



**DHANALAKSHMI SRINIVASAN ENGINEERING COLLEGE
(AUTONOMOUS)**

(Approved by AICTE & Affiliated to Anna University, Chennai)
Re-Accredited by NAAC with 'A' Grade
Accredited by NBA for AERO, BME, CSE, ECE, EEE, IT & MECH.
PERAMBALUR-621212, TAMILNADU, INDIA.
Website: www.dsengg.ac.in



COURSE PLAN(2025-2026 ODDSEMESTER)

Designation/Department	AP/AI&DS			
Course Code/Name	U23AIT53/DATAEXPLORATIONANDVISUALIZATION			
Year/Section/Department	III / B / AI&DS			
Credits Details	L:3	T:0	P:0	C:3
Total Contact Hours Required	45			

Syllabus:

UNIT I CREATING AND MANIPULATING NUMPY ARRAYS	9
Creating arrays, indexing and slicing, mathematical operations with NumPy arrays	
UNIT II DATA MANIPULATION WITH PANDAS	9
Series and Data Frame objects; importing and exporting data from various file formats into pandas Data Frame; Data selection and filtering indexing, slicing, conditional filtering using boolean indexing	
UNIT III GROUPING AND AGGREGATION WITH PANDAS	9
Grouping data using Pandas, applying aggregation functions such as sum, mean, count, etc., to grouped data, using pivot tables and cross-tabulation for data summarization	
UNIT IV DATA VISUALIZATION WITH MATPLOTLIB AND SEABORN	9
Introduction to Matplotlib and Seaborn to plot data using figures and subplots, Plots-Line plots, scatter plots, and Bar plots, visualizing distributions using histogram and box plots, Customizing plot aesthetics and adding annotations	
UNIT V NOSQL DATAMANAGEMENT FOR BIGDATA AND VISUALIZATION	9
Introduction to Plotly library for interactive visualization; Creating interactive line plots, scatter plots, and bar plots; Adding interactivity with hover effects, zooming, and panning	

Objective:

The objective of this syllabus is to equip students with foundational and practical knowledge in data manipulation, analysis, and visualization using popular Python libraries. The course is designed to:

- ❖ Introduce the creation and manipulation of multidimensional arrays using NumPy for efficient numerical computing.
- ❖ Enable students to handle structured data using Pandas for importing, cleaning, filtering, and transforming datasets.
- ❖ Develop skills in grouping and aggregating data to derive insights using Pandas functions and pivoting techniques.
- ❖ Teach effective data visualization using Matplotlib and Seaborn to represent data through various static plots.
- ❖ Familiarize students with interactive data visualization using Plotly and introduce basic concepts of NoSQL data management relevant to big data scenarios

Text Book:

1. VanderPlas, J. "Python data science handbook: Essential tools for working with data". O'Reilly Media, Inc.", 2nd edition.
2. McKinney W. "Python for Data Analysis: Data Wrangling with Pandas, NumPy and IPython". 2nd edition. O'Reilly Media, 2018.
3. Molin S. Hands-On Data Analysis with Pandas, Packt Publishing, 2019.
4. Rahman K. "Python Data Visualization Essentials Guide: Become a Data Visualization expert by building strong proficiency in Pandas", Matplotlib, Seaborn, Plotly, Numpy, and Bokeh, BPB 2021.

Reference Book:

1. Chen D. Y, "Pandas for Everyone: Python Data Analysis", Pearson, 2018. Lars George, "HBase: The Definitive Guide", O'Reilley, 2011

Website:

- W1:** <https://yourusername.github.io/python-syllabus>
W2: <https://your-syllabus-name.netlify.app>

Online Mode of Study (if Any):

NPTEL details can be listed.

