



**(Established under section 3 of the UGC Act 1956, by notification
No.F.9-12/2001-U3 Government of India)**

Accredited by NAAC with 'A' Grade

Founder: Prof. Dr. S. B. Mujumdar, M.Sc.,Ph.D. (*Awarded Padma Bhushan and
Padma Shri by President of India*)

**Sub Committee Engineering – Computer Science/IT for Curriculum
Development**

Course Title: Skill Development Lab

Course Code:

Number of Credits : 4

Level : 4

Learning Outcome (s):

The students are able to:

1. Evaluate problems and analyze data using current technologies in a wide variety of business and organizational contexts.
2. Create data-driven web applications
3. Incorporate best practices for building applications
4. Employ Integrated Development Environment(IDE) for implementing and testing of software solution
5. Construct software solutions by evaluating alternate architectural patterns.

Pre-learning:

Knowledge of C, C++ and Basic JAVA

For this laboratory total five Skill Development Modules are provided as below:

SD Module-II:	PYTHON and DATA Science with R
SD Module-V:	Python and NO SQL Data Stores

Course Outline

Course Contents		
SD Module-II	PYTHON and DATA Science with R	
	Theory Content for Lab	
<p>PYTHON:</p> <p>Python Basics : Data types, Statements and Expressions, Operators and Math's, Conditionals, Loops, Strings, List , Tuples , Set Operation, Dictionary (Dict), Date and Times.</p> <p>Functions, Packages and Classes: Lambda function, Regular expression, Packages, Files, Exception Handling, Classes ,Objects, Method ,class and instance variable, constructor, destructor, inheritance.</p> <p>Numpy and Matplotlib :Array operations, Numpy Side Effects, 2D Numpy Arrays , Numpy Basic Statistics, Universal Function, Matplotlib: Introduction, Simple plots, Line API, Legend API, Figures, Subplots, Axes and Ticks.</p> <p>Pandas: Look Ups, Selections and Indexing, Filling Methods, Series operation, Handling NaN values, Mapping, DataFrames, Reading Files, Plotting,Joins, Correlation, Histograms, Rolling calculation, DateTime indexing, Grouping, Aggregate Functions, pandas.IO.Data, Panel.</p>		
<p>DATA SCIENCE WITH R:</p> <p>Introduction to Data Science- What is Data Science? Current landscape of perspectives, Skill sets needed, The Data Science Process life cycle, Role of Data Scientist. Data pre-processing. ETL – extract, transform, and load.</p> <p>Introduction to R-What is R? Installation of R. Basic features of R. R Objects. Creating Vectors and Matrices. Getting Data in and out of R. Using different packages related to data science. Managing Data frames and Functions.</p> <p>Descriptive Statistics using R - Discrete and continuous random variables, densities and distributions .Data Summarization: Measures of Central Tendency, Measures of Dispersion (quartiles, five number summary, variance, standard deviation), Measures of shape (skewness, kurtosis), Measures of association (covariance, correlation), Outliers. Using R for descriptive statistics and data visualization using ggplot2 package.</p> <p>Predictive Analysis using Machine Learning Techniques using R: Machine learning - what, how, where. Supervised, unsupervised and semi-supervised learning. Training, validation, testing, generalization, over fitting. Building a Regression model using R.Features and feature engineering. Using Decision trees, Linear classifiers, Naïve Bayes, Nearest neighbor methods in R packages.</p>		
Books:		

Text:

1. Learn PYTHON The Hard Way, Zed A. Shaw, Pearson, ISBN: 978-93-325-8210-1, Third Edition, 2017
2. Fundamentals of PYTHON, Kenneth A Lambert and B L Juneja, CENGAGE Learning, ISBN:978-81-315-2903-4
3. Peng, Roger D., and Elizabeth Matsui. "The Art of Data Science." A Guide for Anyone Who Works with Data. Skybrude Consulting 200 (2015): 162.
4. Evans, James R., and Carl H. Lindner. "Business analytics: the next frontier for decision sciences." Decision Line 43.2 (2012): 4-6.

