

## Department of Computer Science & Engineering (Data- Science)

### **COURSE MODULE OF THE SUBJECT TAUGHT FOR THE SESSION 2024-25 (ODD SEM)**

#### **Course Syllabi with CO's**

<b>Faculty Name: Dr. Anitha D B</b>			<b>Academic Year: 2024 - 2025</b>				
<b>Department: CSE- Data Science</b>							
<b>Course Code</b>	<b>Course Title</b>	<b>Core / Elective</b>	<b>Prerequisite</b>	<b>Contact Hours</b>			<b>Total Hrs/ Sessions</b>
				<b>L</b>	<b>T</b>	<b>P</b>	
21AD71	DATA VISUALIZATION	Core	Python Programming	3	-	-	40L
<b>Course Learning Objectives</b>	CLO 1. Understand and use various plot types with Python CLO 2. Explore and work with different plotting libraries CLO 3. Create effective visualizations CLO 4. Implement exemplary applications related to Network Programming and Web Service CLO 5. Exhibit the awareness of the importance and limitation of the exploratory data analysis paradigm						

#### **Topics Covered as per Syllabus**

#### **Module-1: Data Visualization and Data Exploration**

**Introduction:** Data Visualization, Importance of Data Visualization, Data Wrangling, Tools and Libraries for Visualization **Overview of Statistics:** Measures of Central Tendency, Measures of Dispersion, Correlation, Types of Data, Summary Statistics

**Numpy:** Numpy Operations - Indexing, Slicing, Splitting, Iterating, Filtering, Sorting, Combining, and Reshaping

**Pandas:** Advantages of pandas over numpy, Disadvantages of pandas, Pandas operation - Indexing, Slicing, Iterating, Filtering, Sorting and Reshaping using Pandas

#### **Text Book 1: Chapter 1**

#### **Module-2: Plots**

**Comparison Plots:** Line Chart, Bar Chart and Radar Chart; **Relation Plots:** Scatter Plot, Bubble Plot, Correlogram and Heatmap; **Composition Plots:** Pie Chart, Stacked Bar Chart, Stacked Area Chart, Venn Diagram; **Distribution Plots:** Histogram, Density Plot, Box Plot, Violin Plot; **Geo Plots:** Dot Map, Choropleth Map, Connection Map; What Makes a Good Visualization?

#### **A Deep Dive into Matplotlib**

Introduction, Overview of Plots in Matplotlib, **Pyplot Basics:** Creating Figures, Closing Figures, Format Strings, Plotting, Plotting Using pandas DataFrames, Displaying Figures, Saving Figures; **Basic Text and Legend**

**Functions:** Labels, Titles, Text, Annotations, Legends; **Basic Plots:** Bar Chart, Pie Chart, Stacked Bar

Chart, Stacked Area Chart, Histogram, Box Plot, Scatter Plot, Bubble Plot; **Layouts:** Subplots, Tight Layout, Radar Charts, GridSpec; **Images:** Basic Image Operations, Writing Mathematical Expressions

**Text Book 1: Chapter 2, Chapter 3**

### **Module-3: Simplifying Visualizations using Seaborn**

Introduction, Advantages of Seaborn **Controlling Figure Aesthetics:** Seaborn Figure Styles, Removing Axes Spines, Contexts; **Color Palettes:** Categorical Color Palettes, Sequential Color Palettes, Diverging Color Palettes; **Interesting Plots in Seaborn:** Bar Plots, Kernel Density Estimation, Plotting Bivariate Distributions, Visualizing Pairwise Relationships, Violin Plots;

### **Text Book 1: Chapter 4**

### **Module-4: Plotting Geospatial Data**

Introduction, Geoplotlib, The Design Principles of Geoplotlib, Geospatial Visualizations, Tile Providers, Custom Layers, Introduction to Folium

**Visualizing Data:** Building a Google map from geocoded data, Visualizing networks and interconnection and Visualizing mail data

**Making Things Interactive with Bokeh:** Introduction, Bokeh, Concepts of Bokeh, Interfaces in Bokeh, Output, Bokeh Server, Presentation, Integrating, Adding Widgets

### **Text Book 1: Chapter5, Chapter 6**

### **Module-5: Networked Programs**

**HyperText Transfer Protocol** – HTTP, The World’s Simplest Web Browser, Retrieving an image over HTTP, Retrieving web pages with urllib, Parsing HTML and scraping the web, Parsing HTML using regular expressions, Parsing HTML using BeautifulSoup, Reading binary files using urllib

