KONGUNADU ARTS AND SCIENCE COLLEGE
(AUTONOMOUS)

COIMBATORE – 641 029

CURRICULUM AND SCHEME OF EXAMINATIONS (CBCS)
(2026-2027 and onwards)

for the programme

PG DIPLOMA IN DATA ANALYTICS

Offered by

DEPARTMENT OF MATHEMATICS

PROGRAMME OUTCOME (PO)

PO 1: To bridge the gap between the industry and the technology.

PO 2: To visualize the real-world challenges and interpret their complexity with the help of data science.

PO 3: To enhance the economic growth of the country by implementing various data analytics techniques.

PROGRAMME SPECIFIC OUTCOME (PSO)

PSO 1: Recalling the various types of data and effectively make use of them to analyze the progress of any sector.

PSO 2: Interpretation of the data to overcome the threats and challenges that arise time to time.

PSO 3: Application of various statistical tools to measure the efficiency, investigate the quality and to predict the stability of any product in the Marketing arena.

PSO 4: Inferring the usage of appropriate analytics techniques to optimize the Profit Loss ratio.

PSO 5: Implementation of Machine Learning and Artificial Intelligence to reduce the man power and time consumption while dealing with complex real world problems.

KONGUNADU ARTS AND SCIENCE COLLEGE (AUTONOMOUS)

COIMBATORE – 641 029

Programme Name : PG Diploma in Data Analytics

Curriculum and Scheme of Examination under CBCS

(Applicable to the students admitted during the Academic Year 2026-2027 and onwards)

Semester	Subject Code	Title of the Paper	Instruction hours/cycle	Exam. Marks			Duration of Exam (Hours)	Credits
				CIA	ESE	TOTAL		
I	26PDD101	Core Paper 1 :Big Data and Data Analytics	2	25	75	100	3	2
	26PDD102	Core Paper 2 :RDBMS and SQL	2	25	75	100	3	2
	26PDD103	Core Paper 3 :Python Programming	2	25	75	100	3	2
	26PDD1CL	Core Practical 1: Python Programming – Lab	2	40	60	100	3	2
		Total	**8**			**400**		**8**
II	26PDD204	Core Paper 4 :Machine Learning and Programming Based on Application Implementation	2	25	75	100	3	2
	26PDD2CL	Core Practical 2 :R Programming- Lab	2	40	60	100	3	2
	26PDD2Z1	Project – Viva Voce	4	40	160	200	-	4
		Total	**8**			**400**		**8**
		Grand Total	**16**			**800**		**16**

Note:

CBCS – Choice Based Credit System
CIA – Continuous Internal Assessment
ESE – End of Semester Examinations

Components of Continuous Internal Assessment (50 Marks)

Components		**Marks**	**Total**
Theory			
CIA I	75	(75+75) converted to 15	25
CIA II	75		
Attendance		5	
Others*****		5	
Practical			
CIA Practical		25	40
Observation Notebook		10	
Attendance		5	
Project			
Review		30	40
Regularity		10	

* Class Participation, Case Studies Presentation, Field Work, Field Survey, Group Discussion, Term Paper, Workshop/Conference Participation. Presentation of Papers in Conferences, Quiz, Report/Content writing. Etc.

****** Two Assignments to be given. (Each 5 marks).

BLOOM'S TAXONOMY BASED ASSESSMENT PATTERN

K1-Remembering; **K2**-Understanding; **K3**-Applying; **K4**-Analyzing; **K5**-Evaluating

1. Theory Examination:

(i) CIA I & II and ESE: 75 Marks

Knowledge Level	Section	Marks	Description	Total
K1 Q1 to 10	A (Answer all)	10 x 1 = 10	MCQ	75
K1 – K5 Q11 to 15	B (Either or pattern)	5 x 5 = 25	Short Answers	
K2 – K5 Q16 to 20	C (Either or pattern)	5 x 8 = 40	Descriptive / Detailed	

2. ESE Practical Examination:

Knowledge Level	Section	Marks	Total
K3	Experiments	50	60
K4	Record Work	10	
K5			

3. ESE Project Viva Voce:

<table>
<tr><th>Knowledge
Level</th><th>Section</th><th>Marks</th><th>Total</th></tr>
<tr><td>K3</td><td rowspan="3">Project Report

