केन्द्रीय विद्यालय संगठन
कोलकाता संभाग
KENDRIYA VIDYALAYA SANGATHAN
KOLKATA REGION

अध्ययन सामग्री

STUDY MATERIAL

कक्षा : बारहवीं

CLASS: XII

इन्फॉर्मैटिक्स प्रैक्टिसेज

Informatics Practices (065)

2023-24

तत् त्वं पूषन् अपावृणु
केन्द्रीय विद्यालय संगठन

कोलकाता संभाग

KOLKATA REGION

Chief Patron :	**Mr. Y. Arun Kumar** **Deputy Commissioner, KVS RO Kolkata**
Patrons :	**1)Mr. Sanjib Sinha** **Assistant Commissioner, KVS RO Kolkata** **2)Mr. Chintapalli Vijaya Ratnam** **Assistant Commissioner, KVS RO Kolkata** **3)Mr. Dibakara Bhoi** **Assistant Commissioner, KVS RO Kolkata** **4) Mr. Amit Baidya** **Assistant Commissioner, KVS RO KOL.**
Co-Ordinator :	**Sh. Amarendra Kumar Jha** **Principal I/C,KV NO-1 AFS Kalaikunda**
Resource Person:	**Sh. Ashish Kumar Singh (PGT Comp. Sc.)** **K. V. CRPF Durgapur**
Contributors :-	**1. Mr.Debjit Biswas, PGT CS, K.V.Ballygunge** **2. Mrs. Madhulika, PGT CS, KVBarrackpore(AFS).** **3. Mrs Lakshmi Kumari, PGT CS, KV IIM JOKA.** **4. Mrs. Nandini Das,PGT CS, Kharagpur No.2 Rly** **5. Mr. Dipayan Sarkar , PGT CS, KV Birbhum** **6. Mrs. Payal Bhattacharjee, PGT CS, KV Berhampur** **7. Mr. Amit Kr. Malla, PGT (CS), K.V Command Hospital** **8. Mr. A K Jha, PGT (CS), KV No-1 Saltlake**

Salient features of this Study Material

1. Targeting 100% pass rate.
2. Ensuring a minimum level of learning.
3. Providing support for improving risers and other students of Class XII in the subject of IP (Informatics Practices).
4. Provide an introduction to the chapter, including its relevance and importance.
5. Summarize the key concepts and takeaways from the chapter.
6. Multiple Choice Questions (MCQs):

   Present a set of multiple-choice questions related to the chapter .Include options for each question and provide correct answers.

7. Assertion and Reason-Based Questions:

   Include assertion and reason-based questions to test logical reasoning. and provide correct answers.

8. Questions of 2 Marks (Knowledge/Understanding/Application-Based)

   Create questions related to identifying errors in code or predicting program output and Include model answers.

9. 3 Marks Questions (Knowledge/Understanding/Application-Based):

   Formulate questions requiring a deeper understanding of the chapter.

   Include questions that test the application of concepts and provide model answers.

10. 4 Marks Questions (Knowledge/Understanding/Application-Based):

    Include comprehensive questions combining various aspects of the chapter.

    Cover topics like application, output, errors, and in-depth understanding and provide detailed model answers

11. Case Study-Based Questions:

    Include case study questions that require students to analyse and apply their knowledge to real-world scenarios. Provided model answers or suggested approaches to solving the case studies.

INDEX

DISTRIBUTION OF MARKS

Informatics Practices
CLASS XII
Code No. 065
2023-2024

1. **Prerequisite:** Informatics Practices – Class XI

2. **Learning Outcomes**
   At the end of this course, students will be able to:
   - Create Series, Data frames and apply various operations.
   - Visualize data using relevant graphs.
   - Design SQL queries using aggregate functions.
   - Import/Export data between SQL database and Pandas.
   - Learn terminology related to networking and internet.
   - Identify internet security issues and configure browser settings.
   - Understand the impact of technology on society including gender and disability issues.

3. **Distribution of Marks and Periods**

Unit No	Unit Name	Marks	Periods Theory	Periods Practical	Total Period
1	Data Handling using Pandas and Data Visualization	25	25	25	50
2	Database Query using SQL	25	20	17	37
3	Introduction to Computer Networks	10	12	0	12
4	Societal Impacts	10	14	-	14
	Project	-	-	7	7
	Practical	30	-	-	-
	Total	100	71	49	120

SYLLABUS

INFORMATICS PRACTICES(065)

SESSION- 2023-24

Unit 1: Data Handling using Pandas -I

Introduction to Python libraries- Pandas, Matplotlib. Data structures in Pandas - Series and Data Frames.

Series: Creation of Series from – ndarray, dictionary, scalar value; mathematical operations; Head and Tail functions; Selection, Indexing and Slicing.

