HINDUSTHAN COLLEGE OF ENGINEERING AND TECHNOLOGY
An Autonomous Institution, Approved by AICTE, New Delhi, Affiliated to Anna University, Chennai
Accredited by NBA (AERO, AUTO, CIVIL, CSE, ECE, EEE, IT, MECH & MCT)
Accredited with 'A++' Grade by NAAC
Valley Campus, Pollachi Highway, Coimbatore - 641 032

DEPARTMENT OF COMPUTER APPLICATIONS

PROGRAMMING LANGUAGE LABORATORY

DETAILS OF HARDWARE

Computer Configuration	No. of Computers
HP i5 processor CPU/G2 280 MT/500GB HDD/8GB RAM	30

DETAILS OF SOFTWARE

| Category	Software	Description
Programming Languages	Python, R	Primary programming languages for data analysis, statistical computing, and machine learning.
Integrated Development Environments (IDEs)	Jupyter Notebook, JupyterLab, RStudio	IDEs for interactive computing, data analysis, and visualization.
Data Manipulation	Pandas, dplyr	Libraries for data manipulation and transformation in Python and R, respectively.
Statistical Analysis	SciPy, StatsModels, R (base functions)	Libraries for conducting statistical analysis and hypothesis testing.
Machine Learning	scikit-learn, TensorFlow, Keras, PyTorch	Libraries for building and deploying machine learning models.
Data Visualization	Matplotlib, Seaborn, Plotly, ggplot2, Tableau, Power BI	Tools and libraries for creating static, interactive, and complex data visualizations.
Big Data Processing	Apache Hadoop, Apache Spark	Frameworks for processing and analyzing large datasets.
Database Management	MySQL, PostgreSQL, MongoDB	Database management systems for storing, querying, and managing data.
Data Cleaning	OpenRefine, Trifacta	Tools for cleaning and transforming raw data.

List of Experiments

Programme	Course Code	Name of the Course	L	T	P	C
MCA	20CA2204	DATA SCIENCE WITH PYTHON PROGRAMMING	3	1	0	4

Course Objective	1. Understand the fundamental concepts of data science 2. Understand the theoretical properties of methods involved in Data Science. 3. knowledge of statistical data analysis techniques 4. Understand various machine learning algorithms used in data science process 5. Use concepts and methods of mathematical disciplines relevant to data Science

Unit	Description	Instructional hours
I	**DATA SCIENCE** DS Introduction - DS What is Data - DS DataBase Table - Ds Python - DS Data Frame - DS Functions - DS DataPreparation	12
II	**DS MATH** DS Linear Functions - DS Plotting Functions - DS Slop & Intercept	12
III	**DS STATISTICS** Stat Introduction - Stat Percentiles - Stat Deviation - Stat Variance - Stat Correlation - Stat Correlation Matrix - Stat Correlation VS Casuality	12
IV	**DS COMPONENTS** DS Statistics - DS Visualization - DS Machine Learning - DS Deep Learning	12
V	**DS ADVANCED** DS Linear Regression - DS Regression Table - DS Regression Info - DS Regression coefficients - DS Regression P-Value - DS Regression R-Squared - DS Linear Regression Case	12
	Total Instructional hours	**60**

Course Outcome	CO1: Acquire fundamental concepts of data science. CO2: Able to Understand the theoretical properties of methods involved in Data Science. CO3: Apply the acquired knowledge of statistical data analysis techniques CO4: Apply the acquired knowledge of machine learning algorithms used in data science process CO5: Apply the concepts and methods of mathematical disciplines relevant in Data Science

R1. Machine Learning for Absolute Beginners, January 1, 2018, Scatterplus.com, Oliver Theobald R2. Introduction to Statistics, Jim Frost, Jim Publishing, 2018

R3. William McKinney Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Ipython, O''Reilly Media, Inc., 2017

R4. ManasaKalaimalai, The Essential Beginner's Guide to Data Science, Notion Press,2020.
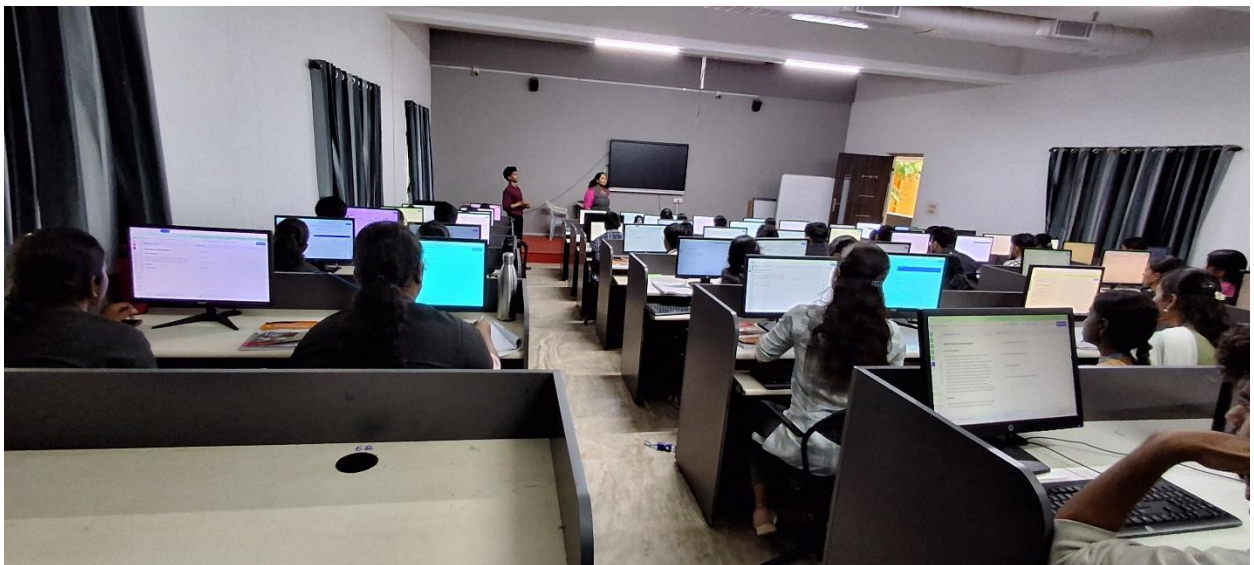
Programming Language Lab & Data Science and Visualization Lab