Viva voce</td><td>120</td><td rowspan="3">160</td></tr>
<tr><td>K4</td><td rowspan="2">40</td></tr>
<tr><td>K5</td></tr>
</table>

26PDD101

Programme Code:02		PG Diploma in Data Analytics		
Title of the Paper		**Core Paper 1 : Big Data and Data Analytics**		
Batch 2026-2027	**Hours / Week** 2	**Total Hours** 30	**Credits** 2	Employability/ Skill Development

Course Objectives

1.To understand and apply scaling up machine learning techniques and associated computing techniques and technologies.
2.To identify the characteristics of datasets and compare the trivial data and big data for various applications.
3.To recognize and implement various ways of selecting suitable model parameters for different machine learning techniques.

Course Outcomes (CO)

K1 to K5	CO1	Understand the different dimensions of digital data.
	CO2	Apply the concept of data classification on different types of data.
	CO3	Analyze the characteristics of different patterns of data.
	CO4	Implement the concept of big data in different scenarios.
	CO5	Compare the various types of data

Unit I **(6 Hours)**

Data Evolution: Data Development Time Line – ICT Advancement - a Perspective – Data Growth - a Perspective – IT Components - Business Process – Landscape - Data to Data Science – Understanding data: Introduction – Type of Data: Numeric – Categorical – Graphical– High Dimensional Data — Data Classification – Hot Data – Cold Data – Warm Data – Thick Data – Thin Data - Classification of digital Data: Structured, Semi-Structured and Un-Structured.

Unit II **(6 Hours)**

Sources of Data: Time Series – Transactional Data – Biological Data – Spatial Data – Social Network Data – Data Evolution – Data Sources.
Data Science: Data Science-A Discipline – Data Science Vs Statistics – Mathematics - Programming Language - Database, - Machine Learning. Data Analytics Relation: Data Science, Analytics, Big Data Analytics.

Unit III **(6 Hours)**

Data Science Components: Data Engineering, Data Analytics-Methods and Algorithm, Data Visualization Big Data: Introduction to Big Data: Evolution - What is Big Data – Sources of Big Data. Characteristics of Big Data 6Vs – Big data- Challenges of Conventional Systems.

Unit IV **(6 Hours)**

Data Processing Models – Limitation of Conventional Data Processing Approaches – Big Data Myths - Data Discovery-Traditional Approach, Big Data Technology: Big Data Exploration - Data Augmentation – Operational Analysis – 360 View of Customers – Security and Intelligence.

Unit V **(6 Hours)**

Big Data Use cases – Big Data Technology Potentials – Limitations of Big Data and Challenges - Big Data Roles Data Scientist, Data Architect, Data Analyst – Skills – Case Study: Big Data – Customer Insights – Behavioral Analysis – Big Data Applications - Marketing – Retails – Insurance – Risk and Security – Health care.

Text Book:

1. V. Bhuvaneswari, T. Devi, (2016), Big Data Analytics: A Practitioner's Approach.

Reference Books:

1. Han Hu, Yonggang Wen, Tat - Seng, Chua, XuelongLi, (2015), Toward Scalable Systems for Big, SeemaAcharya, SubhashniChellappan, Big Data Analytics, Wiley.
2. Paul C. Zikopoulos, Chris Eaton, Dirk deRoos, Thomas Deutsch, George Lapis, (2012), Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data, McGraw Hill

e-Resources

1. https://www.geeksforgeeks.org/big-data/what-is-big-data/
2. https://www.geeksforgeeks.org/data-science/data-science/
3. https://www.youtube.com/watch?v=ePeWwDFhTLs
4. https://www.youtube.com/watch?v=Q0rbO8H2Zv0
5. https://www.geeksforgeeks.org/big-data/applications-of-big-data/

26PDD101

Mapping

PSO / CO	PSO1	PSO2	PSO3	PSO4	PSO5
CO1	S	S	H	M	S
CO2	H	S	S	H	S
CO3	H	H	S	M	H
CO4	M	H	H	M	S
CO5	H	M	S	H	S

S - Strong; **H**-High; **M**-Medium; **L**-Low

Content prepared by (Based on Inputs) Signature of the Course in-charge	 Signature of the HOD

