Table of Contents

|  |  |  |
| --- | --- | --- |
| **S.No.** | **List** | **No. of days** |
| 1 | **Introduction**   * 1. Introduction to the course   2. What is data science?   3. Why is data science important and where is it used?   4. Python in data science | 1 |
| 2 | **PYTHON**  **2.1 Introduction to python**  **2.2 IDEs, Installing and getting started**  **2.3 Numbers, Strings and Boolean**  Find the solution of a Quadratic equation.  A quadratic equation is of the form ax^2 + bx + c = 0. Values of a, b, c will be given as input.    Make a Basic Calculator.  It will perform 6 functions : Add, Subtract, Multiply, Divide, Power of x to y, Square. Input will be two numbers, so the square option will square both the numbers separately.    Basic Strings and Boolean operations.    **2.4 Lists, Dictionaries, Tuples, Sets**  Program for Dictionaries.  Add another value to an existing dictionary.    Create a list. Convert it into a tuple using inbuilt function. Then, use count() function to count how many times does a number appear in the tuple.    **2.5 Control flow, functions**  Program to generate and print a dictionary that contains a number (between 1 and n) in the form (x, x\*x).    Insertion Sort on an Array    Bubble Sort on an array.    Perform Linear Search on an array.  Linear Search searches for a given element in an array element-wise starting from the first element till last. If the element is not present, it returns -1.    **2.6 OOP in Python**  Make a kid class with first name, last name, address(city). Make another class which inherits from kid class. This class, called the student class, contains the roll no and class od the kid as well. Add methods to print inforamtion in both the classes.    Make a class revstring. Define a method in it to reverse a string.    Polymorphism in a class. | 3 |
| 3 | **Python – II**  **3.1 Lambda functions**  Sorting list of dictionaries with lambda function.    Finding intersection of two array using lambda function.    Add two lists using lambda and map function.    **3.2 Generators, Decorators**  Sum of fibonnaci numbers using generators    Use of decorators.    **3.3 Errors and Exceptional Handling**  Making the same basic calculator, but this time using error handling to handle error in case of division by zero.    We define a class derived from ‘Exception’ class. Then define a new exception in it.    **3.4 Regular Expressions**  Change the date format from YYYY-MM-DD to DD-MM-YYYY.    Extract an email from a string using regular expressions.    To check whether a string contains any special characters or not.    **3.5 Modules and packages**  Use of Random module.    **3.6 Name and main**  Finding the longest word in a file.    Find the size of a file. | 3 |
| 4 | **Numpy**  **4.1 Introduction, Arrays in Numpy**    Making various dimensioned array. Also, Finding the dimensions of arrays of different dimensions.    **4.2 Basic Operations, slicing**  Slicing 1-D array and 2-D array in various ways.    Reversing the entire array.    Make a numpy array of length 15. Slice it into two array x, y. where x is the array containing elements at even places and y is the array containing elements at odd places.    Basic functions on NumPy array.    **4.3 NumPy for basic arithmetic**  Performing basic arithmetic operations on two numpy arrays.  Performing matrix multiplication on two numpy arrays (matrices).    Performing basic operations on complex numpy array.  **4.4 Solve equation with NumPy**  Find out the eigen values and eigen vectors of a matrix.    Find out the Determinant of a matrix.    Find out the inverse of a matrix.    Find out the solution of a given set of equations using the above methods and mathematical logic.    Find out the solution of a set of equations using built in function in NumPy.    **4.5 NumPy for statistical Operations**  Find out the Maximum and the minimum values in a multidimensional array.    Use of ptp() function.    Find the mean and median of the array.    Find out the average and standard deviation. | 1 |
