.

Printed Page:-4 **Subject Code:- ACSAI0713/ACSAIH0713**

Roll. No:

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA
(An Autonomous Institute Affiliated to AKTU, Lucknow)
B.Tech
SEM: VII - THEORY EXAMINATION (2025 - 2026)
Subject: Programming for Data Analytics

Time: 3 Hours **Max. Marks: 100**

General Instructions:

IMP: *Verify that you have received the question paper with the correct course, code, branch etc.*

1. *This Question paper comprises of* ***three Sections -A, B, & C.*** *It consists of Multiple Choice Questions (MCQ's) & Subjective type questions.*

2. *Maximum marks for each question are indicated on right -hand side of each question.*

3. *Illustrate your answers with neat sketches wherever necessary.*

4. *Assume suitable data if necessary.*

5. *Preferably, write the answers in sequential order.*

6. *No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked.*

SECTION-A 20

1. Attempt all parts:-

1-a. List down the following is not a method to read data into a DataFrame in Pandas? (CO1,K1) 1

(a) pd.read_csv()
(b) pd.read_excel()
(c) pd.read_table()
(d) pd.read_html()

1-b. Select method can be used to drop rows with missing values in pandas? (CO1,K1) 1

(a) df.fillna()
(b) df.dropna()
(c) df.replace()
(d) df.interpolate()

1-c. The Dplyr function used to select specific columns is (CO2, K2) 1

(a) select()
(b) mutate()
(c) arrange()
(d) filter()

1-d. R-Shiny is primarily used for (CO2, K2) 1

(a) Data import
(b) Model training
(c) Building interactive web applications
(d) Cleaning datasets

.

1-e. Command to create a connection object in SQLite.(CO3, K3) 1

(a) connect.db('sqlite3')
(b) sqlite3.connect('mydb.db')
(c) sqlite.connect('mydb')
(d) sql.connect()

1-f. INSERT INTO students VALUES (...) is a _____ command.(CO3, K1) 1

(a) DML
(b) DDL
(c) DCL
(d) XML

1-g. Select the following is a key feature of TensorFlow?(CO4, K21) 1

(a) AutoML
(b) Reinforcement learning
(c) Distributed computing
(d) Data visualization

1-h. Select the following operations is performed by a convolutional layer?(CO4, K1) 1

(a) Matrix multiplication
(b) Element-wise multiplication
(c) Convolution
(d) Activation

1-i. Tick the correct one among of the following is a popular deep learning framework?(CO5,K2) 1

(a) Scikit-learn
(b) TensorFlow
(c) Keras
(d) Java

1-j. Select correct one of the following is a commonly used optimization algorithm for training deep neural networks?(CO5,K2) 1

(a) Linear regression
(b) Gradient descent
(c) Naive Bayes
(d) RNN

2. Attempt all parts:-

2.a. Explain a p-value in hypothesis testing? (CO1,K1) 2

2.b. Classify different data structures used in R programming. (CO2, K2) 2

2.c. Mention the role of cursors in SQLite database operations.(CO3, K2) 2

2.d. Specify the role of placeholders in TensorFlow?(CO4, K3) 2

2.e. Explain concept of transfer learning?(CO5,K2) 2

SECTION-B 30

3. Attempt all parts:-

3.a. Answer any one of the following:-

3.a.(i)	Use Pandas functions to handle missing values in a dataset. Demonstrate how dropna() and fillna() work differently. (CO1,K3)	6
3.a.(ii)	Demonstrate how to merge two DataFrames using Pandas. Show how merge() and concat() can be used with examples.(CO1,K3)	6

3.b. Answer any one of the following:-

3.b.(i)	Describe how attributes in R objects influence object behavior, providing code-based examples.(CO2, K2)	6
3.b.(ii)	Demonstrate of the process of creating an interactive dashboard using Flexdashboard and explaining its benefits for data reporting..(CO2, K3)	6

3.c. Answer any one of the following:-

3.c.(i)	Explain the creation of tables and insertion of records in SQLite using Python code including commit and close operations.(CO3, K2)	6
3.c.(ii)	Discuss the structure and key features of MongoDB and explain how it differs from relational databases like SQLite.(CO3, K2)	6

3.d. Answer any one of the following:-

3.d.(i)	Explain concept of backpropagation, and how is it used to train neural networks in TensorFlow?(CO4, K2)	6
3.d.(ii)	Explain the role of the computational graph in TensorFlow.(CO4, K2)	6

3.e. Answer any one of the following:-

3.e.(i)	Explain working of an improved GAN work?(CO5,K2)	6
3.e.(ii)	Explain the key components of a variational autoencoder (VAE).(CO5,K2)	6

SECTION-C 50

4. Answer any one of the following:-

4-a.	Steps for some common measures of central tendency and variability, and how are they used in descriptive statistics?(CO1,K3)	10
4-b.	Explain the difference between a Pandas Series and DataFrame. Provide examples of when each would be used in data analysis. (CO1, K2)	10

5. Answer any one of the following:-

5-a.	Design and implement a complete R script that imports a dataset, cleans it using Dplyr, manipulates text with Stringr, and summarizes data for reporting(CO2, K6)	10
5-b.	Construct a workflow for performing EDA, from data import to visualization, using R packages like Dplyr, ggplot2, and Stringr. Explain each step.(CO2, K6)	10

6. Answer any one of the following:-

6-a.	Discuss the design and implementation of a Python program that creates multiple tables in SQLite, performs transactions, and ensures data consistency using commit and rollback mechanisms.(CO3, K2)	10
6-b.	Analyze a scenario where a Python application performs large-scale data manipulation using SQLite. Explain how batching with executemany() improves performance.(CO3, K4)	10

7. Answer any one of the following:-

7-a.	An HR firm wants to analyze job descriptions and candidate resumes. How can you	10

.

	use word embeddings in TensorFlow to find semantic similarity between texts?(CO4, K2)	
7-b.	Explain how TensorFlow's Keras API simplifies deep learning model building. Discuss its advantages compared to low-level TensorFlow coding.(CO4, K2)	10

8. Answer any one of the following:-

8-a.	Explain the different types of autoencoders, such as basic, denoising, and variational, and how do they differ from each other?(CO5,K2)	10
8-b.	Explain some common architectures used in GANs, such as deep convolutional GANs and Wasserstein GANs, and how are they used in practice? (CO5,K2)	10