

**Savitribai Phule Pune University  
(Formerly University of Pune)**



**Department of Technology  
Board of Studies Electronics and Electrical Technology (EE)**

**STRUCTURE OF ONE YEAR FULL TIME POST GRADUATE DIPLOMA IN  
Data Science and AI (PGD-DS&AI)**

**Semester 1**

Sr. No.	Course Code	Course Name	Teaching Scheme			Credits
			L	T	P	
1.	PGDS101	Introduction to Programming	3	1	0	4
2.	PGDS102	Statistics Essential for Data Science	2	1	1	4
3.	PGDS103	R Programming for Data Science	2	1	1	4
4.	PGDS104	Data Science with R	0	2	2	4
5.	PGDS105	Python for Data Science				4
		<b>Total Credits</b>				<b>20</b>

**Semester 2**

Sr. No.	Course Code	Course Name	Teaching Scheme			Credits
			L	T	P	
1.	PGDS201	Data Science with Python	3	1		4
2.	PGDS202	Machine Learning	3	1		4
3.	PGDS203	Tableau	3	1		4
4.	PGDS204	Natural Language Processing	3	2	2	4
5.	PGDS205	Data Science Capstone Project	3	2	2	4
		<b>Total Credits</b>				<b>20</b>

## 1 – Introduction to Programming

- Course Introduction
- Programming Basics

Reference:

- Introduction to Object Oriented Programming (English, Paperback, S. Sureshkumar)
- A Textbook of Basics of C-Language Programming (English, Paperback, K Bikshalu)
- Python Programming Fundamentals - A Beginner's Handbook (Nischay kumar Hegde)

## 2 - Statistics Essential for Data Science

- Introduction
- Sample or Population Data?
- The Fundamentals of Descriptive Statistics
- Measures of Central Tendency, Asymmetry, and Variability
- Practical Example: Descriptive Statistics
- Distributions
- Estimators and Estimates
- Confidence Intervals: Advanced Topics
- Practical Example: Inferential Statistics
- Hypothesis Testing: Introduction
- Hypothesis Testing: Let's Start Testing!
- - Practical Example: Hypothesis Testing
- The Fundamentals of Regression Analysis
- Subtleties of Regression Analysis
- Assumptions for Linear Regression Analysis
- Dealing with Categorical Data

- Practical Example: Regression Analysis

Reference:

- Statistics for Data Science (Miller James D.)
- Practical Statistics for Data Scientists ( Bruce Peter)
- Statistical Data Book 2nd Edition (R. S. Nagarazan, C. Muralidharan)

### **3 - R Programming for Data Science**

- RBasics
- Data Structures in R
- R Programming Fundamentals
- Working with Data in R
- Stings and Dates in R

Reference:

- R PROGRAMMING FOR BEGINNERS (, Sandip Rakshit)
- R for Data Science, 520 Pages - Import, Tidy, Transform, Visualize, and Model Data (Hadley Wickham, Garrett Golemund)

### **4 - Data Science with R**

- Introduction to Business Analytics
- Introduction to R Programming
- Data Structures
- Data Visualization
- Statistics for Data Science I
- Statistics for Data Science II
- Regression Analysis
- Classification
- Clustering
- Association

Reference:

- Data Analysis with R (Fischetti Tony)
- Data Science and Analytics (with Python, R and SPSS Programming) (, V.K. Jain)
- Programming with R (Sekhar S. R. Mani)

## 5 - Python for Data Science

- Python Basics
- Python Data Structures
- Python Programming Fundamentals

- Working with Data in Python

- Working with NumPy Arrays

References: Python Data Science Handbook - Essential Tools for Working with Data (Jake VanderPlas) DATA SCIENCE AND ANALYTICS WITH PYTHON (JESUS ROGEL - SALAZAR)

Mastering Python for Data Science (Madhavan Samir)

## 6 - Data Science with Python

- Data Science Overview
- Data Analytics Overview
- Statistical Analysis and Business Applications
- Python Environment Setup and Essentials
- Mathematical Computing with Python (NumPy)
- Scientific Computing with Python (Scipy)
- Data Manipulation with Pandas
- Machine Learning with Scikit-Learn
- Natural Language Processing with Scikit Learn
- Data Visualization in Python using Matplotlib
- Web Scraping with BeautifulSoup
- Python Integration with Hadoop MapReduce and Spark

Reference: 'Python Crash Course' by Eric Matthews

Big Data Analytics with Hadoop 3 (Alla Sridhar)

Scikit-Learn : Machine Learning Simplified

## **7 -Machine Learning**

- Introduction to Artificial Intelligence and Machine Learning
- Data Wrangling and Manipulation
- Supervised Learning
- Feature Engineering
- Supervised Learning Classification
- Unsupervised Learning
- Time Series Modeling
- Ensemble Learning
- Recommender Systems
- Text Mining

Reference: Machine Learning Using Python (Pradhan Manaranjan)

BASICS OF ARTIFICIAL INTELLIGENCE & MACHINE LEARNING (Dr. Dheeraj Mehrotra)

Machine Learning Fundamentals (Saleh Hyatt)

- **8 – Tableau**
- Getting Started with Tableau
- Core Tableau in Topics
- Creating Charts in Tableau
- Working with Metadata
- Filters in Tableau
- Applying Analytics to the worksheet
- Dashboard in Tableau
- Modifications to Data Connections
- Introduction to Level of Details in Tableau (LODS)

Reference: Visual Data Storytelling with Tableau, (4 Color), 1st Edition by Pearson (Lindy Ryan)

Mastering Tableau 2019.1 (Meier Marleen)

Tableau Your Data! - Fast and Easy Visual Analysis With Tableau Software (Murray Dan)

- **9 - Natural Language Processing**

- Introduction to Natural Language Processing
- Feature Engineering on Text Data
- Natural Language Understanding Techniques
- Natural Language Generation
- Natural Language Processing Libraries
- Natural Language Processing with Machine Learning and Deep Learning
- Speech Recognition Technique

Reference: Natural Language Processing and Information Retrieval First Edition ( TIWARY, U.S, SIDDIQUI ,TANVEER)

Python Natural Language Processing (Thanaki Jalaj)

## **10 - Data Science Capstone Project**

### **Projective objective:**

- **Data Processing** - In this step, you will apply various data processing techniques to make raw data meaningful.
- **Model Building** - You will leverage techniques such as regression and decision trees to build Machine Learning models that enable accurate and intelligent predictions. You may explore Python, R, or SAS to develop your model. You will follow the complete model-building exercise from data split to test and validate data using the k-fold cross-validation process.
- **Model Fine-tuning** - You will apply various techniques to improve the accuracy of your model and select the champion model that provides the best accuracy.
- **Dash boarding and Representing Results** - As the final step, you will be required to export your results into a dashboard with meaningful insights using Tableau.