

Story telling with data visualization.

SciPy and Statistics

Without knowledge of probability and statistics, data science is like food without nutrition. This module will start with classical definition of probability. Participants will understand SciPy and use it for probability and statistical analysis.

- Introduction to Probability.
- Law of Probability

- Random Numbers
- Introduction to random numbers and variables
- Concept of Probability Distributions
- Discrete Random Variables
- Probability Distribution
- Concept of Expectation
- Concept of Variance
- Different discreet Random Variables
  - Bernoulli
  - Geometric
  - Binomial
  - Poisson
- Continuous Random Variables
- Probability Density function
- Cumulative density function
- Concept of expectation
- Concept of variance
- Some Important contiguous distribution
  - Exponential Distribution
  - Normal Distribution
  - Standard Normal Distribution
  - t-Distribution
  - Chi-Square Distribution
- Independence.
- Bayes theorem
- Confidence Interval
  - Introduction to Confidence Interval
  - Confidence interval for mean
  - z-Distribution
  - t-Distribution
- Hypothesis Testing
  - Introduction to Hypothesis testing
  - How to define Null Hypothesis
  - z-test
  - Test of normality
  - t-test
  - chi-square test
- test for variance

Machine Learning and Scikit-Learn :

Machine learning is one of the most important part of data science. This section will discuss about machine learning primitives. Some importation topics of data preprocessing such as data transformation and categorical transformations will be discussed.

- Introduction to Data Mining and Machine learning
- Supervised and unsupervised learning and semisupervised .
- Some use cases on machine learning.
- Introduction to sklearn
- Different submodules of sklearn
- Data normalization and scaling in sklearn
- Transforming string categorical value to numerical
  - Label Encoding
  - One Hot Encoding
- Steps in data analysis.
- Dimensionality Reduction And PCA

Simple Linear Regression Multiple Linear Regression

Old methods but still Simple linear regression and Multiple linear regression is very important pillar in Data Science. This section will enable participants to understand statistical concepts about Linear regression. Implementing Linear regression models using sklearn python package.

- Introduction to simple linear regression.
- Assumption to simple linear regression.
- Parameter estimation
- Understanding Least square estimation
- Simple linear regression with Scikit-Learn
- Multiple linear regression with Scikit-Learn
- Model verification and linear regression assumption testing.
- Data transformation and polynomial regression
- Introduction to Ridge regression
- Ridge regression and Scikit Learn
- Introduction to Lasso Regression
- Lasso Regression and Scikit Learn
- Measure the performance of a Regressor
  - R square
  - Adjusted R Square
  - Mean square error
  - Mean absolute error

Classification :

If you want that your computer program should classify among birds and mammals. Then this chapter is certainly help you. Participants will understand theoretical aspects and programming implementation of classification algorithms.

- Introduction to classification
- Classification use cases
- K Nearest neighbor
  - Introduction to K Nearest neighbor
  - K Nearest neighbor in Scikit Learn
  - Strength and weakness of K Nearest neighbor
- Logistic Regression
  - Introduction to logistic regression
  - Use cases of logistic regression
  - Logistic Regression Algorithm
    - Logit Function
    - Logistic Function
    - Maximum Likelihood Estimator
  - Logistic regression with Scikit Learn
  - Performance of classifier
    - Confusion Matrix
    - Precision
    - Recall
    - Accuracy
    - ROC curve
    - Area under curve
    - F1 Score
- Naive Bayes Classifier
  - Introduction to Bayes theorem
  - Bayes theorem in classification
  - What is Naive in Naive Bayes's
  - Naive Bayes' Classification Algorithm
  - Bayes classifier with Scikit Learn
- Decision Tree
- Introduction to decision tree.
- Use cases of decision tree
- Partition algorithms for decision tree
  - ID3
  - Gini Impurity

- Cart
- Tree pruning
- Scikit Learn and decision tree
- Ensemble methods
- Different type of Ensemble Method
  - Bagging Method
  - Boosting Method
  - Random Forest
- Bagging Method
  - Introduction to Bagging
  - How bagging improves accuracy
- Boosting Methods
  - Adaboost
  - Gradient boosting
- Random forest
- How randomforest decrease variance and improve accuracy

More Algorithm on Regression

Regression analysis can be performed with decision tree and ensemble algorithms.

- Decision tree for regression
  - Variance : use to create decision tree regressor
- Different type of Ensemble Method
  - Bagging Method
  - Boosting Method
  - Random Forest
- Bagging Method
  - Introduction to Bagging
  - How bagging improves accuracy
- Boosting Methods
  - Adaboost
  - Gradient boosting
- Random forest
- How randomforest decrease variance and improve accuracy

Neural Networks :

Deep learning is now a days, becoming popular day by day. The basic building block of deep networks, is artificial neuron. In this tutorial, we are going to learn perceptron, artificial neurons, and multiple layer perceptron. Neural networks can be used for classification problems as well as Regression. In fact neural networks can be used in unsupervised algorithms too.

- Introduction to Neural Network
- Introduction to Perceptron
- Perceptron Structure
- Input, Aggregator, Activation function
- Perceptron learning algorithm
- Perceptron as a classifier
- Perceptron on logic gates
- A simple artificial Neuron
- Activation function revisit
- Different type of Activation Function
- Multi-layer Perceptron
- MLP learning algorithm
- Neural network classifier with sklearn
- Neural network regressor with sklearn

Clustering

Your company wants to segment customers so that different product promotions can be send to the customer groups, then this section certainly help participants. This is not all the learning of clustering can help participants to make cluster of flowers to cluster of galaxy.

Introduction to clustering
Use cases of clustering
K-means clustering
Hierarchical clustering
  ○ Different linkage type: Ward, complete and average linkage

DBSCAN

Clustering performance evaluation

- ○ Adjusted Rand index
- ○ Mutual Information based scores
- ○ Homogeneity, completeness and V-measure
- ○ Fowlkes-Mallows scores
- ○ Silhouette Coefficient

NLTK

Text mining is becoming popular day by day as it is solving new and complex problems. In this section we are going to learn Text Mining using NLTK. Participants will create a spam sms classifier.

Introduction to natural language processing.

Introduction to NLTK

Installing NLTK

Text analysis basics

- ○ Tokenization
- ○ Stemming
- ○ Stop words
- ○ Part of speech tagging
- ○ Lemmatization

NLTK Corpora

Clustering