Data Frames: creation - from dictionary of Series, list of dictionaries, Text/CSV files; display; iteration; Operations on rows and columns: add, select, delete, rename; Head and Tail functions; Indexing using Labels, Boolean Indexing; Importing/Exporting Data between CSV files and Data Frames.

Data Visualization Purpose of plotting; drawing and saving following types of plots using Matplotlib – line plot, bar graph, histogram Customizing plots: adding label, title, and legend in plots.

Unit 2: Database Query using SQL

Revision of database concepts and SQL commands covered in class XI

Math functions: POWER (), ROUND (), MOD ().

Text functions: UCASE ()/UPPER (), LCASE ()/LOWER (), MID ()/SUBSTRING ()/SUBSTR (), LENGTH (), LEFT (), RIGHT (), INSTR (), LTRIM (), RTRIM (), TRIM ().

Date Functions: NOW (), DATE (), MONTH (), MONTHNAME (), YEAR (), DAY (), DAYNAME ().

Aggregate Functions: MAX (), MIN (), AVG (), SUM (), COUNT (); using COUNT (*).

Querying and manipulating data using Group by, Having, Order by. Working with two tables using equi-join

Unit 3: Introduction to Computer Networks

Introduction to networks,

Types of network: PAN, LAN, MAN, WAN.

Network Devices: modem, hub, switch, repeater, router, gateway

Network Topologies: Star, Bus, Tree, Mesh.

Introduction to Internet, URL, WWW, and its applications- Web, email, Chat, VoIP.

Website: Introduction, difference between a website and webpage, static vs dynamic web page, web server and hosting of a website.

Web Browsers: Introduction, commonly used browsers, browser settings, add-ons and plug-ins, cookies.

Unit 4: Societal Impacts

Digital footprint, net and communication etiquettes, data protection, intellectual property rights (IPR), plagiarism, licensing and copyright, free and open source software (FOSS), cybercrime and cyber laws, hacking, phishing, cyber bullying, overview of Indian IT Act.

E-waste: hazards and management. Awareness about health concerns related to the usage of technology.

Name of Chapter: Pandas Series

Topics covered : Series: Creation of Series from – ndarray, dictionary, scalar value; mathematical operations; Head and Tail functions; Selection, Indexing and Slicing.

KEY POINTS:

- **Python libraries contain a collection of built-in modules that allow us to perform many actions without writing detailed programs for it.**
- **A LIBRARY in python can be used by using import <library_name>**
- **NumPy, Pandas and Matplotlib are three well-established Python libraries for scientific and analytical use.**
- **NumPy stands for 'Numerical Python',**
- **Pandas stands for 'PAnelDAta'**
- **PANDAS is a high-level data manipulation tool used for analyzing data**
- **The Matplotlib library in Python is used for plotting graphs and visualization**
- **Pandas has three important data structures, namely – Series, DataFrame and Panel to make the process of analyzing data organized, effective and efficient**
- **Before using of the above libraries one needs to install as follows:**
- **To install pandas library ->pip install pandas**
- **To install numpy library ->pip install numpy**
- **To install matplotlib library ->pip install matplotlib**
- **Data Structure - A data structure is a collection of data values and operations that can be applied to that data**
- **A Series is a one-dimensional array containing a sequence of values of any data type (int, float, string, list etc.)**
- **By default the values of a Series have numeric data labels starting from zero.**
- **A series can be created from any sequence data type or dictionary.**
- **The data label associated with a particular value in a series is called its index**

1.1 Creation of Series:

(A) Creation of Series from Scalar Values:

```
>>> import pandas as pd #import Pandas with alias pd
>>> series1 = pd.Series([10,20,30]) #create a Series
>>> print(series1) #Display the series
```

Output:

```
0   10
1   20
2   30
dtype: int64
```

```
>>>series2=pd.Series([1,2,3],index=['Jan','Feb','Mar'])
>>>print(series2)
```

Output:

```
Jan   1
Feb   2
Mar   3
dtype: int64
```

(B) Creation of Series from NumPy Arrays (ndarray):

```
>>> import numpy as np # import NumPy with alias np
>>> import pandas as pd
>>> array1 = np.array([1,2,3,4])
>>> series3 = pd.Series(array1)
>>> print(series3)
```

Output:

```
0 1
1 2
2 3
3 4
dtype: int32
```

Note:- The index labels and values of Series must be of same size otherwise it will result in a ValueError.

