## 22ANL203: Python for Data Science

| Subject Code: | **22 ANL 203** | C I A Marks | 30 |
|---|---|---|---|
| No. of Lecture Hours / Week | 03 | End Exam Marks | 70(35+35) |
| Practical Component Hour/Week | 02 | Total Marks | 100 |
| Total Number of Lecture Hours | 75 | Exam Hours | 03 |
| Course Focus | **Employability** | **E**ntrepreneurship | **Skill Development** |

**Course Outcomes:**

**By the end of the course, student will able to**

CO-1    Learn how to use lists, tuples, and dictionaries in Python programs. PO2, PO7

CO-2    Learn how to use indexing and slicing to access data in Python programs. PO2, PO7

CO-3    Effectively visualize results. PO1, PO5

CO-4    Apply how to manipulate and analyze un-organized datasets. PO6, PO7

CO-5    Conduct statistical analysis and machine learning methods. PO2, PO6,PO7

| Contribution of Course Outcomes towards achievement of Program Outcomes & Strength of correlations (H-High, M-Medium, L-Low) | | PO1 | PO2 | PO3 | PO4 | PO5 | PO6 | PO7 |
|---|---|---|---|---|---|---|---|---|
| **22ANL 203** | **CO1** | | 3 | | | | | 2 |
| | **CO2** | | 2 | | | | | 3 |
| | **CO3** | 3 | | | | 2 | | |
| | **CO4** | | | | | | 2 | 3 |
| | **CO5** | | 3 | | | | 1 | 2 |

## Course Content

**UNIT-I Introduction to Python**

Python Processes, Python in the Real World – Installing Python Interactive Interpreter, Built in functions Python Built-in Data types, Numbers, Strings and Variables- Variables, Names, and Objects., List and Operators, Tuples and Operators -Dictionaries Operators.          **(9 Hours)**

**UNIT- II Function and its application to business**

Functions ( applying it to the basic business principles ), Working With Files Numpy, Numpy Array, Numpy Array Operations, Indexing, Slicing, Numpy Array, Numpy Exercise Pandas, Introduction To Series, Introduction To Data Frame, Loc, Iloc, Split, Merge and Append, Read, Write .Csv, .Html, Excel File.                                              **(9 Hours)**

**UNIT- III Python Libraries ( Matplotib, Seaborn )**

Visualization of Data with Pandas Matplot-Lib, Basic Plotting Plotting Terminology, Subplots, Special Plot Ploty, Basic Plotting – Plotly, Extend Basic Plot, Plotly Scatter and Line Chart, Bubble Chart, Histogram and Distribution Plot.                                              **(9 Hours)**

**UNIT- IV Introduction to data analysis (Theory) package (Stats Models, Scipy)**

Application of Statistical Techniques Using Python Data Pre Processing, Fillna Groupna, Missing Values, Outliers, Duplicates, Descriptive Statistics, Correlation (Bivariate, Cross Tabulation) , Multivariate Analysis ( Cross Matrix), **Simple** Linear Regression.                     **(9 Hours)**

**UNIT- V Multivariate analysis, and EDA,**

Exploratory Data Analysis using Dtale and Python Profiling and SweetViz**,** Introduction to Machine Learning (theory), Machine learning Concepts, Types of Machine Learning, Machine Learning Application.                                              **(9Hours)**

**Suggested Reading**

1. McKinney, W. (2017). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. United States: O'Reilly Media.
2. Unpingco, J. (2021). Python Programming for Data Analysis. Germany: Springer International Publishing.
3. Albon, C. (2018). Machine Learning with Python Cookbook: Practical Solutions from Preprocessing to Deep Learning. Japan: O'Reilly Media.
4. VanderPlas, J. (2016). Python Data Science Handbook: Essential Tools for Working with Data. United States: O'Reilly Media.

**List of Practical's**                                              **30Hrs**

1. Write a list of Operators in Python
2. Define a Programme for LISTS
   a. Length of a List
   b. List and Operators
   c. Joining of Two List

    d. Other List Operators

3. Define a Programme for Sets and Dictionaries and perform various operators for it

4. How to Write a Function in python prepare a function in writing all arithmetic operators

5. Write a programme for Tuple, Assignment operators and comparison operators and execute with the examples.

6. Frame steps involving handling a Data frame, Handling Missing Data, dropna, fillna, Grouping data, Read, Write .csv, .html, excel file ,

7. Write a programme for plotting various graphs using Matplot lib

    a. Scatter and line chart

    b. Bubble Chart

    c. Histogram and Distribution plot

    d. Trend line

8. Write a programme for plotting various graphs using Seaborn

9. Write a programme Categorical Data, Splitting Data Testing Set Normalize Data

10. Write a programme for application of Statistical techniques using Python

### Internal Examination Pattern
**22ANL203: Python for Data Science**

**Max.Marks : 30**                                            **Max.Time : (90min)**

(I IA Test )    (Descriptive Type) – Max Marks: 30Marks

        (i) Section A (10Marks): Set two questions from Unit I & II with internal Choice. Each question carries 5 marks.

        (ii) Section B (20Marks): Set two questions from Unit I & II with internal Choice. Each question carries 10 marks

(II IA Test )    (Descriptive Type) – Max Marks: 30Marks

        (i) Section A (10Marks): Set two questions from Unit III & IV with internal Choice. Each question carries 5 marks.

        (ii) Section B (20Marks): Set two questions from Unit III & IV with internal Choice. Each question carries 10 marks

**22ANL203: Python for Data Science**

**Max.Marks : 70(35TH +35P)**                     **Max.Time : (1.5Hrs+1.5 Hr)**

**(A)**     **Semester –End Examination (Descriptive Type) – Max Marks : 35Marks**

   **(i) Section A (10Marks): Set two questions from each unit with internal Choice. Each question carries 2 marks.**

   **(ii) Section B (25Marks): Set two questions from each unit with internal Choice. Each question carries 5 marks**

**(B)**     **Semester –End Examination (Practical Examination) – Max Marks: 35**

   **Evaluation Procedure**

   **(i) Experiments (Exam & Execution)**                **15 Marks**

   **(ii) Viva**                                          **10 Marks**

   **(iii) Continuous Assessment          :**             **10 MARKS**

   10 marks for the continuous assessment (Day to day work in the laboratory shall be evaluated for 10 marks by the concerned laboratory teacher based on the regularity/ record/viva). Laboratory teachers are mandated to ensure that every student completes 80%-90% of the lab assessments.

   **TOTAL: (A)+(B) =  70 MARKS**