Data Visualization-Link: <http://nptel.ac.in/courses/106/106/106106179/>

Course Plan:

Topic Number	Topic	Reference Detail	Page Number	Mode of teaching	Number of Periods Required	Cumulative Period
UNIT I CREATING AND MANIPULATING NUMPY ARRAYS						
1	Creating arrays	T1	3-20	BB	1	1
2	Creating arrays	T2	25-45	BB	1	2
3	Creating arrays	T1	23-35	BB	1	3
4	Indexing and slicing	T3	50-70	BB	1	4
5	Indexing and slicing	T1	40-55	BB	1	5
6	Indexing and slicing	R1	90-110	BB & VIDEO	1	6
7	mathematical operations with NumPy arrays	R2	120-140	BB & VIDEO	1	7
8	mathematical operations with NumPy arrays	T1	60-80	BB	1	8
9	mathematical operations with NumPy arrays	T2	100-120	BB	1	9
Outcome of Unit I:						
CO1 : Create and manipulate NumPy arrays to perform data analysis.						
UNIT II DATA MANIPULATION WITH PANDAS						
10	Series and Data Frame objects	T3	80-100	BB & VIDEO	1	10
11	Series and Data Frame objects	T1	85-105	BB	1	11
12	Series and Data Frame objects	T1	110-130	BB	1	12
13	Importing and exporting data from various file formats into pandas Data Frame	R1	140-160	BB & VIDEO	1	13
14	Importing and exporting data from various file formats into pandas Data Frame	T1	135-155	BB	1	14
15	Importing and exporting data from various file formats into pandas Data Frame	T2	165-185	BB	1	15
16	Data selection and filtering indexing	T3	110-130	BB	1	16

17	Slicing	T1	140-160	BB	1	17
18	Conditional filtering using boolean indexing	T2	190-210	BB	1	18
Outcome of Unit II:						
CO2 : Use Pandas methods to import, export, and preprocess data from various sources						
UNIT III GROUPING AND AGGREGATION WITH PANDAS						
19	Grouping data using Pandas	T1	170-190	BB	1	19
20	Grouping data using Pandas	R2	220-240	BB & VIDEO	1	20
21	Grouping data using Pandas	T1	200-220	BB	1	21
22	Grouping data using Pandas	T1	230-250	BB	1	22
23	Applying aggregation functions such as sum, mean, count, etc.to grouped data	T3	140-160	BB	1	23
24	Applying aggregation functions such as sum, mean, count, etc.to grouped data	T1	260-280	BB	1	24
25	Applying aggregation functions such as sum, mean, count, etc.to grouped data	T2	230-250	BB	1	25
26	Applying aggregation functions such as sum, mean, count, etc.to grouped data	T1	270-290	BB	1	26
27	Using pivot tables and cross-tabulation for data summarization	T3	170-190	BB	1	27
Outcome of Unit III:						
CO3: Perform basic data manipulation tasks, including data cleaning, filtering, sorting, and merging on Pandas objects.						
UNIT IV DATA VISUALIZATION WITH MATPLOTLIB AND SEABORN						
28	Introduction to Matplotlib and Seaborn to plot data using figures and subplots	T1	300-320	BB	1	28
29	Plots	T1	320-340	BB	1	29
30	Line plots	R1	250-270	BB & VIDEO	1	30
31	scatter plots, and bar plots	R1	280-300	BB & VIDEO	1	31
32	visualizing distributions using histogram and box plots	R1	310-330	BB	1	32
33	visualizing distributions using histogram and box plots	T1	350-370	BB	1	33
34	Customizing plot aesthetics and adding annotations	T1	380-400	BB	1	34
35	Customizing plot aesthetics and adding annotations	T2	270-290	BB	1	35

