

# P P SAVANI UNIVERSITY

Fifth Semester of B.Sc.IT Examination

December 2022

SSIT3021 Data Science

22.11.2022, Tuesday

Time: 10:00 a.m. To 12:30 p.m.

Maximum Marks: 60

## Instructions:

1. The question paper comprises of two sections.
2. Section I and II must be attempted in separate answer sheets.
3. Make suitable assumptions and draw neat figures wherever required.
4. Use of scientific calculator is allowed.

## SECTION - I

Q - 1	MCQ/Short Question/Fill in the Blanks	[05]	CO	BTL
(i)	What is the role of data scientist?		1	1
(ii)	How Statistics plays vital role in data science?		3	2
(iii)	_____ type of data is used to label variables without any order or quantitative value.		1	1
(iv)	Data from different websites are considered as which type of data?		2	1
(v)	Linear regression analysis is used to predict the value of a variable based on the value of another variable. (T/F)		2	2
Q - 2 (a)	Draw Data Science Life Cycle and Explain each process step in detail.	[05]	1	2
Q - 2 (b)	What is Data Exploration? Explain its importance in Data Science.	[05]	2	2
<b>OR</b>				
Q - 2 (a)	Explain how we can collect data from multiple data sources with example.	[05]	2	3
Q - 2 (b)	Explain real life application of Primary Data in Data Science.	[05]	1	3
Q - 3 (a)	What is SVM, elaborate this algorithm with suitable example.	[05]	3	3
Q - 3 (b)	Draw a neat Graphical representation of Linear Regression and explain it in your own words	[05]	3	2
Q - 4	State down the usage of the following libraries/Packages a) NumPy b) Matplotlib c) Pandas d) Sklearn e) SciPy	[05]	3	1

**SECTION - II**

<b>Q - 1</b>	<b>MCQ/Short Question/Fill in the Blanks (Any Five)</b>	<b>[05]</b>		
(i)	Which are pros of data visualization? a. It can be distracting b. It can be accessed quickly by a wider audience. c. It can misrepresent information d. It can misrepresent information		4	1
(ii)	Line charts are the most basic and commonly used techniques (True / False)		4	2
(iii)	What makes data visualization good?		5	4
(iv)	Enlist five application of Data science.		6	1
(v)	What are the four stages of data visualization?		5	1
(vi)	Line charts are used to represent how much each category contributes to a total sum, such as percentages or totals, over time. (True / False)		5	2
(vii)	Mention 3 advantages of using data visualization.		5	1
<b>Q - 2 (a)</b>	Explain visual Encoding	<b>[05]</b>	6	1
<b>Q - 2 (b)</b>	Explain factors affecting the data visualization choices.	<b>[05]</b>	6	1
<b>OR</b>				
<b>Q - 2 (a)</b>	What are retinal variables? Explain in detail.	<b>[05]</b>	4	2
<b>Q - 2 (b)</b>	Explain various visualization techniques	<b>[05]</b>	6	1
<b>Q - 3 (a)</b>	Explain any 3 advantage of data visualization.	<b>[05]</b>	4	1
<b>Q - 3 (b)</b>	Explain application of data science in finance.	<b>[05]</b>	6	1
<b>OR</b>				
<b>Q - 3 (a)</b>	Explain any 2 data science techniques in detail.	<b>[05]</b>	5	5
<b>Q - 3 (b)</b>	Explain application of data science in medical field.	<b>[05]</b>	6	5

Q - 4

**Attempt anyone.**

[05]

(i)

How data visualization is helpful in cricket matches?

6

2

(ii)

How data visualization is helpful in Movies?

6

2

\*\*\*\*\*

CO: Course Outcome Number

BTL: Blooms Taxonomy Level

Level of Bloom's Revised Taxonomy in Assessment

1: Remember	2: Understand	3: Apply
4: Analyze	5: Evaluate	6: Create