26PDD102

Programme Code:02		PG Diploma in Data Analytics		
Title of the Paper		**Core Paper 2 : RDBMS and SQL**		
Batch 2026-2027	**Hours / Week** 2	**Total Hours** 30	**Credits** 2	Employability/ Skill Development

Course Objectives

1. To develop the knowledge in various Database concepts, queries, normalization and reports.
2. To be able to construct a new normalized database.

Course Outcomes (CO)

K1 to K5	CO1	Remember the basic concepts of database management systems and database techniques.
	CO2	Understand Data constraints and CODDs rules, DML and DDL statements of ORACLE
	CO3	Apply various DDL and DML statements, joins queries, PL / SQL statements.
	CO4	Analyze the granting and revoking permissions, cursors
	CO5	Explain BCNF.

Unit I **(6 Hours)**

Introduction: Purpose of Database Systems - View of Data - Data Models - Database Languages - Database Administrator - Database Users. Entity Relationship Model: Basic concepts –Mapping Cardinalities - Entity Relationship Diagram- E-R Features - Relational Model: Structure of Relational Databases - Relational Algebra.

Unit II **(6 Hours)**

Interactive SQL : Invoking SQL * Plus- data definition- data manipulation in DBMS – The oracle data types –DML and DDL statements-Data constraints- arithmetic, logical operators- oracle functions- grouping data from tables -manipulating dates- union, intersect and minus clause-Granting permissions- Revoking permissions- Codd’s Rules.

Unit III **(6 Hours)**

PL/SQL: Introduction, PL/SQL syntax, understanding PL/SQL block structure, oracle transactions, cursors, stored procedures, stored functions, database triggers – **Creating Default Tabular Report***.

Unit IV **(6 Hours)**

SQL: Nested Sub queries - Derived Relations – **Views** * - Joined Relations. Integrity Constraints: Domain Constraints- Referential Integrity - Assertions.

Unit V **(6 Hours)**

Functional Dependencies - Relational Database Design: Pitfalls – Normalization-First Normal Form, Second Normal Form, Third Normal Form and BCNF.

Text Book:

1. Abraham Silberschatz, Henry F. Korth, S. Sudharshan, (2006), Database System Concepts, Fifth Edition, Tata McGraw Hill Publication. (Unit I, IV and V)
2. Ivan Bayross, (2007), Commercial application development using ORACLE developer 2000, First Edition. (Unit II and III)

Reference Books:

1. Bipin.C.Desai, (2000), An Introduction to database systems, First Edition, Galgotia Publication.
2. Ivan Bay Ross(1995), Oracle 7 The Complete Reference, First Edition, BPB Publications, Chennai.

e-Resources

1. https://www.geeksforgeeks.org/dbms/introduction-of-dbms-database-management-system-set-1/
2. https://www.w3schools.com/sql/
3. https://www.tutorialspoint.com/plsql/index.htm
4. https://www.youtube.com/playlist?list=PLBlnK6fEyqRiyryTrbKHX1Sh9luYI0dhX
5. https://www.geeksforgeeks.org/dbms/boyce-codd-normal-form-bcnf/

Mapping

PSO / CO	PSO1	PSO2	PSO3	PSO4	PSO5
CO1	H	S	H	M	S
CO2	H	H	S	H	S
CO3	S	H	S	M	H
CO4	M	S	H	M	S
CO5	S	H	S	M	S

S - Strong; **H**-High; **M**-Medium; **L**-Low

Content prepared by (Based on Inputs) Signature of the Course in-charge	 Signature of the HOD

26PDD103

Programme Code:02		PGDiploma in Data Analytics		
Title of the Paper		**Core Paper 3 : Python Programming**		
Batch	**Hours / Week**	**Total Hours**	**Credits**	Employability/ Skill Development
2026-2027	2	30	2	

Course Objectives

1. To introduce the fundamentals of Python Programming.
2. To teach about the concept of Functions in Python.
3. To impart the knowledge of Lists, Tuples, Files and Directories.

