|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **DATA VISUALIZATION LABORATORY**  **(Effective from the Academic Year-2022-23)**  **SEMESTER - VII** | | | | | | | | | | | | | | | | | | | | | |
| **Course Code** | | | | | | | | | **21DSL75** | | | | | **CIA Marks** | | | | | **50** | | |
| **Number of Contact Hours/Week (L: T:P:S)** | | | | | | | | | **0:0:2:0** | | | | | **SEE Marks** | | | | | **50** | | |
| **Total Hours of Pedagogy** | | | | | | | | | **24 P+ 12 S** | | | | | **Exam Hours** | | | | | **3 Hours** | | |
| **CREDITS –** | | | | | | | | | | | | | | | | | | | | | |
| **Prerequisites:**   * Knowledge of Mathematics and Python Programming. * Usage of IDEs like Jupyter Notebook/PyCharm, VSCODE | | | | | | | | | | | | | | | | | | | | | |
| **Course Objectives:**   * Make use of Data sets in implementing the data visualization techniques * Implement the data visualization techniques * Integrate machine learning libraries and mathematical and statistical tools that are suitable for the Data Science applications under consideration | | | | | | | | | | | | | | | | | | | | | |
| **Descriptions:**  Installation procedure of the required software must be demonstrated, carried out in group sand documented in the journal. | | | | | | | | | | | | | | | | | | | | | |
| **LIST OF EXPERIMENTS** | | | | | | | | | | | | | | | | | | | | | |
| **Exp.**  **No.** | **Experiment Description** | | | | | | | | | | | | | | | | | | | | |
| 1 | Load the specified dataset a) Using numpy and compute mean, median, variance and Standard deviation and illustrate Indexing, Slicing, Splitting, Iterating, Filtering, Sorting, Combining, and Reshaping b) Using pandas and compute mean, median, variance and Standard deviation and illustrate Indexing, Slicing, Iterating, Filtering, Sorting and Reshaping | | | | | | | | | | | | | | | | | | | | |
| 2 | For a given set of training data examples stored in a .CSV file, implement and demonstrate the (Note: Import Matplotlib)  a) Visualizing through a Line Plot  b) Creating a Bar Plot  c) Creating a Stacked Bar Plot to Visualize a specified parameter  d) Comparing specific parameters using Stacked Area Chart | | | | | | | | | | | | | | | | | | | | |
| 3 | For a given set of training data examples stored in a .CSV file, implement and demonstrate the (Note: Import Matplotlib)  a)Histogram and a Box Plot to Visualize the given parameter  b) Scatter Plot to Visualize Correlation  c) Scatter Plot with Marginal Histograms d) Plotting Multiple Images in a Grid. | | | | | | | | | | | | | | | | | | | | |
| 4 | Implement the non-parametric Locally Weighted Regression algorithm in order to fit data points. Select appropriate data set for your experiment and draw graphs | | | | | | | | | | | | | | | | | | | | |
| 5 | Write a program to parse HTML using Beautifulsoup /equivalent Library support. | | | | | | | | | | | | | | | | | | | | |
| 6 | For a given set of training data examples stored in a .CSV file, implement and demonstrate the (Note: Import Seaborn)   1. a) To Generate surface temperature analysis 2. b) Using heat map to find patterns. | | | | | | | | | | | | | | | | | | | | |
| 7 | Using Folium in Python, create a choropleth map that displays the population density of different Indian states. Color the states based on population density and add a legend. | | | | | | | | | | | | | | | | | | | | |
| 8 | Create a Bokeh plot to visualize the mean car prices of different manufacturers. | | | | | | | | | | | | | | | | | | | | |
| 9 | Create a simple application that prompts the user for a search string, utilizes the Google geocoding API for geolocation data retrieval, parses and extracts relevant information from the JSON response, and displays or utilizes this information in a useful way for the user. | | | | | | | | | | | | | | | | | | | | |
| **COURSE OUTCOMES** | | | | | | | | | | | | | | | | | | | | | |
| Upon completion of this course, the students will be able to: | | | | | | | | | | | | | | | | | | | | | |
| **CO**  **No.** | **Course Outcome Description** | | | | | | | | | | | | | | | | | | | **Bloom’s Taxonomy Level** | |
| CO1 | Demonstrate proficiency in numeric computation using Numpy and structured data handling with Pandas. | | | | | | | | | | | | | | | | | | | CL3 | |
| CO2 | Demonstrate a variety of visualizations using Matplotlib and Seaborn, including line plots, bar plots, histograms, and scatter plots. | | | | | | | | | | | | | | | | | | | CL3 | |
| CO3 | Implement the Locally Weighted Regression algorithm, selecting an appropriate dataset and drawing graphs to visualize the fitted data points. | | | | | | | | | | | | | | | | | | | CL3 | |
| CO4 | Demonstrate Proficiency in Web Scraping and Data Extraction. | | | | | | | | | | | | | | | | | | | CL3 | |
| CO5 | Design geospatial analyses using Folium, generating surface temperature analysis and creating choropleth maps. | | | | | | | | | | | | | | | | | | | CL3 | |
| **CO-PO-PSO MAPPING** | | | | | | | | | | | | | | | | | | | | | |
| **CO**  **No.** | **Programme Outcomes (PO)** | | | | | | | | | | | | | | | | | | | **Programme Specific Outcome**  **(PSO)** | |
| **1** | **2** | | **3** | **4** | **5** | **6** | | | **7** | **8** | | **9** | | **10** | | **11** | **12** | | **1** | **2** |
| **CO1** | 3 | 3 | | 2 |  | 3 |  | | |  | 2 | | 3 | | 3 | | 2 | 3 | |  |  |
| **CO2** | 3 | 3 | | 2 |  | 3 |  | | |  | 2 | | 3 | | 3 | | 2 | 3 | |  |  |
| **CO3** | 3 | 3 | | 2 |  | 3 |  | | |  | 2 | | 3 | | 3 | | 2 | 3 | |  |  |
| **CO4** | 3 | 3 | | 2 |  | 3 |  | | |  | 2 | | 3 | | 3 | | 2 | 3 | |  |  |
| **CO5** | 3 | 3 | | 2 |  | 3 |  | | |  | 2 | | 3 | | 3 | | 2 | 3 | |  |  |
| **3: Substantial (High)** | | | | | | | | **2: Moderate (Medium)** | | | | | | | | **1: Poor (Low)** | | | | | |
| **Assessment Strategy:**   * Assessment will be both CIA and SEE. * All laboratory experiments should be included for practical examination. * Students are allowed to pick one experiment from the lot. * Lab test should be conducted for 25 Marks. * Marks Distribution: Procedure + Conduction + Viva = 05 + 15 + 05 = 25 Marks. * Change of experiment is allowed only once and marks allotted to the procedure part to be made zero. | | | | | | | | | | | | | | | | | | | | | |
| **Sl. No.** | | | **Assessment Description** | | | | | | | | | **Weightage (%)** | | | | | | | | **Max. Marks** | |
| **1** | | | **Continuous Internal Assessment (CIA)** | | | | | | | | | **100 %** | | | | | | | | **50** | |
| Weekly Assessment | | | | | | | | | 50 % | | | | | | | | 25 | |
| Lab Test | | | | | | | | | 50 % | | | | | | | | 25 | |
| **2** | | | **Semester End Examination (SEE)** | | | | | | | | | **100 %** | | | | | | | | **50** | |
| **Learning Resources:**   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 | | | | | | | | | | | | | | | | | | | | | |