| 5 | **SkiPy**  **5.1 Introduction**  **5.2 Sub package for integration and optimization**  Performing simple integration using skipy.    Performing double integration using skipy.    Using skipy for optimisation using Nelder-Mead Simplex Method or Powell’s Method.    **5.3 Calculating Eigenvectors, eigenvalues**  Finding Eigenvectors using skipy.    **5.4 Subpackage – static, weave, IO**  Opening a MATLAB file.    **5.5 Linear Algebra using SkiPy**  Finding the determinant of a matrix using SkiPy.    Using skipy for linear Algebra. | 2 |
| 6 | **Pandas**  **6.1 Introduction, dataframes**  Introduction to dataframes.    Adding a new column to dataframes.    Selecting a column data from dataframes for display.    Adding a new row to dataframes.    Selecting a row data from dataframes for display.    **6.2 Missing data, group by**  Viewing missing values    Finding missing data using isnull(), notnull()    Filling missing values    Dropping missing Values    Split the data into groups and view the groups.    Split the data into groups and Iterate through the groups.    **6.3 Merging, Joining and concatenating**  Merge two dataframes using key = id.    Merge two dataframes using multiple keys.    Depict left, right and outer join.    Concatinating two dataframes without keys, using keys.    **6.4 Operations**  Performing mathematical operations using pandas. We first read a csv file, then perform mathematical operations on it. We assume this file only contains mathematical values.    Performing mathematical operations using pandas. We create two numpy arrays. Then, we convert them into dataframes and apply few mathematical operations. | 3 |
| 7 | **Matplotlib**  **7.1 What is data Visualization? Its Importance.**  **7.2 Introduction to matplotlib**  Introduction to Matplotlib. Making a simple plot with x, y labels.    Using Matplotlib to make plots of sine wave, cos wave.    Making plots of trigonometric functions, many in one.    **7.3 Line chart, Histogram, Scatterplot**  Plotting Line chart.    Plotting Histogram.    Making a Scatter Plot.    **7.4 Bar chart, Pie chart, Boxplot**  Plotting Bar chart.    Using Box Plot for plotting values.    Plotting Pie chart. | 2 |
| 8 | **Seaborn**   * 1. **Introduction to Visualisation with Seaborn**   Introduction to Seaborn.    Importing datasets in Seaborn.    Importing datasets in Seaborn (using dataframes).     * 1. **Distribution Plots, Categorical Plots**   Distribution Plots in Seaborn.    Categorical Plots in Seaborn using stripplot().    Categorical Plots in Seaborn using swarmplot().     * 1. **Matrix Plots, Regression Plots**   Matrix Plots in Seaborn.    Regression Plots in Seaborn.     * 1. **Grids, Style and Colour**   Grids and style in Seaborn.      Colors in Seaborn. | 2 |
| 9 | **Data Visualisation using Plotly and Cufflinks** | 1 |
| 10 | **Machine Learning**  10.1 Introduction  10.2 ML with Python  10.3 Why is it important? | 1 |
| 11 | **Linear Regression**  **11.1 Introduction**  **11.2 ScikitLearn**  Introduction to Scikitlearn.    Splitting Datasets in Scikitlearn.    **11.3 Linear regression**  Linear Regression in Scikitlearn.    **11.4 Logistic regression**  Logistic Regression in Scikitlearn. | 2 |
| 12 | **KNN and SVM**  **12.1 K Nearest Neighbours introduction**  **12.2 KNN theory**  **12.3 KNN with Python**    **12.4 SVM Classification – Linear, Non linear**  **12.5 Support Vector Regression**    **12.6 K Means Clustering**  **12.7 K means Algorithm** | 4 |
| 13 | **Natural Language Processing**  **13.1 NLP Introduction**  **13.2 NLP Theory**  **13.3 NLP with Python**  Naïve-Bayes Classifier Algorithm    Random Forest Algorithm | 1 |
| 14 | **Recommendation System with Python** | 2 |
| 15 | **Big Data and Spark with Python**  **15.1 Big Data Introduction**  **15.2 Local Spark Set-up, Spark Intro**  **15.3 PySpark setup**  **15.4 Spark with Python**    **15.5 RDD Transformation and actions** | 3 |
| 16 | **Extra Algorithms**  Decision trees Algorithm    Apriori algorithm    PCA    AdaBoost algorithm    Mean Shift Algorithm |  |