Course Outcomes (CO)

K1 – K5	CO1	Remembering the concept of operators, data types, Loops and control statements in Python programming.
	CO2	Understanding the concepts of Input / Output operations in file.
	CO3	Applying the concept of functions and exception handling
	CO4	Analyzing the structures of list, tuples and maintaining dictionaries.
	CO5	Compare mutable and immutable data types.

Syllabus

Unit I **(6 Hours)**

Introduction to Python**:** Introduction - Python Overview - Getting Started with Python – Comments - Python Identifiers - Reserved Keywords – Variables - Standard Data Types. Operators - Statement and Expression - String Operations - Boolean Expressions - Control Statements - Iteration - While Statement - Input from Keyboard.

Unit II **(6 Hours)**

Problem solving strategies: Problem Analysis – Algorithms – Flow Charts – ***Examples of Algorithms and Flow Charts.**

Functions: Built-in functions - Composition of functions - User defined functions - Parameters and Arguments - Function Calls - The return Statement - Python Recursive Function - The Anonymous Functions – Writing Python Scripts.

Unit III **(6 Hours)**

Strings**:** Compound Data Type - Len Function - String Slices - Strings are Immutable - String Traversal - Escape Characters - String Formatting Operator - String Formatting Functions.

Lists: Values and Accessing Elements - Lists are Mutable - Deleting Elements from List- Built-in list Operators - Built-in List Methods.

26PDD103

Unit IV **(6 Hours)**

Tuples**:** Creating Tuples - Accessing values in Tuples - Tuples are Immutable - Tuple Assignment - Tuples as Return Values - Variable Length Argument Tuples - Basic Tuple Operations - Built-in Tuple Functions.
Dictionaries: Creating a Dictionary - Accessing values in a Dictionary - Updating Dictionary - Deleting elements from Dictionary - Properties of Dictionary keys - Operations in Dictionary - Built-in Dictionary methods.

Unit V **(6 Hours)**

Classes and Objects: Overview of OOP - Class Definition - Creating Objects - Objects as arguments - Objects as Return Values - Built-in class attribute - Inheritance - Method Overriding - Data Encapsulation - Data Hiding.

* Self Study and questions for examinations may be taken from the self study portions also.

Text Book:

1. E. Balagurusamy,Introduction to Computing andProblem Solving using Python Programming, First Edition, 2016, McGraw-Hill Education (India) Pvt. Ltd, Chennai.

Reference Books:

1. Ashok NamdevKamthane, Amit Ashok Kamthane, Programming and Problem Solving with Python, McGraw-Hill, First Edition, 2017.
2. Martin Jones,Python for Complete Beginners, Createspace Independent Publisher, First Edition, 2015.
3. S.A. Kulkarni, Problem Solving and Python Programming, Yes Dee Publishing Pvt. Ltd, First Edition, 2017.

e-Resources

1. https://www.w3schools.com/r/r_data_structures.asp
2. https://www.geeksforgeeks.org/machine-learning/machine-learning/
3. https://www.geeksforgeeks.org/machine-learning/decision-tree-introduction-example/
4. https://www.appliedaicourse.com/blog/computational-learning-theory-in-machine-learning/
5. https://www.geeksforgeeks.org/machine-learning/introduction-to-recurrent-neural-network/

26PDD103

Mapping

PSO / CO	PSO1	PSO 2	PSO 3	PSO 4	PSO 5
CO1	S	H	M	S	S
CO2	S	S	H	H	H
CO3	H	M	S	S	H
CO4	M	S	M	H	S
CO5	S	H	M	S	M

S – Strong **H** – High **M** – Medium **L** – Low

Content prepared by (Based on Inputs) Signature of the Course in-charge	 Signature of the HOD

26PDD1CL

Programme Code:02		PGDiploma in Data Analytics		
Title of the Paper		**Title :Core Practical 1: Python Programming- Lab**		
Batch 2026-2027	**Hours / Week** 2	**Total Hours** 30	**Credits** 2	Employability/ Skill Development

Course Objectives

1. To gain knowledge about the concepts of Python programming.
2. To implement various operators in python programming in problem solving.
3. To enhance the students to develop the program writing skills for computational problems.