(C) Creation of Series from Dictionary:

```
>>>import pandas as pd
>>>dict1 = {'Rajasthan': 'Jaipur', 'West Bengal':'Kolkata', 'Assam': 'Dispur'}
>>>print (dict1)
{'Rajasthan': 'Jaipur', 'West Bengal': 'Kolkata', 'Assam': 'Dispur'}
>>>series4=pd.Series(dict1)
>>>print(series4)
```

output:

```
Rajasthan     Jaipur
West Bengal   Kolkata
Assam         Dispur
dtype: object
```

1.2 Accessing Elements of a Series:

There are two common ways for accessing the elementsof a series: **Indexing** and **Slicing.**

(A) Indexing:

- It is used to access elements in a series.

Indexes are of two types: **positional index** and **labelled index**. Positional index takes

```
>>> import pandas as pd
```

- an integer value that corresponds to its position in the series starting from 0, whereas labelled index takes any user-defined label as index.

(i) Accessing elements using positional index:

```
>>>seriesvow = pd.Series(['a','e','i','o','u'])
>>>seriesnow
    0  a
    1  e
    2  i
    3  o
    4  u
    dtype: object
>>>seriesvow[2]
    'i'
>>>seriesvow[5]
    KeyError: 5
>>>seriesvow[-1]
    KeyError: -1
```

(iI) Accessing elements using labelled index:

```
>>> import pandas as pd
>>>srcapcountry=pd.Series(['INDIA','JAPAN','GERMANY','RUSSIA'],index=['NEW
DELHI','TOKYO','BERLIN','MOSCOW'])
>>>srcapcountry

NEW DELHI     INDIA
TOKYO         JAPAN
BERLIN      GERMANY
MOSCOW       RUSSIA
dtype: object

>>>srcapcountry['TOKYO']

'JAPAN'

>>>srcapcountry[1]  #POSITIONAL INDEX ALSO WORKS

'JAPAN'
```

Accessing more than one element using positional / labelled index:

```
>>>srcapcountry=pd.Series(['INDIA','JAPAN','GERMANY','RUSSIA'],index=['NEW
        DELHI','TOKYO','BERLIN','MOSCOW'])

>>>srcapcountry[[3,2]]          #INDEXING CAN BE USED IN ANY ORDER

MOSCOW     RUSSIA
BERLIN    GERMANY
dtype: object

>>>srcapcountry[['MOSCOW','BERLIN']]      #INDEXING CAN BE USED IN ANY ORDER

MOSCOW     RUSSIA
BERLIN    GERMANY
dtype: object

>>>srcapcountry[['TOKYO','MOSCOW']]

TOKYO      JAPAN
MOSCOW    RUSSIA
dtype: object
```

(B) **Slicing:**

- It is used to extract a part of a series.
- Syntax: <seriesname> = [startindex : endindex]
- When we use positional indices for slicing, the value at the endindex position is excluded
- By default step is 1 while displaying using positional index

```
>>>srcapcountry=pd.Series(['INDIA','JAPAN','GERMANY','RUSSIA'],index=['NEW
        DELHI','TOKYO','BERLIN','MOSCOW'])
>>>srcapcountry[2:4]    # elements at index position 2 to 3 will be displayed

BERLIN    GERMANY
MOSCOW     RUSSIA
dtype: object

>>>srcapcountry['NEW DELHI':'BERLIN']  # elements  at index position TOKYO to BERLIN

NEW DELHI      India
TOKYO         JAPAN
BERLIN       GERMANY
dtype: object
```

To display the series in reverse order:

```
>>>srcapcountry[::-1]

MOSCOW       RUSSIA
BERLIN       GERMANY
TOKYO         JAPAN
NEW DELHI     INDIA
dtype: object

>>>srcapcountry[0:3:2]  #step is 2 here (applicable for positional index only)

NEW DELHI     INDIA
BERLIN       GERMANY
dtype: object
```

Updating the values in a series:

```
>>>import numpy as np
>>>import pandas as pd
>>>seriesAlph = pd.Series(np.arange(1,7),index = ['a', 'b', 'c', 'd', 'e', 'f'])
>>>seriesAlph
a   1
b   2
c   3
d   4
e   5
dtype: int32

>>>seriesAlph['a']=10
>>>seriesAlph  #value at index position 'a' is modified
a   10
b    2
c    3
d    4
e    5
f    6
dtype: int32

>>>seriesAlph[2]=30  #value at index position 2 is modified
>>>seriesAlph
a   10
b    2
```

```
c   30
d    4
e    5
f    6
dtype: int32

>>>seriesAlph[1:3]=25  #values at index position 1 to 2 are modified (slicing method)
>>>seriesAlph
a   10
b   25
c   25
d    4
e    5
f    6
dtype: int32

>>>seriesAlph[[1,3]]=50  #values at index position 1 and 3 are modified (specific index)
>>>seriesAlph
a   10
b   50
c   25
d   50
e    5
f    6
dtype: int32
```

