

KJ'S Educational Institutes

**TRINITY ACADEMY OF ENGINEERING,
PUNE**

(Accredited with 'A+' grade by NAAC)

Department of MCA



LABORATORY MANUAL

Data Science Laboratory

(Subject Code: 410908)

For the Academic Year 2025 - 26

SYMCA- Semester III

TeachingScheme: PR:04
Hours/Week: 4

Credit 02

ExaminationScheme:
TW:25
Marks PR: 50 Marks

Course Objectives	
CO No.	Program Objectives:
CO1	To learn basics about Data Analytics Tool for Data Science

Course Outcomes	
CO1	Describe framework of any Data Analytics Tool
CO2	Write basic applications using the fundamentals of any Data Analytics Tool.
CO3	Apply Modeling techniques using any Data Analytics Tool.
CO4	Implement Mining techniques using any Data Analytics Tool
CO5	Employ data analysis using graphs.
CO6	Implement Data Visualization

LAB OBJECTIVE:

Guidelines for Student Journal:

- The laboratory assignments are to be submitted by student in the form of journal. Journal consists of prologue, Certificate, table of contents, and **handwritten write-up** of each assignment (Title, Objectives, Problem Statement, Outcomes, software & Hardware requirements, Date of Completion, Assessment grade/marks and assessor's sign, Theory-Concept in brief, algorithm, flowchart, Design, test cases, conclusion/analysis).
- **Program codes with sample output of all performed assignments are to be submitted as softcopy.** As a conscious effort and little contribution towards Green IT and environment awareness, attaching printed papers as part of write-ups and program listing to journal may be avoided. Use of DVD containing students programs maintained by lab In-charge is highly encouraged. For reference one or two journals may be maintained with program prints at Laboratory.

Content of Lab Experiments

S.No.	Title of Program
1	Installation and study of any one Data Analytics Tool Framework.
2	Write a python program to demonstrate the use of Numpy.
3	Design and develop at least 10 problem statements which demonstrate the use of data structure, functions, Importing / Exporting Data in any data analytics tool.
4	Design and develop at least 5 problem statements which demonstrate the use of Control Structures of any data analytics tool.
5	Implement KNN Classification techniques using any data analytics tool.
6	Implement Naive Bayes Classification techniques using any data analytics tool.
7	Implement K means Clustering techniques using any data analytics tool.
8	Implement DBScan Clustering techniques using any data analytics tool.
9	Implement Eclat Association Rule Mining techniques using any data analytics tool.
10	Implement Apriori Association Rule Mining techniques using any data analytics tool.
11	Visualize all the statistical measures (mean, mode, median, range, inter quartile range, etc.) using Histograms, Boxplots, scatter plots, etc.

Books:

Text Books:

1. Jeffrey S.Saltz,Jeffre M. Stanton, “An Introduction to Data Science”, Sage Publications,2018
2. Seema Acharya ,”Data Analytics using R “ , McGraw Hill, 2018
3. Cathy O’Neil and Rachel Schutt. Doing Data Science, Straight Talk From The Frontline. O'Reilly.
4. Jiawei Han, Micheline Kamber, “Data mining: concepts and techniques”, Morgan Kaufmann Publisher, second edition.
5. Jure Leskovec, Anand Rajaraman and Jeffrey Ullman. Mining of Massive Datasets. v2.1,Cambridge University Press

Reference Books:

1. Bharti Motwani, “Data Analytics with R”, Wiley 2019.
2. Hadley Wickham, “R for Data Science: Import, Tidy, Transform, Visualize, and ModelData”, First Edition,O'Reilly Media Publisher, ISBN: 9781491910399, 2017.
3. T. Hastie, R. Tibshirani, J. H. Friedman, The Elements of Statistical Learning: Data Mining, Inference,

- and Prediction. Springer, 2013.
4. Tom Mitchell, Machine Learning. McGraw-Hill, 1997.
 5. Peter Flach, Machine Learning: The Art and Science of Algorithms that Make Sense of Data. Cambridge University Press, 2012.
 6. Carl Edward Rasmussen and Christopher K. I. Williams, Gaussian Processes for Machine Learning. MIT Press, 2005.
 7. Daphne Koller and N. Friedman, Probabilistic Graphical Models: Principles and Techniques. MIT Press, 2009.
 8. Christopher Bishop, Pattern Recognition and Machine Learning. Springer, 2007.
 9. Laura Igual and Santi Segui, Introduction to Data Science: A Python Approach to Concepts, Techniques and Applications, Springer; 1st ed. 2017 edition

Websites links

1. [http:// https://towardsdatascience.com/](http://https://towardsdatascience.com/)
2. [https:// www.simplilearn.com](https://www.simplilearn.com)