DATA ANALYTICS WITH PYTHON (Minor Project)

Chapter1-Introduction to Python

Learning Objectives: You will get a brief idea of what Python is and touch on the basics.

Topics:

1.1Overview of Python
1.2The Companies using Python
1.3Different Applications where Python is used
1.4Discuss Python Scripts on UNIX/Windows
1.5Values, Types, Variables
1.6Operands and Expressions
1.7Conditional Statements
1.8Loops
1.9Command Line Arguments
1.10Writing to the screen

Hands On/Demo:

Creating "Hello World" code
Variables
Demonstrating Conditional Statements
Demonstrating Loops

Chapter2-Sequences and File Operation

Learning Objectives: Learn different types of sequence structures, related operations and their usage. Also learn diverse ways of opening, reading, and writing to files.

Topics:
2.1Python files I/O Functions
2.2Numbers
2.3Strings and related operations
2.4Tuples and related operations
2.5Lists and related operations
2.6Dictionaries and related operations
2.7Sets and related operations

Hands On/Demo:
Tuple - properties, related operations, compared with a list
List - properties, related operations
Dictionary - properties, related operations
Set - properties, related operations

Skills:
File Operations using Python
Working with data types of Python

DATA ANALYTICS WITH PYTHON (Minor Project)

Chapter 3- Functions,Loops,Modules,Errors,Exceptions

Learning Objectives: In this Module, you will learn how to create generic python scripts, how to address errors/exceptions in code and finally how to extract/filter content using regex.

Topics:

3.1Functions
3.2Function Parameters
3.3Global Variables
3.4Variable Scope and Returning Values
3.5Lambda Functions
3.6Object-Oriented Concepts
3.7Standard Libraries
3.8Modules Used in Python
3.9The Import Statements
3.10Module Search Path
3.11Package Installation Ways
3.12Errors and Exception Handling

Hands On/Demo:

Functions - Syntax, Arguments, Keyword Arguments, Return Values
Lambda - Features, Syntax, Options, Compared with the Functions
Sorting - Sequences, Dictionaries, Limitations of Sorting
Errors and Exceptions - Types of Issues, Remediation
Packages and Module - Modules, Import Options, sys Path

Skills:

Error and Exception management in Python

Working with functions in Python

Chapter 4-Introduction to Numpy ,Pandas and Matplotlib

Learning Objectives: This Module helps you get familiar with basics of statistics, different types of measures and probability distributions, and the supporting libraries in Python that assist in these operations. Also, you will learn in detail about data visualization.

Topics:

4.1NumPy - arrays
4.2Operations on arrays
4.3Indexing slicing and iterating
4.4Reading and writing arrays on files
4.5Pandas - data structures & index operations
4.6Reading and Writing data from Excel/CSV formats into Pandas
4.7matplotlib library

4.8Grids, axes, plots
4.9Markers, colours, fonts and styling
4.10Types of plots - bar graphs, pie charts, histograms

Hands On/Demo:
NumPy library- Creating NumPy array, operations performed on NumPy array
Pandas library- Creating series and dataframes, Importing and exporting data
Matplotlib - Using Scatterplot, histogram, bar graph, pie chart to show information, Styling of Plot

Skills:
Probability Distributions in Python
Python for Data Visualization

Chapter 5- Data Manipulation

Learning Objective: Through this Module, you will understand in detail about Data Manipulation

Topics:
5.1Basic Functionalities of a data object
5.2Merging of Data objects
5.3Concatenation of data objects
5.4Types of Joins on data objects
5.5Exploring a Dataset
5.6Analysing a dataset

Hands On/Demo:
Pandas Function- Ndim(), axes(), values(), head(), tail(), sum(), std(), iteritems(), iterrows(), itertuples()
GroupBy operations
Aggregation
Concatenation
Merging
Joining

Skills:

Python in Data Manipulation

Chapter 6-Advanced Packages for Statistical Analysis

Learning Objectives: In this module, you will learn the Advanced Packages and their implementation, for Various Statistical Analysis

6.1 SciPy: SciPy is a Python-based ecosystem of open-source software for mathematics, science, engineering and technical computing.
SciPy builds on the NumPy array object.
6.2 Scikit-Learn: Scikit-learn are a machine learning library for the python programming language.
It features various classification, regression and closeting algorithms.

6.3 Statsmodels: is a Python module that allows users to explore data, estimate statistical models, and perform statistical tests.

An extensive list of descriptive statistics, statistical tests, plotting functions, and result statistics are available for different types of data and each estimator.

Skills:

Python in Statistical Analysis

Chapter 7-Descriptive Statistics

7.1 Central Tendency: Identify situations in which the center of a distribution is valuable, different ways the center of a distribution can be calculated for symmetric and asymmetric distributions, handling outliers
7.2 Measures of Variability: How the spread of random variables can be meaningfully summarized in the context of the distribution shape, handling bias in estimators
7.3 Distributions of Random Variables: Computing and summarizing distribution of discrete and continuous random variables, parametric fitting.
7.4 Correlations: Auto and cross-correlations between random variables, correlation matrix, correlation function, relationship with statistical dependence and causation

Chapter 8-Supervised Learning

Learning Objective: In this module, you will learn Supervised Learning Techniques and their implementation

8.1 Classification: Introduction to frequentist and Bayesian algorithms such as linear discriminant, *k* nearest neighbors, decision tree, naïve Bayes classifier

8.2 Regression: Introduction to linear and logistic regression extending to generalized linear models, clarifying the notions of principal components and diagnostics

Skillset:You will be able to make the comparison between various Predictive models which we can use for Supervised Learning

Hands-on
Predict the Sales Revenue based on Advertisement in various media

Project: Regression Analysis

How to assess if you are paying correct price or not while buying a property? Price is very important function for any business. Correct price can create a real gap between profit and loss. In this case study we will take an example of property pricing to gain a deeper understanding of regression analysis.
Step – 1: Data Preparation A. checking the outlier B. Checking Missing Values and how to treat them. C. Basic bivariate and univariate analysis i.e. checking correlations, how the variables are distributed. **Step – 2: Traditional Regression Analysis with variable selection**