

Program: B.Com Economics and Analytics				Semester: II	
Course: Data Analysis and Visualization				Course Code:	
AY:2024-25					
Teaching Scheme			Evaluation Scheme		
Lecture (Hours per week)	Practical (Hours per week)	Tutorial (Hours per week)	Credit	Continuous Assessment (CA) (Marks - 20)	Semester End Examinations (SEE) (60 Marks - in Question Paper)
3 LABS	-	-	3		
Course Learning Objectives:					
<ol style="list-style-type: none"> To understand the importance of data analysis and develop a strong foundation in Python's data analysis ecosystem for extracting and communicating insights from data To develop the ability to design compelling visualizations through Python's visualization tools and communicate information effectively to both technical and non-technical audiences To enhance critical thinking skills by analysing real-world data to draw conclusions and interpret visualizations to make data-driven decisions 					
Course Outcomes:					
<p>After completion of the course, learners would be able to:</p> <p>CO1: Apply principles of data analysis to extract insights from raw data with Python's data analysis libraries like NumPy and Pandas while being able to manipulate, clean and analyse data effectively</p> <p>CO2: Create engaging and informative visualizations using Python's tools such as Matplotlib and Seaborn while summarizing all of the data to communicate data-driven insights</p> <p>CO3: Analyse real-world datasets by identifying patterns and trends, develop meaningful conclusions and make informed decisions based on the interpretation of the data</p>					
Pedagogy:					
Hands-On Learning, Computer Laboratory Based Learning					
Outline of Syllabus: (per session plan)					
Module	Description				No of Hours
1	Introduction to Essential Data Analysis Libraries				15
2	Data Manipulation and Analysis Techniques				15
3	Interactive Data Visualization				15
Total					45
Practicals					-

Unit	Topic	No. of Hours
Module 1	<p>Introduction to Data Analysis and Essential Python Library</p> <p>Foundations of Data Analysis Understanding Data Analysis, The Standard Process of Data Analysis, The KDD Process, The Five Phases of Data Analysis and Visualization, Ways to do Data Analysis, The IDEs for Python Data Analysis</p> <p>Introduction to Basic NumPy Understanding Data Types in Python, The Basics of NumPy Arrays: 1D, Computation on NumPy Arrays: Universal Functions, Aggregations: min, max, and Everything in Between, Computation on Arrays: Broadcasting, Basic Indexing and Slicing, Sorting, Transposing, and Swapping Axes of Arrays, File Input and Output with Arrays</p> <p>Advanced NumPy Comparisons, Masks, and Boolean Logic, Multidimensional Array, Array Manipulations, Linear Algebra, Structured Data: NumPy's Structured Arrays, Exploring Numba</p>	15
Module 2	<p>Data Manipulation and Analysis Techniques with Python</p> <p>Getting Started with Pandas Introduction to Pandas Data Structures: Series and DataFrames, Essential Functionality, Hierarchical Indexing, Handling Missing Data, Data Transformation, String Manipulation</p> <p>Data Aggregation and Grouping Operations Aggregating, Combining, Grouping and Joining Pandas DataFrame, Pivot Tables, Summarizing and Computing Descriptive Statistics, High Performance Pandas: eval and query, Exploratory Data Analysis</p>	15
Module 3	<p>Interactive Data Visualization with Python and Case Studies</p> <p>Introduction to Visualization and its Libraries Why Data Visualization, A Brief Matplotlib API Primer, 3D Visualization with Matplotlib, Plotting with Pandas and Seaborn, Interactive Visualization with Bokeh, Other Visualization Libraries</p> <p>Examples and Case Studies for Analysis and Visualization Bitly Data from 1.USA.gov, MovieLens 1M Dataset, US Baby Names 1880-2010, USDA Food Database, Federal Election Commission Database, Financial Analysis, and More as Per Recent Dataset Availability</p>	15

Reference Books:

- [1]. Wes McKinney, “Python for Data Analysis”, 3rd Edition, O’Reilly Media Inc., August 2022
- [2]. Avinash Navlani, Armando Fandango, Ivan Idris, “Python Data Analysis”, 3rd Edition, Packt Publishing, February 2021
- [3]. Scott McCoy, “Murach’s Python for Data Analysis”, Murach, August 2021
- [4]. Jake VanerPlas, “Python Data Science Handbook”, 2nd Edition, O’Reilly Media Inc., December 2022
- [5]. Dr. Ossama Embarak, “Data Analysis and Visualization Using Python”, Apress, November 2018
- [6]. Peter Bruce, Andrew Bruce, Peter Gedeck, “Practical Statistics for Data Scientists” 2nd Edition, O’Reilly Media Inc., May 2020
- [7]. Jack Dougherty, Ilya Ilyankou, “Hands-On Data Visualization”, O’Reilly Media Inc., April 2021
- [8]. Ayodele Oluleye, “Exploratory Data Analysis with Python Cookbook”, 1st Edition, Packt Publishing, 2023

EXAM PATTERN FOR THE COURSE

1. Internal Class Test - 10 Marks Paper Pattern (Pen Paper Based)

Question 1	Fill in the Blanks	4 Marks
Question 2	Descriptive Question	3 Marks
Question 3	Descriptive Question	3 Marks

2. 10 Marks Internal Assignments / Practical Study / Case Study / Mini Project

3. 60 Marks Final Exam Paper Pattern

- Examination shall be conducted in machine test form in Batches (Max. Batch Size – 30 Students) in the computer laboratory
- One external examiner must be present along with the internal examiner (subject faculty in-charge) for the conduct of examination
- Question paper should have maximum number of distinct sets, kept faced down on table, from which student will pick up one question paper
- Duration of the examination is 2.5 Hrs
- Marks distribution is as follows:

1	Practical Question Two questions for 20 marks each = $2 * 20 = 40$ and One question for 10 marks = $1 * 10 = 10$	50 Marks
2	Viva Voce	05 Marks
3	Coursework Journal	05 Marks

Evaluation shall be done by the examiners, both internal and external, on machine in the computer laboratory

Students must prepare answer book during the examination with the code and output in it, which further must be printed