

**Sixth Semester**  
**Computer Science and Engineering**  
**Scheme OCBC 2019**  
**DATASCIENCE**

**Time : Three Hours**

**Maximum Marks : 70**

**Note :** i) All 7 Questions are **Compulsory**. Internal choices has been given in each LO (Learning Outcome)


सभी 7 प्रश्न अनिवार्य हैं। आंतरिक विकल्प प्रत्येक LO (लर्निंग आउटकम) में दिए गए हैं।

ii) In case of any doubt or dispute, the English version question should be treated as final.

किसी भी प्रकार के संदेह अथवा विवाद की स्थिति में अंग्रेजी भाषा के प्रश्न को अंतिम माना जायेगा।

Q.	LO	Questions	Marks
1.	LO1	Define data science. Explain the advantages disadvantages and its application in industry. डाटा साइंस को परिभाषित कीजिए। इसके लाभ, हानि व अनुप्रयोगों (इंडस्ट्री में) की व्याख्या कीजिए।	10
		OR/अथवा Explain the history and roles of data science. डाटा साइंस के इतिहास के बारे में बताइए एवं डाटा साइंस के रोल्स के बारे में व्याख्या कीजिए।	10
2.	LO2	Explain various stages of data mining. डाटा माइनिंग की विभिन्न स्टेज की व्याख्या कीजिए।	10
		OR/अथवा Explain data pre-processing techniques. डाटा प्री-प्रोसेसिंग टेक्नीक की व्याख्या कीजिए।	10

Q.	LO	Questions	Marks
3.	LO4	<p>Differentiate between continuous and discrete data.</p> <p>कन्टीन्यूअस एवं डिस्क्रीट डाटा के मध्य अंतर कीजिए।</p> <p>OR/अथवा</p> <p>Explain various types of attributes with appropriate example.</p> <p>विभिन्न प्रकार के एट्रीब्यूट्स को उपयुक्त उदाहरण की सहायता से समझाइए।</p>	<p>10</p> <p>10</p>
4.	LO5	<p>Explain mean and median with the help of suitable example.</p> <p>मीन एवं मीडियन की उदाहरण सहित व्याख्या कीजिए।</p> <p>OR/अथवा</p> <p>Explain the ways to calculate mode of any data sample.</p> <p>किसी डाटा सेम्पल के मोड को केलकुलेट करने के तरीकों की व्याख्या कीजिए।</p>	<p>10</p> <p>10</p>
5.	LO6	<p>The students of a class have given two tests of the subject data science. In the first test, 60% of the students pass while only 40% of the students passed both the test. What percentage of students who passed the first test, also passed the second test?</p> <p>डाटा साइंस विषय की कक्षा में विद्यार्थियों को इस विषय के दो टेस्ट दिये गए। पहले टेस्ट में कुल 60% विद्यार्थी उत्तीर्ण हुये जबकि 40% विद्यार्थी दोनों टेस्ट में उत्तीर्ण हुये। कितने प्रतिशत विद्यार्थी ऐसे थे जिन्होंने पहला टेस्ट एवं दूसरा टेस्ट भी उत्तीर्ण किया?</p> <p>OR/अथवा</p>	10

Q.	LO	Questions	Marks																										
		Explain experiment, sample space and event. एक्पेरीमेंट, सॅपल स्पेस एवं इवेंट की व्याख्या कीजिए।	10																										
6.	LO8	Explain the Apriori algorithm. एप्रीआरी एल्गोरिथम की व्याख्या कीजिए। OR/अथवा Explain support and confidence with suitable example. सपोर्ट एवं कान्फीडेंस की उदाहरण सहित व्याख्या कीजिए।	10  10																										
7.	LO13	Table 1 shows the respective heights of x and y of a sample of 12 fathers and their oldest sons. Construct a scatter diagram. टेबल 1 में 12 पिताओं और उनके सबसे बड़े पुत्रों के नमूने की x और y की संबंधित ऊँचाई दिखाती है। एक स्कैटर आरेख का निर्माण करें। Table 1 <table border="1" style="margin-left: auto; margin-right: auto;"> <tbody> <tr> <td>Height of Father (inches)</td> <td>65</td> <td>63</td> <td>67</td> <td>64</td> <td>68</td> <td>62</td> <td>70</td> <td>66</td> <td>68</td> <td>67</td> <td>69</td> <td>71</td> </tr> <tr> <td>Height of son (inches)</td> <td>68</td> <td>66</td> <td>68</td> <td>65</td> <td>69</td> <td>66</td> <td>68</td> <td>65</td> <td>71</td> <td>67</td> <td>68</td> <td>70</td> </tr> </tbody> </table>	Height of Father (inches)	65	63	67	64	68	62	70	66	68	67	69	71	Height of son (inches)	68	66	68	65	69	66	68	65	71	67	68	70	10
Height of Father (inches)	65	63	67	64	68	62	70	66	68	67	69	71																	
Height of son (inches)	68	66	68	65	69	66	68	65	71	67	68	70																	
	LO13	Differentiate between linear and multiple regression. लीनियर एवं मल्टीपल रिग्रेशन के मध्य अंतर कीजिए। 	10																										