Reference:

1. Think PYTHON, Allen B Downey, O’Rielly, ISBN: 13:978-93-5023-863-9, 4th Indian Reprint 2015
2. "Data Mining: Concepts and Techniques", Jiawei Han and Micheline Kamber, Morgan Kaufman, ISBN 978-81-312-0535-8, 2nd Edition
3. Learning R, Richard Cotton, O’Reilly, ISBN: 13:978-93-5110-286-1, First Edition, Fourth Indian Reprint 2015

Suggested List of Laboratory Assignments on PYTHON		
1.	Getting Started with Python (Example Word count exercise)	
2.	Build the Hangman Game using Python.	
3.	Write python code loads the any dataset (example Game_medal.csv), and plot the graph.	
4.	Write python code loads the any dataset (example Game_medal.csv), and does some basic data cleaning. Ask some questions on that data set.	
Suggested List of Laboratory Assignments on DATA Science with R		
1.	Getting Started with R installation, R objects and basic statistics.	
2.	Using R for data preprocessing, exploratory analysis, visualization.	
3.	Using R for correlation and regression analysis.	
4.	Data analysis case study using R for readily available data set using any one machine learning algorithm	
Suggested Mini Project on PYTHON and DATA Science with R		
<ol style="list-style-type: none"> 1. Implementing a simple Recommender System based on user buying pattern. 2. Twitter Sentiment Analysis in Python 3. Applying linear regression model to a real world problem. 		
SD Module-IV	PYTHON and NO SQL Data Stores	
	Theory Content for Lab	

PYTHON:

Python Basics : Data types, Statements and Expressions, Operators and Math's, Conditionals, Loops, Strings, List , Tuples , Set Operation, Dictionary (Dict), Date and Times.

Functions, Packages and Classes: Lambda function, Regular expression, Packages, Files, Exception Handling, Classes ,Objects, Method ,class and instance variable, constructor, destructor, inheritance.

Numpy and Matplotlib :Array operations, Numpy Side Effects, 2D Numpy Arrays , Numpy Basic Statistics, Universal Function, Matplotlib: Introduction, Simple plots, Line API, Legend API, Figures, Subplots, Axes and Ticks.

Pandas: Look Ups, Selections and Indexing, Filling Methods, Series operation, Handling NaN values, Mapping, DataFrames, Reading Files, Plotting,Joins, Correlation, Histograms, Rolling calculation, DateTime indexing, Grouping, Aggregate Functions, pandas.IO.Data, Panel.

Suggested List of Laboratory Assignments on PYTHON	
1.	Getting Started with Python (Example Word count exercise)
2.	Build the Hangman Game using Python.
3.	Write python code loads the any dataset (example Game_medal.csv), and plot the graph.
4.	Write python code loads the any dataset (example Game_medal.csv), and does some basic data cleaning. Ask some questions on that data set.

NoSQL:

1. Amazon SimpleDB :Data Model, Amazon SimpleDB API
2. MongoDB : Data Model Amazon SimpleDB API
3. Hadoop HBase

Suggested List of Laboratory Assignments on NoSQL	
1.	Implementation of Amazon SimpleDB API and CRUD Operations
2.	Implementation of Hadoop HBase API and CRUD Operations
3.	Implemenattion of MongoDB API

Books:

Text:

1. Learn PYTHON The Hard Way, Zed A. Shaw, Pearson, ISBN: 978-93-325-8210-1, Third Edition, 2017
2. Fundamentals of PYTHON, Kenneth A Lambert and B L Juneja, CENGAGE Learning, ISBN:978-81-315-2903-4
3. Peng, Roger D., and Elizabeth Matsui. "The Art of Data Science." A Guide for Anyone Who Works with Data. Skybrude Consulting 200 (2015): 162.
4. Evans, James R., and Carl H. Lindner. "Business analytics: the next frontier for decision sciences." Decision Line 43.2 (2012): 4-6.

Pedagogy

1. Demonstration in labs
2. Case Study

Suggested Assessment/ Evaluation Methods

A) Continuous Assessment

- a) Assignments
- b) Mini Project

B) End Semester Examination

- a) Lab Exam

Benchmarked against similar courses in other national/ international universities /organizations

S. No.	Name of the Course	Name of University where it is offered
1	Skill Development Lab	University of Pune

Name of Member	Dr. Shraddha Phansalkar	Dr. Swati Ahirrao
Designation	Associate Professor (CS & IT)	Associate Professor (CS & IT)
Org. / Inst.	SIT	SIT
Signature		

Name of Experts	Dr. M. L. Dhore
Designation	Professor
Org. / Inst.	VIT, Pune
Signature	

Signature of Dean:

Date: