Total No. of Questions: 6 *Total No. of Printed Pages:3*

Enrollment No.....

Faculty of Management Studies
End Sem Examination Dec 2024
MS5CO38 Data Privacy & Ethics

Programme: MBA Branch/Specialisation: Business Analytics

Duration: 3 Hrs. **Maximum Marks: 60**

Note: All questions are compulsory. Internal choices, if any, are indicated. Answers of Q.1 (MCQs) should be written in full instead of only a, b, c or d. Assume suitable data if necessary. Notations and symbols have their usual meaning.

			Marks	BL	PO	CO	PSO
Q.1	i.	What is the primary purpose of data pre-processing? (a) To delete data (b) To structure data for analysis (c) To distort data intentionally (d) To avoid data governance	**1**	1	1	1	
	ii.	Which of the following is NOT a data format? (a) CSV (b) JSON (c) HTML (d) RAM	**1**	1	1	1	
	iii.	What is "data minimization"? (a) Collecting as much data as possible (b) Reducing collected data to essentials (c) Avoiding data storage (d) Maximizing data usage	**1**	2	2	2	
	iv.	Which term refers to ensuring fairness in algorithms? (a) Anonymity (b) Algorithmic fairness (c) Metadata management (d) Data sharing	**1**	2	2	2	
	v.	Why is data anonymization important? (a) Ensures data is completely deleted (b) It prevents identification of individuals (c) It simplifies data sorting (d) It removes ethical considerations	**1**	2	3	3	

vi.	Which of the following is a feature of unbiased data? (a) Transparency (b) Ownership (c) Fabrication (d) Inconsistency	**1**	2	3	3		
vii.	What does metadata describe? (a) Data security algorithms (b) Data ownership (c) Structural and descriptive aspects of data (d) Data pre-processing	**1**	2	4	4		
viii.	What is Big Query primarily used for? (a) Data anonymization (b) Managing small data sets (c) Processing large datasets (d) Designing databases	**1**	2	4	4		
ix.	What is a key principle of privacy design? (a) Ignoring user consent (b) Maximizing data sharing (c) Transparency (d) Removing compliance	**1**	2	8	5		
x.	Which of the following is a key aspect of data protection laws? (a) Ignoring third-party data sharing (b) Ensuring transparency and compliance (c) Avoiding physical security measures (d) Eliminating anonymization techniques	**1**	2	8	5		
Q.2	i. Explain the difference between structured and unstructured data with an example.	**2**	2	1	1		
	ii. What are the key steps in the pipeline of data access?	**3**	3	4	4		
	iii. Discuss the importance of data pre-processing in ensuring data credibility, privacy, and ethics.	**5**	4	5	3		
OR	iv. Explain data modeling techniques.	**5**	2	4	4		
Q.3	i. What is data minimization, and why it is an essential principle in data privacy?	**2**	2	8	2		

	ii. Discuss the modern risks to data privacy and explain how organizations can mitigate these risks.	**8**	4	8	2		
OR	iii. Explain the impact of errors in data processing and model design on ensuring fairness and accountability in data systems.	**8**	4	5	3		
Q.4	i. What is the importance of data ethics?	**3**	2	5	3		
	ii. Explain what is meant by "fairness" in data ethics. Why is it important for data to be unbiased?	**7**	3	5	3		
OR	iii. What do you mean by open data usage? Explain it with its features and characteristics.	**7**	3	5	3		
Q.5	i. What is metadata, and why is it important in managing databases?	**4**	2	4	4		
	ii. Explain the role of relational databases in organizing and securing data. Provide an example of their use.	**6**	2	4	4		
OR	iii. What is data integration, and how does it help in accessing and managing data from different sources?	**6**	2	4	4		
Q.6	Attempt any two:						
	i. What is personal data, and why is transparency important in its protection?	**5**	2	8	5		
	ii. Explain the importance of balancing security and analytics in data protection.	**5**	3	8	5		
	iii. What are privacy laws, and how do they help in ensuring compliance with data protection standards?	**5**	2	8	5		

******

Marking Scheme
(MS5CO38) Data Privacy and Ethics-(T)

Q.1	i)	b. To structure data for analysis	**1**
	ii)	d. RAM	**1**
	iii)	b. Reducing collected data to essentials	**1**
	iv)	b. Algorithmic fairness	**1**
	v)	b. It prevents identification of individuals	**1**
	vi)	a. Transparency	**1**
	vii)	c. Structural and descriptive aspects of data	**1**
	viii)	c. Processing large datasets	**1**
	ix)	c. Transparency	**1**
	x)	b. Ensuring transparency and compliance	**1**
Q.2	i.	difference between structured and unstructured data with an example – 2 marks	**2**
	ii.	key steps in the pipeline of data access – 3 marks	**3**
	iii.	importance of data pre-processing in ensuring data credibility (2 marks), privacy (1 marks), and ethics (2 marks)	**5**
OR	iv.	data modelling techniques -5 marks	**5**
Q.3	i.	What is data minimization- 1 mark, its importance – 1mark	**2**
	ii.	modern risks to data privacy- 4 risk (1 mark each), mitigation these risks – 4 marks.	**8**
OR	iii.	impact of errors in data processing- 4 marks, model design – 4 marks.	**8**
Q.4	i.	importance of data ethics- 3 points 3 marks	**3**
	ii.	fairness in data ethics – 4 marks. Why is it important for data to be unbiased – 3 marks.	**7**
OR	iii.	open data usage. – 3 marks its features and characteristics – 4 marks	**7**
Q.5	i.	What is metadata, -2 marks, important in managing databases- 2 marks.	**4**
	ii.	role of relational databases -3 marks, example – 3 marks	**6**
OR	iii.	What is data integration -2 marks, accessing and managing – 4 marks	**6**
Q.6			
	i.	What is personal data -2 marks, importance -3 marks	**5**
	ii.	At least 5 points, 1 mark each	**5**
	iii.	Privacy laws – 2 marks, ensuring compliance – 3 marks	**5**

******