36	Customizing plot aesthetics and adding annotations	T1	410-430	BB	1	36
Outcome of Unit IV: CO4 : Use grouping and aggregation operations in Pandas to summarize data in Series and Data Frame objects						
UNIT V NOSQL DATA MANAGEMENT FOR BIG DATA AND VISUALIZATION						
37	Introduction to Plotly library for interactive visualization	T1	440-460	BB	1	37
38	Introduction to Plotly library for interactive visualization	T1	470-490	BB	1	38
39	Creating interactive line plots	R3	290-310	BB & VIDEO	1	39
40	Creating interactive line plots	T2	300-320	BB	1	40
41	scatter plots and bar plots	T1	330-350	BB	1	41
42	scatter plots and bar plots	R1	360-380	BB & VIDEO	1	42
43	Adding interactivity with hover effects	R3	390-410	PPT	1	43
44	Adding interactivity with hover effects	R3	420-440	PPT	1	44
45	Adding interactivity with hover effects	R3	450-470	PPT	1	45
Outcome of Unit V: CO5: Analyze and interpret data based on grouped and aggregated results.						

Course Outcome:**COURSE OUTCOMES:**

At the end of the course the students would be able to:

CO 1: Create and manipulate NumPy arrays to perform data analysis.

CO 2: Use Pandas methods to import, export, and preprocess data from various sources.

CO 3: Perform basic data manipulation tasks, including data cleaning, filtering, sorting, and merging on Pandas objects.

CO 4: Use grouping and aggregation operations in Pandas to summarize data in Series and Data Frame objects.

CO 5: Analyze and interpret data based on grouped and aggregated results.

CO 6: Use Matplotlib and Seaborn to create static visualizations and plot to create interactive visualizations of data to communicate data insights

Course Outcome Vs Program Outcome Mapping:

Course Outcomes (COs)	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1:	3	2	1	2	3	–	–	–	–	1	–	2
CO2:.	3	3	2	2	3	–	–	–	–	1	–	2
CO3:.	3	3	2	2	3	–	–	–	–	1	–	2
CO4:	3	3	2	2	3	–	–	–	–	1	–	2
CO5.	3	3	2	3	3	–	–	–	–	2	–	3
CO6:	3	2	2	2	3	–	1	–	–	3	–	2

[Levels of correlation: 3 (High), 2 (Medium), 1(Low)]

Content Beyond Syllabus:

1. MultiIndexing in Pandas (Beyond UNIT II & III)
2. Plotly Dash for Web-Based Dashboards (Beyond UNIT V)

Internal Evaluation Components:

Webportal	Assignment	Components	Topic Number with Topic / Unit Details	Relevance to CO
Webportal 1	--	Assessment – I (60)	Unit I and II	CO 1 & CO2
	1	Assignment – Handwritten (20)	Mathematical operations with NumPy arrays	
	2	Assignment – Poster Presentation / PPT (20)	Importing and exporting data from various file formats into pandas Data Frame	
Webportal 2	--	Assessment – II (60)	Unit III and IV	CO3 & CO4
	3	Seminar (20)	Introduction to Matplotlib and Seaborn to plot data using figures and subplots	
	4	Case Study Report (20)	scatter plots and bar plots	
Webportal 3	--	Model Exam (75)	Unit I to V	CO1 to CO6
	5	MCQ (15)	Unit I to V	CO1 to CO6

	-	Course Attendance (10)	--	--
--	---	---------------------------	----	----

Submission Details:

Phase1(Before AT1)		Phase2(Before AT 2)		Phase3 (Model)
Assignment1	Assignment2	Assignment3	Assignment4	Assignment5

Google Class Code Details:

Class Name: DEV

Link: <https://classroom.google.com/c/Nzg5OTczMjg5NTg3?cjc=nseyuyp4>**PLAN OF ASSESSMENT TEST DISTRIBUTION OF MARKS:**

TEST	CO-MARKWISE DISTRIBUTION						BLOOM'S LEVEL MARK WISE DISTRIBUTION					
	CO 1	CO 2	CO 3	CO 4	CO 5	CO 6	BTL 1	BTL 2	BTL 3	BTL 4	BTL 5	BTL 6
AT-1												
AT-2												
MODEL												

Prepared By

Verified By

Approved By

PRINCIPAL