**POORNIMA UNIVERSITY, JAIPUR**

**END SEMESTER EXAMINATION, APRIL 2023**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **1BC2101** | Roll No. | Total Printed Pages: 2 |
| **1BC2101** |  |
| BCA I Year II-Semester (Main/Back) End Semester Examination, May 2023  **(AIDS)** | |
| **BASCCA2105 : Introduction to Data Science** | | | |

# Time: **3**Hours. Total Marks: **60**

Min. Passing Marks: **21**

Attempt **five** questions selecting one question from each Unit. There is internal choice from Unit I to Unit V. Marks of each question or its parts are indicated against each question / parts. Draw neat sketches wherever necessary to illustrate the answer. Assume missing data suitably (if any) and clearly indicate the same in the answer.

Use of following supporting material is permitted during examination for this subject.

# **1.--------------------------Nil--------------------** **2.------------------Nil-----------------------**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | **UNIT-I (CO1)** | **Marks** | **Bloom Level** |
| **Q.1** | **(a)** | What does data science mean to you? Describe what advantages and disadvantages are. | **(6)** | **Knowledge** |
|  |  |  |  |  |
|  | **(b)** | What do you mean by facets? Describe all the distinct types. | **(6)** | **Knowledge** |
|  |  |  |  |  |
|  |  | **OR** |  |  |
|  |  |  |  |  |
| **Q.2** | **(a)** | Give a step-by-step explanation of the data science process using an example. | **(6)** | **Knowledge** |
|  |  |  |  |  |
|  | **(b)** | Describe data science. What does the term "data scientist" mean to you? Where is the need for data science? | **(6)** | **Knowledge** |
|  |  |  |  |  |
|  |  | **UNIT-II (CO2)** |  |  |
|  |  |  |  |  |
| **Q.3** | **(a)** | What do you understand by Frequency Distribution? Explain all types of Frequency Distribution with example. | **(6)** | **Knowledge** |
|  |  |  |  |  |
|  | **(b)** | What do you mean when you say "outliers"? Find out the outliers in this series:  22, 25, 17, 19, 33, 64, 23, 17, 20, 18 | **(6)** | **Application** |
|  |  |  |  |  |
|  |  | **OR** |  |  |
|  |  |  |  |  |
| **Q.4** | **(a)** | Find out the Mean and Median   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | Class | 0 - 10 | 10 - 20 | 20 - 30 | 30 -40 | 40-50 | | Frequency | 3 | 5 | 4 | 8 | 2 | | **(6)** | **Application** |
|  |  |  |  |  |
|  | **(b)** | The annual salaries of a group of an employee are given in following table. Find out the Variance, Standard Deviation, Coefficient of Standard Deviation and Coefficient of variance:   |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | X | 45 | 50 | 55 | 60 | 65 | 70 | 75 | 80 | | f | 3 | 5 | 8 | 7 | 9 | 7 | 4 | 7 | | **(6)** | **Application** |
|  |  |  |  |  |
|  |  | **UNIT-III (CO3)** |  |  |
|  |  |  |  |  |
| **Q.5** | **(a)** | What do you understand by Machine learning? Explain its type with example. | **(6)** | **Knowledge** |
|  |  |  |  |  |
|  | **(b)** | Create a linear regression model using this data.   |  |  |  |  |  | | --- | --- | --- | --- | --- | | x | 1 | 2 | 3 | 4 | | y | 3 | 4 | 5 | 7 | | **(6)** | **Application** |
|  |  |  |  |  |
|  |  | **OR** |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| **Q.6** | **(a)** | By KMeans Clustering Algorithm, what do you mean? List all the steps a model  Takes to identify clusters in a dataset along with this example. when k = 2  Dataset = { 2, 3, 4, 10, 11, 12, 20, 25,30 } | **(6)** | **Application** |
|  |  |  |  |  |
|  | **(b)** | What is Decision Tree Algorithm? List down the different types of nodes in Decision Tree. Explain its advantages and disadvantage also. | **(6)** | **Knowledge** |
|  |  |  |  |  |
|  |  | **UNIT-IV (CO4)** |  |  |
|  |  |  |  |  |
| **Q.7** | **(a)** | What is Data Visualization? Why it is so important in data science? Explain its advantages and disadvantages. | **(6)** | **Knowledge** |
|  |  |  |  |  |
|  | **(b)** | What is Spatial Data? How we use this data in data visualization? | **(6)** | **Knowledge** |
|  |  |  |  |  |
|  |  | **OR** |  |  |
|  |  |  |  |  |
| **Q.8** | **(a)** | Describe each step of the Data Visualization Process. | **(6)** | **Knowledge** |
|  |  |  |  |  |
|  | **(b)** | Describe Six graphs that show effective data visualization. | **(6)** | **Application** |
|  |  |  |  |  |
|  |  | **UNITV (CO5)** |  |  |
|  |  |  |  |  |
| **Q.9** | **(a)** | Why did Python become so popular in data science? Describe its attributes or features. | **(6)** | **Knowledge** |
|  |  |  |  |  |
|  | **(b)** | Write down the output of this code:  thislist = ["apple", "banana", "cherry"]  thislist.insert(2, "watermelon")  print(thislist)  print(thislist[0:3])  thislist.append("orange")  print(thislist)  thislist.remove("banana")  print(thislist)  del thislist  print(thislist)  thislist = ["apple", "banana", "cherry"]  thistuple = ("kiwi", "orange")  thislist.extend(thistuple)  print(thislist) | **(6)** | **Application** |
|  |  |  |  |  |
|  |  | **OR** |  |  |
|  |  |  |  |  |
| **Q.10** | **(a)** | Write down the output of this code:  arr = np.array([1, 2, 3, 4, 5, 6, 7], ndmin=5)  print(arr)  print('number of dimensions :', arr.ndim)  print(arr[0])  print(arr[2] + arr[3])  print(arr[1:5])  print(arr[4:])  print(arr[:4])  print(arr[-3:-1])  print(arr[1:5:2])  print(arr[::2])  print(arr[1:5:3])  print(arr [:-5]) | **(6)** | **Application** |
|  |  |  |  |  |
|  | **(b)** | What does "set," "list," "tuple," and "dictionary" mean in Python? Explain with coding. | **(6)** | **Knowledge** |