Course Outcomes (CO)

K3- K5	CO1	Finding the GCD of two integers using Python program.
	CO2	Demonstration of classes and their attributes.
	CO3	Utilizing Python program to solve complex mathematical problems.
	CO4	Analyzing the GCD and LCM of integers using Python programs.
	CO5	Applying, compiling and debugging programs with the help of Python.

LIST OF PRACTICAL PROGRAMS

1. Program to determine the Greatest Common Divisor (GCD) of any two integers.
2. Program to find the Least Common Multiple (LCM) of any two integers using functions.
3. Program to find the number of instances of different digits in a given number.
4. Program to find the number of vowels and consonants in a text string.
5. Program to get a list of words and sort them in Alphabetical Order.
6. Program to find permutation of all characters in a string.
7. Program to compute the Average of Best of Two of Three Assignment Tests. Maximum Marks for each Assignment is 25. Fractional Marks in Final Average are Rounded off to the Nearest and Highest Whole Number.
8. Program to sort an integer list using Bubble sorting technique.
9. Program to Draw a Histogram using Turtle Graphics
10. Program to demonstrate Classes and their Attributes.

26PDD1CL

Distribution of Marks in ESE			**CIA**		
Experiment	:	50	CIA Practical	:	25
Record	:	10	Exam		
Total		**60**	Attendance	:	5
			Observation Note book	:	10
			Total		**40**

To be awarded jointly by the internal and external examiners.

Mapping

PSO / CO	PSO1	PSO2	PSO3	PSO4	PSO5
CO1	S	S	H	H	S
CO2	M	M	H	H	M
CO3	H	S	S	S	M
CO4	S	H	S	M	S
CO5	H	S	S	H	M

S - Strong; **H**-High; **M**-Medium; **L**-Low

Content prepared by (Based on Inputs) Signature of the Course in-charge	 Signature of the HOD

26PDD204

Programme Code:02		PGDiploma in Data Analytics		
Title of the Paper		**Core Paper 4 :Machine Learning and Programming Based on Application Implementation**		
Batch 2026-2027	**Hours / Week** 2	**Total Hours** 30	**Credits** 2	Employability/ Skill Development

Course Objectives

1. To understand the concepts of Machine Learning.
2. To implement various programming applications.
3. To know more about Neural networks.

Course Outcomes (CO)

K1 to K5	CO1	Remembering the basic concepts of R Data Structures.
	CO2	Applying the notions of Machine Learning and Computational Learning.
	CO3	Exerting the fundamental concepts of Artificial Neural Networks.
	CO4	Evaluating the concept of Tree and Probabilistic Model.
	CO5	Investigating the problem solving approach of artificial intelligence.

Unit I **(6 Hours)**

Introducing to R: Introducing to R – R Data Structures – Help Functions in R – Vectors – Scalars – Declarations – Recycling – Common Vector Operations

Unit II **(6 Hours)**

Introduction to Learning: ML Fundamentals - Algorithmic models of learning, Learning classifiers, functions, relations, grammars, value functions.
ML- Models: Parameter Estimation, sufficient statistics, decision trees, neural networks, support vector machines, Bayesian networks.

Unit III **(6 Hours)**

Tree & Probabilistic Model: Tree and Probabilistic Models – Learning with Trees – Decision Trees – Constructing Decision Trees, Probability and Learning – Data into Probabilities – Basic Statistics.

Unit IV **(6 Hours)**

Computational Learning: Computational Learning theory, mistake bound analysis, sample complexity analysis, VC dimension, Occam learning, accuracy and confidence boosting.
Unsupervised Learning: Unsupervised Learning: Clustering, mixture models, k-means clustering, hierarchical clustering, and distributional clustering, Reinforcement learning.

Unit V **(6 Hours)**

Introduction to Artificial Neural Networks: Introduction- History of neural networks- multilayer perceptions- -What are RNNs – Introduction to RNNs.

Artificial Intelligence and Knowledge Representation: Introduction – Definition – Future of Artificial Intelligence – Characteristics of Intelligent Agents – Typical Intelligent Agents – Problem Solving Approach.