**Using Web Services-** eXtensibleMarkup Language – XML, Parsing XML, Looping through nodes, JavaScript Object Notation – JSON, Parsing JSON

### **Text Book 2: Chapters 12 and Chapter 13**

### **List of Textbooks**

#### **Text Books**

1. Data Visualization workshop, Tim Grobmann and Mario Dobler, Packt Publishing.
2. Python for Everybody: Exploring Data Using Python
3. Charles R. Severance, Create Space Independent Publishing Platform, 1st Edition, 2016

#### **Reference:**

1. “Data Visualization”: A Successful Design Process, Kirk, Andy, Packt Publishing Ltd,2012
2. Think Python: How to Think Like a Computer Scientist, Allen B. Downey, Green Tea Press, 2nd Edition, 2015
3. Interactive Data visualization for the Web, Murray, Scott, O’Reilly Media, Inc., 2013 27.09.2022
4. Visualizing Data: Exploring and Explaining Data with The Processing Environment, Fry, Ben, O’Reilly Media, Inc., 2007

#### **Web links and Video Lectures (e-Resources)**

1. <https://www.youtube.com/watch?v=eFBYJkA3ti4>
2. <https://www.youtube.com/watch?v=JhK2qVi5dC4>
3. <https://www.youtube.com/watch?v=UjYzNhBVIvY>
4. <http://book.visualisingdata.com/>
5. <https://matplotlib.org/>
6. <https://docs.python.org/3/tutorial/>
7. <https://www.tableau.com/>

<b>Course Outcomes</b>	CO 1. Demonstrate the data visualization techniques. CO 2. Analyze data represented in the form of graphs & charts CO 3. Experiment with different visualization tools CO 4. Identify geospatial data and interconnection of data. CO 5. Make use of the web for data extraction
<b>Assessment Details (both CIE and SEE):</b>	
The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures not less than 35% (18 Marks out of 50) in the semester-end examination(SEE), and a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together	
<b>Continuous Internal Evaluation:</b>	
<b>Three Unit Tests</b> each of 20 Marks (duration 01 hour) 1. First test at the end of 5th week of the semester 2. Second test at the end of the 10th week of the semester 3. Third test at the end of the 15th week of the semester	
<b>Two assignments</b> each of 10 Marks 4. First assignment at the end of 4th week of the semester 5. Second assignment at the end of 9th week of the semester 6. <b>Activity:</b> Group discussion/Seminar/quiz any one of three suitably planned to attain the COs and POs for 20 Marks (duration 01 hours) At the end of the 13th week of the semester	
<b>The sum of three tests(60M), two assignments(20M), and quiz/seminar/group discussion(20) will be out of 100 marks and will be scaled down to 50 marks</b>	

### The Correlation of Course Outcomes (CO's) and Program Outcomes (PO's)

Subject Code	21AD71	TITLE: DATA VISUALIZATION												Faculty Name	Dr Anitha D B			
List of Course Outcomes	Program Outcomes													Total				
	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12						
CO-1	2	-	-	-	3	-	-	-	-	-	-	-	5					
CO-2	2	2	2	-	3	-	-	-	-	-	-	-	9					
CO-3	2	-	2	-	3	-	-	-	-	-	-	-	7					
CO-4	2	-	2	-	3	-	-	-	-	-	-	-	5					
CO-5	2	2	2	-	3	-	-	-	-	-	-	-	9					
<b>Total</b>	<b>10</b>	<b>04</b>	<b>08</b>	-	<b>15</b>	-	-	-	-	-	-	-	<b>35</b>					

**Note:** 3 = Strong Contribution, 2 = Average Contribution, 1 = Weak Contribution, - = No Contribution

### The Correlation of Course Outcomes (CO's) and Program Specific Outcomes (PSO's)

Subject Code	21AD71	TITLE: DATA VISUALIZATION			Faculty Name	Dr Anitha D B	
List of Course Outcomes	Program Specific Outcomes						Total
	PSO-1		PSO-2		PSO-3		
CO-1	2		2		-		4
CO-2	2		2		-		4
CO-3	2		2		-		4
CO-4	2		2		-		4
CO-5	2		2		-		4
<b>Total</b>	<b>10</b>		<b>10</b>		-		<b>20</b>