

Roll No																			
---------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--



PRESIDENCY UNIVERSITY BENGALURU

SCHOOL OF ENGINEERING END TERM EXAMINATION - JAN 2023

Semester : Semester V - 2020

Course Code : ECE3106

Course Name : Sem V - ECE3106 - Introduction to Data Analytics

Program : B.Tech. Electronics and Communication Engineering

Date : 13-JAN-2023

Time : 9.30AM - 12.30PM

Max Marks : 100

Weightage : 50%

Instructions:

- (i) Read all questions carefully and answer accordingly.
- (ii) Question paper consists of 3 parts.
- (iii) Scientific and non-programmable calculator are permitted.

PART A

ANSWER ALL THE FIVE QUESTIONS

5 X 2 = 10M

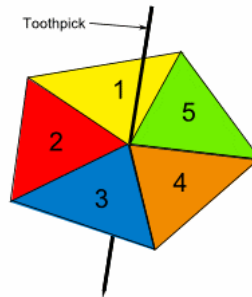
1. The data files are the files that store data pertaining to a specific application for later use. Data files can be used in two ways: Text Files and Binary Files. What does the `rstrip()` function do?
(CO1,CO2) [Knowledge]
2. While defining an array, we can use `arrange` function followed by the starting number and the end number. `Reshape` function is used for converting the matrix into a particular size. Find $A * B$ for the matrices $A = \text{np.arange}(0,9).\text{reshape}(3,3)$ and $B = \text{np.ones}((3,3))$
(CO3,CO2) [Knowledge]
3. Probability is simply how likely something is to happen. Many events can't be predicted with total certainty. The best we can say is how likely they are to happen, using the idea of probability. What is the probability of rolling a "4" with a die
(CO4,CO3) [Knowledge]
4. An array is a collection of items of same data type stored at contiguous memory locations. While defining an array, we can use `arrange` function followed by the starting number and the end number. We can define some logical conditions for such array and matrices. For the following matrix,

```
>>> A = np.random.random((4, 4))
>>> A
array([[ 0.03536295,  0.0035115 ,  0.54742404,  0.68960999],
       [ 0.21264709,  0.17121982,  0.81090212,  0.43408927],
       [ 0.77116263,  0.04523647,  0.84632378,  0.54450749],
       [ 0.86964585,  0.6470581 ,  0.42582897,  0.22286282]])
```

Find $A < 0.5$. Write True if the condition satisfies and false otherwise

(CO3,CO2) [Knowledge]

5. Probability refers to how likely something is to happen. Many events can't be predicted with total certainty. The best we can say is how likely they are to happen, using the idea of probability. The diagram shows a spinner made up of a piece of card in the shape of a regular pentagon, with a toothpick pushed through its center. The five triangles are numbered from 1 to 5. Each time, the spinner is spun until it lands on one of the five edges of the pentagon. If X represents the number on that edge what is $P(X \text{ is prime})$?



(CO3,CO2) [Knowledge]

PART B

ANSWER ALL THE TWO QUESTIONS

2 X 15 = 30M

6. A random variable is a mathematical formalization of a quantity or object which depends on random events. It is a mapping or a function from possible outcomes in a sample space to a measurable space. Four coins are tossed.
- If Y represents the number of tails, what is $P(Y \leq 1)$?
 - List all the outcomes Write the Sample Space. What is the probability of $Y=1$, $Y=2$ and $Y=3$?
- (CO3,CO2) [Comprehension]
7. a) File handling is an important part of any web application. Python has several functions for creating, reading, updating, and deleting files . You have a file named "demo.txt" located in "C:/Users/lovely/desktop" in your computer. Write python instructions to perform the following operations
- Open the file and read all the characters in the file
 - Open the file and read only first 10 the characters in the file
 - Open the file and read all the lines in the file
 - Open the file and read a line in the file
 - Without deleting the data in the file add your name and your ID number.
 - overwrite the file with the following message " PRESIDENCY UNIVERSITY"
- b) You can split a matrix in a number of ways. NumPy uses the concept of splitting, providing a number of functions in this regard. Explain with suitable examples the different splitting functions of NumPy.
- (CO3,CO2) [Comprehension]

PART C

ANSWER ALL THE THREE QUESTIONS

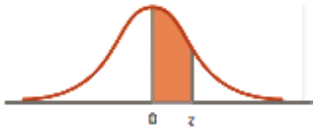
3 X 20 = 60M

8. *Normal distribution* is a continuous probability distribution wherein values lie in a symmetrical fashion mostly situated around the mean. There are many cases where the data tends to be around a central value with no bias left or right, and it gets close to a "Normal Distribution".
- Your score in a recent test was 0.5 standard deviations above the average, how many people scored lower than you did?
 - Your score in a recent test was 1.5 standard deviations above the average, how many people scored more than you did?

(CO1,CO2) [Application]

9. Z-score is a statistical measurement that describes a value's relationship to the mean of a group of values. Z-score is measured in terms of standard deviations from the mean. For the following sequence find
- the z-score values and
 - the corresponding probability values using the following z- score table
- 26, 33, 65, 28, 34, 55, 25, 44, 50

Table II Normal Curve Areas



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Source: Abridged from Table I of A. Hald, *Statistical Tables and Formulas* (New York: Wiley), 1952.

(CO2,CO3) [Application]

10. Normally, an *array* is a collection of similar type of elements which has contiguous memory location. While defining an array, we can use arrange function followed by the starting number and the end number. Reshape function is used for converting the matrix into a particular size. Find the following for the array `a=np.arange(10,16)`

- `a[1:5]`
- `a[1:5:2]`
- `a[:,2]`
- `a[:5:2]`
- `a[:5:]`

(CO3,CO2,CO4) [Application]