Text Book

1. Stuart J. Russell and Peter Norvig, Artificial Intelligence: A Modern Approach, PearonPublisher, 2018.

2. Brett Lantz, Machine Learning with R - Second Edition,Packt Publisher, 2015.

Mapping

PSO / CO	PSO1	PSO2	PSO3	PSO4	PSO5
CO1	S	H	S	M	S
CO2	H	S	S	H	S
CO3	H	S	H	M	H
CO4	M	H	H	M	S
CO5	S	M	H	M	S

S - Strong; **H**-High; **M**-Medium; **L**-Low

Content prepared by (Based on Inputs) Signature of the Course in-charge	 Signature of the HOD

26PDD2CL

Programme Code:02		PG Diploma in Data Analytics		
Title of the Paper		**Core Practical 2: R Programming – Lab**		
Batch 2026-2027	**Hours / Week** 2	**Total Hours** 30	**Credits** 2	Employability/ Skill Development

Course Objectives

1. To gain knowledge about the concepts of R programming.
2. To solve Probability Distributions, Correlation and Regression Analysis using R programs.
3. To enhance the students to develop the Data AnalyticswithMachine Learning.

Course Outcomes (CO)

K3 - K5	CO1	Remembering the basic terms used in Data Structures.
	CO2	Understanding and maintenance of data.
	CO3	Applying the notions of Data Visualization and Probability Distributions.
	CO4	Exerting the fundamental concept of Artificial Neural Networks.
	CO5	Evaluating the concept of ML in R with spark and R with H_2O

Program List

1. Creating Data Structures
2. Manage, Explore and Understanding Data
3. Data Visualization
4. Probability Distributions
5. Correlation and Regression Analysis
6. Decision Trees
7. Cluster Analysis
8. Artificial Neural Networks
9. ML in R with Spark
10. ML in R with H2O

Reference Book :

1. Machine Learning Using R, Karthik Ramasubramanian, Abhishek Singh, Apress, 2017

26PDD2CL

Distribution of Marks in ESE			**CIA**		
Experiment	:	50	CIA Practical	:	25
Record	:	10	Exam		
Total		**60**	Attendance	:	5
			Observation Note book	:	10
			Total		**40**

To be awarded jointly by the internal and external examiners.

Mapping

PSO / CO	PSO1	PSO2	PSO3	PSO4	PSO5
CO1	S	S	H	M	S
CO2	S	S	S	H	H
CO3	H	H	S	M	H
CO4	S	H	S	M	S
CO5	H	S	M	S	H

S - Strong; **H**-High; **M**-Medium; **L**-Low

Content prepared by (Based on Inputs) Signature of the Course in-charge	 Signature of the HOD

26PDD2Z1

Programme Code:02		PG Diploma in Data Analytics		
Title of the Paper		**Project – Viva Voce**		
Batch 2026-2027	**Hours / Week** 4	**Total Hours** 60	**Credits** 4	Employability/ Skill Development

Course Objectives

1. To understand the research methodology
2. To get exposed to real time data and study them.
3. To explore the tools and analytics techniques.

Course Outcomes (CO)

K3 to K5	CO1	Choosing the appropriate research area.
	CO2	Illustration of data.
	CO3	Applying relevant techniques to study the scenario.
	CO4	Analyzing the impact factors over the time that lead the fluctuation.
	CO5	Concluding the study by providing appropriate measures and suggestions.

Distribution of Marks in ESE

Dissertation : 120

Viva voce : 40

Total 160

Internal

Project Review : 30

Attendance : 10

Total 40

To be awarded jointly by the internal and external examiners

26PDD2Z1

Mapping

PSO / CO	PSO1	PSO2	PSO3	PSO4	PSO5
CO1	S	S	S	H	H
CO2	H	S	S	M	S
CO3	S	H	H	S	M
CO4	S	H	S	S	M
CO5	H	M	S	H	S

S - Strong; **H**-High; **M**-Medium; **L**-Low

Content prepared by (Based on Inputs) Signature of the Course in-charge	 Signature of the HOD