

INDIRA UNIVERSITY, PUNE

SCHOOL OF INFORMATION TECHNOLOGY- BSC DS

Term End Examination (2025 Pattern) April- May – 2026- Semester – II

Course Name: Data Analytics
Course Code: 25BDS212T

Max. Marks: 25
Time: 1 Hr. 30 Minutes

Instructions:

- All Questions are compulsory.
- Figures to the right indicates full marks.

Q1.	Attempt any 5 out of 7. a) Define data analytics. b) What is raw data? c) What is predictive analytics? d) Define Machine Learning. e) What is clustering? f) Define Social Media Analytics. g) What is tokenization?	(5 Marks)	CO1												
Q2.	Attempt any 2 out of 4. a) Explain the types of data analytics with suitable examples. b) Describe supervised learning and its types. c) Explain data characterization with steps and example. d) Explain the seven layers of Social Media Analytics.	(10 Marks)	CO2												
Q.3	Attempt all questions. (5 marks each) a) Describe key techniques used in text analytics such as tokenization, Bag of Words, n-grams, stop word removal. b) Given transaction data, find frequent item sets using minimum support=2. <table><tr><td>Transaction ID</td><td>Item Purchased</td></tr><tr><td>T1</td><td>Bread, Butter</td></tr><tr><td>T2</td><td>Bread, Milk</td></tr><tr><td>T3</td><td>Milk, Butter</td></tr><tr><td>T4</td><td>Bread, Butter, Milk</td></tr><tr><td>T5</td><td>Bread, Butter</td></tr></table> OR	Transaction ID	Item Purchased	T1	Bread, Butter	T2	Bread, Milk	T3	Milk, Butter	T4	Bread, Butter, Milk	T5	Bread, Butter	(10 Marks)	CO3
Transaction ID	Item Purchased														
T1	Bread, Butter														
T2	Bread, Milk														
T3	Milk, Butter														
T4	Bread, Butter, Milk														
T5	Bread, Butter														

Q.3	Attempt all questions. (5 marks each)	(10 Marks)	CO3										
a) Perform the following NLP tasks on the given Instagram comment. Show all steps clearly.													
Comment: “This phone camera is amazing!!!”													
a) Tokenization b) Stop word removal c) Identify bigrams d) POS (Parts of Speech) tagging													
b) A dataset shows the relationship between study hours and marks:													
<table><tr><td>Hours (X)</td><td>Marks (Y)</td></tr><tr><td>2</td><td>40</td></tr><tr><td>4</td><td>50</td></tr><tr><td>6</td><td>65</td></tr><tr><td>8</td><td>80</td></tr></table>				Hours (X)	Marks (Y)	2	40	4	50	6	65	8	80
Hours (X)	Marks (Y)												
2	40												
4	50												
6	65												
8	80												
1. Find the regression equation ($Y = a + bX$). 2. Predict marks when a student studies for 5 hours.													
