Savitribai Phule Pune University
T.Y.B.Sc. (Computer Science) Sem – V
Course Type: DSEC - II **Course Code: CS - 358**
Course Title : Practical Course based on CS - 353 and CS - 354

Teaching Scheme:	No. of Credits:	Examination Scheme:
5 Lect/ week Batch Size : 12	2	IE : 15 marks UE: 35 marks

Course Objectives:

- To Design dynamic and interactive Web pages.
- To Learn Core-PHP, Server Side Scripting Language
- To Learn PHP- Database handling
- To apply statistical, data preprocessing and visualization techniques on data sets

Course Outcomes:

- Understand how to develop dynamic and interactive Web Page
- Prepare data for use with a variety of statistical methods and recognize how the quality of the data may affect conclusions.
- Perform exploratory data analysis

Guidelines:

Operating Environment for web technologies: HTML5.0, PHP 5.0 and above , Webserver
Operating Environment for Data Science: Linux + python

List of Assignments on web technologies:

1 : HTML and HTML5.0
2 : CSS, Box Model, Navigation Bar
3 : Bootstrap
4 : Function and String
5 : Arrays
6: Files
7: Databases (PHP-PostgreSQL)

Suggested Assignments for Foundations of Data Science

Assignment 1: The Data Science environment
Getting introduced to Python IDLE, command line, online tools like google colaboratory and essential packages like NumPy, SciPy, pandas, scikit-learn, matplotlib, jupyter, beautiful-soup, etc.

Assignment 2: Loading the dataset
Select a dataset from a list of publicly available datasets at UCI Machine Learning Repository and load it using Pandas. (Import different dataformat files like .CSV,.htm,.json etc. Briefly describe what the dataset is about and size of the dataset (e.g. number of tables, number of instances and attributes, etc.)

Assignment 3: Basic statistical operations
Apply basic statistical operations on a dataset. For example - compute the mean, median, mode, range, quartiles, and variance for one or more attributes.

Assignment 4: Data preprocessing
Apply data preprocessing techniques that are likely required for the dataset.
1) Partition them into appropriate number of bins by equal-frequency as well as equal-width partitioning.
2) Use smoothing by bin means to smooth the data based on the above partitioning,
3)Normalize the attribute based on min-max normalization and z-score normalization.
Comment on which method you would prefer to use for partitioning, smoothing, and normalization for the given attribute.

Assignment 5: Data Visualization with matplotlib
View the data using various 2-D, 3-D plots and charts, setting styles, saving the figures, customizing the legends, multiple subplots,