



Sardar Patel Institute of Technology

Bhavan's Campus, Munshi Nagar, Andheri (W), Mumbai : 400058, India

(Autonomous College of Affiliated to University of Mumbai)

End Semester Examination

December 2022

Maxi Marks : 100

Class : BEIT/BE COMP




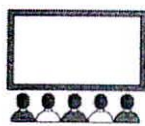


Course code: CS414

Name of the course : Big Data Analytics and Information retrieval

Duration : 3 hours

Semester : VII

Branch : IT/Computer

Q No		Max Marks	CO	BL
1	<p>Observe the following statistics about BookMyShow and justify why you think it is an example of Big Data. Ensure your answer is based on the details mentioned in the case study.</p> <p style="text-align: center;">Quick Facts</p> <div style="display: flex; justify-content: space-around; align-items: flex-end;"> <div style="text-align: center;">  30 Million+ <small>CUSTOMERS</small> </div> <div style="text-align: center;">  5 <small>COUNTRIES</small> </div> <div style="text-align: center;">  650+ <small>TOWNS AND CITIES</small> </div> <div style="text-align: center;">  5000+ <small>SCREENS</small> </div> </div> <div style="display: flex; justify-content: space-around; align-items: flex-end; margin-top: 20px;"> <div style="text-align: center;"> 50 Million+ <small>APP DOWNLOADS</small> </div> <div style="text-align: center;">  15 Million+ <small>TICKETS A MONTH</small> </div> <div style="text-align: center;">  2 Billion+ <small>PAGE VIEWS A MONTH</small> </div> </div>	10M	CO1	4
2	<p>Consider a case study mentioned in question no 1. You are appointed as a data analyst in BookMyShow and assigned some task of data analyst.</p> <p>List down the important columns/features which you think are important in the dataset. Describe your dataset in brief.</p> <p>Write 6 questions about this dataset which will give you useful insights.</p> <p>What type of data visualization you will use to support these insights. Draw diagrams.</p>	03M 03M 04M	CO1	5
3	<p>Consider the FIFA 20 dataset. Part of the dataset is shown below. Write Hadoop map reduce code in java to answer following queries.</p>		CO2	3

For official use only (not for students)

CO Number	CO statement	Marks allotted
CO1	Identify challenges in big data management and inadequacy of existing technology to analyze big data problems.	20
CO2	Apply scalable algorithms based on Hadoop and Map Reduce to perform Big Data Analytics.	20
CO3	Use big data tools for information retrieval.	20
CO4	Use stream data model to provide real time analysis of big data.	20
CO5	Analyze data using graphs.	20

PI chart for percentage CO marks