DHANALAKSHMI SRINIVASAN ENGINEERING COLLEGE
(AUTONOMOUS)
(Approved by AICTE & Affiliated to Anna University, Chennai)
Re-Accredited by NAAC with 'A' Grade
Accredited by NBA for AERO, BME, CSE, ECE, EEE, IT & MECH.
PERAMBALUR-621212, TAMILNADU, INDIA.
Website: www.dsengg.ac.in

LABORATORY COURSE PLAN (2025-2026 ODD SEMESTER)

LAB COURSE TITLE	DATA SCIENCE LABORATORY			
LAB COURSE CODE	**U23AIP42**			
LAB COURSE STRUCTURE	LECTURE	TUTORIAL	PRACTICAL	CREDIT
	0	0	4	2
REGULATION	BRANCH	YEAR	SEMESTER	ACADEMIC YEAR
2020	IT	III	V	2025-2026
COURSE INCHARGE				

SYLLABUS

COURSE OBJECTIVE:

The main learning objective of this course is to prepare the students:

- To understand the techniques and processes of data science
- To apply descriptive data analytics
- To visualize data for various applications
- To understand inferential data analytics
- To analysis and build predictive models from data

LIST OF EXPERIMENTS:

1. Study of Basic function in Excel
2. Study of Basic Data Science Libraries in Python
3. Working with Range Names and Tables
4. Cleaning Data with Text Functions
5. Cleaning Data containing Data Values
6. Working with VLOOKUP functions and Pivot Table.
7. Demonstration of Data Visualization in Excel.
8. Demonstration of Data Visualization in Python.
9. Importing Data from External Source Using Excel & Python
10. Creating a data model
11. Create a dashboard for a given requirement
12. Implement a data analytics for the real time data set

TOTAL: 60 PERIODS

BIBLIOGRAPHY

TEXT/REFERENCE BOOKS:

1. Joel Grus,Explains the core concepts of data science by building tools from scratch using Python. Helps understand the "why" behind lab exercises.
2. Sebastian Raschka, hands-on book covering machine learning in Python with step-by-step lab-oriented
3. Trevor Hastie, Robert Tibshirani, Jerome Friedman The Elements of Statistical Learning

HARDWARE:

Standalone Desktop

SOFTWARE:

- Microsoft Excel
- Python 3

Web link for resource &Virtual lab reference link

https://www.digimat.in/nptel/courses/video/106105191/L01.html.
https://www.javatpoint.com/java-tutorial

EXP. NO.	NAME OF THE EXPERIMENTS	NO. OF PERIODS	CUMULATIVE PERIODS
1.	Study of Basic function in Excel	4	4
2.	Study of Basic Data Science Libraries in Python	4	8
3.	Working with Range Names and Tables	8	16
4.	Cleaning Data with Text Functions	8	24
5.	Cleaning Data containing Data Values	4	28
6.	Working with VLOOKUP functions and Pivot Table.	4	32
7.	Demonstration of Data Visualization in Excel.	4	36
8.	Demonstration of Data Visualization in Python.	4	40
9.	Importing Data from External Source Using Excel & Python	4	44
10.	Creating a data model	4	48
11.	Create a dashboard for a given requirement	8	56
12.	Implement a data analytics for the real time data set	4	60

COURSE OUTCOME

Upon Completion of the course, the students will be able to:

At the end of the course the students would be able to

CO1: Understand the real-world data and information.
CO2: Apply data science using excel & Python.
CO3: Design of mathematical model for problem solving
CO4: Interpret various tools and its advantages.
CO5: Illustrate the different opportunities in industries.
CO6: Apply data modelling for real-world applications.

CO-PO mapping:

CO	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12
CO1	3	2	2	2	2	1	-	-	2	2	2	2
CO2	3	2	2	2	2	1	-	-	2	2	2	1
CO3	3	2	2	1	2	1	-	-	2	1	2	1
CO4	3	2	2	2	2	1	-	-	2	1	2	2
CO5	2	1	1	1	1	-	-	-	2	1	2	1
CO6	1	1	1	0	2	-	-	-	1	1	1	1
AVG	2.5	1.67	1.67	1.33	1.83	0.67	-	-	1.83	1.33	1.83	1.33

ADDITIONAL EXPERIMENTS		
EXP. NO.	**NAME OF THE EXPERIMENTS**	**Identified Resource link**
1	Data Cleaning Using Real-World Dataset	https://knreddycse.weebly.com/java-lab-programs.html.
2	Perform EDA on a dataset and extract insights.	https://knreddycse.weebly.com/java
3	Build an interactive dashboard for a sales or HR dataset.	https://knreddycse.weebly.com/java
4	Create meaningful visualizations from dataset.	https://knreddycse.weebly.com/java
5	Combine data from multiple sources and perform aggregation.	https://knreddycse.weebly.com/java

MODEL LAB DETAILS

BATCH	REGISTER NO.	MODE OF LAB CONDUCT	DATE	TIMING
1	810424205127-810424205189	Offline		

LIST OF QUESTIONS

1. Write a Python program to fill missing numeric values with the mean of the column.
2. Write a code snippet to remove special characters (like $ or ,) from a price column and convert it to float.
3. Write a program to group data by a category and find the average of numeric columns.
4. Write a Python program to detect outliers using Z-score.
5. Write code to train a simple logistic regression model on a dataset.
6. Write a VLOOKUP formula to find the price of an item in Excel.

VIVA QUESTIONS

1. What is data cleaning, and why is it important in data science?
2. Explain the difference between structured and unstructured data.
3. What are missing values? How can they be handled?
4. Will the program run if we write static public void main?
5. What is a Pandas DataFrame?
6. How do you check for null values in a DataFrame?
7. What is the difference between `loc[]` and `iloc[]` in Pandas?
8. How do you merge two datasets in Python?
9. How do you detect and treat outliers in a dataset?
10. Explain the difference between normalization and standardization.
11. How do you convert a column to datetime format in Pandas?
12. What types of charts are useful for categorical data?
13. How can you visualize the relationship between two numeric variables?
14. What is a heatmap and when would you use it?
15. What is a Pivot Table and why is it useful?
16. Explain how VLOOKUP works in Excel.
17. How do slicers enhance data analysis in Excel dashboards?
18. What is the difference between supervised and unsupervised learning?
19. What is overfitting in machine learning?
20. Explain precision and recall.

PREPARED BY **VERIFIED BY**

AP/IT HOD/IT

APPROVED BY
PRINCIPAL