1.3 Mathematical Operations on series:

- The following mathematical operations can be performed on two series:
  Addition (+), subtraction (-), multiplication (*) and division (/)
- Other operations : floor division (//), remainder (%), exponentiation (**)
- The operation is done on each corresponding pair of elements
- While performing mathematical operations on series, index matching is implemented and all missing values are filled in with **NaN** by default.
- Equivalent operations for +, -, *, / are : s1.add(s2), s1.sub(s2), s1.mul(s2) and s1.div(s2), where s1, s2 are series

Let us consider the following series: ser1, ser2 and ser3

```
>>>import pandas as pd
>>>ser1=pd.Series([10,20,30],index=['a','b','c'])
>>>ser2=pd.Series([4,5,6],index=['a','b','c'])
>>>ser3=pd.Series([2,3,4],index=['a','c','d'])
```

```
>>>ser1+ser2  # Equivalent statement ser1.add(ser2)
a   14
b   25
c   36
dtype: int64

>>>ser1+ser3
a   12.0
b    NaN
c   33.0
d    NaN
dtype: float64

>>>ser1/ser3
a    5.0
b    NaN
c   10.0
d    NaN
dtype: float64

>>>ser2//ser3   #floor division
a   2.0
b   NaN
c   2.0
d   NaN
dtype: float64

>>>ser1.mul(ser2)
a    40
b   100
c   180
dtype: int64

>>>ser1%ser2
a   2
b   0
c   0
dtype: int64
```

```
>>>ser1**ser3
a     100.0
b       NaN
c   27000.0
d       NaN
dtype: float64
```

1.4 Methods of Series: head(), tail()

```
>>>import pandas as pd
>>>import numpy as np
>>>ser1=pd.Series(np.arange(5,20,2))
>>>ser1
0    5
1    7
2    9
3   11
4   13
5   15
6   17
7   19
dtype: int32
```

note:- arange() is used to used to generate an array with evenly spaced values within a specified interval. It is very similar to range in python.

Methods	Explanation	Example
head(n)	Returns the first n members of the series. If the value for n is not passed, then by defaultn takes 5 and the first five members aredisplayed.	>>>ser1.head(3) 0 5 1 7 2 9 dtype: int32 >>>ser1.head() 0 5 1 7 2 9 3 11 4 13 dtype: int32
tail()	Returns the last n members of the series. If the value for n is not passed, then by defaultn takes 5 and the last five members aredisplayed.	>>>ser1.tail(2) 6 17 7 19 dtype: int32 >>>ser1.tail()

		3 11 4 13 5 15 6 17 7 19 dtype: int32

15 Objective Question (1 Mark)

Q1.	Which of the following statement is not correct for Pandas? (a) Pandas is open source built in library (b) Pandas offers high-performance, easy to use data structures (c) Pandas provides tools for backup and recovery (d) Pandas provides tools for data analysis
Ans	(c) Pandas provides tools for backup and recovery
Q2.	Which of the following statement will import pandas library? (a) Import pandas as pd (b) import Pandas as py (c) import pandas as pd (d) import panda as pd
Ans	(c) import pandas as pd
Q3.	You can not create a pandas Series using: (a) List and Dictionary (b) List and Tuple (c) Array and string (d) Dataframe
Ans	(d) Dataframe
Q4.	To display last five rows of a series object 'S', you may write: (a) S.Head() (b) S.Tail(5) (c) S.Head(5) (d) S.tail()
Ans	(d) S.tail()
Q5.	What type of error is returned by the following statement: import pandas as pd print(pd.Series([1,2,3,4],index=['a','b','c'])) (a) ValueError(b) SyntaxError (c) NameError(d) LogicalError
Ans	(a) ValueError
Q6.	In Python Pandas, while performing mathematical operations on series, index matching is implemented and all missing values are filled in with _____by default. (a) Null (b) Blank (c) NaN (d) Zero
Ans	(c) NaN
Q7.	To create an empty Series object, you can use: (a) pd.Series(NaN)

	(b) pd.Series(empty) (c) pd.Series(np.NaN) (d) pd.Series(None)
Ans	(d) pd.Series(None)
Q8.	To display First 15 rows of a series object 'Ser1', you may write: (a) Ser1.Head(15) (b) Ser1.head() (c) Ser1.head(15) (d) Ser1.Head()
Ans	(c) Ser1.head(15)
Q9.	Which of the following is the correct statement to access index 3rd and 5th values using positional index for series S? (a) S[3,5] (b) S[[3,5]] (c) S[3:5] (d) S[[3],[5]]
Ans	(b) S[[3,5]]
Q10.	A series contains a total of 10 elements including a missing value. The output of count and len functions when applied to this Series respectively will be (a) 10,10 (b) 9,10 (c) 9,9 (d) 10,9
Ans	(b) 9,10
Q11	PANDAS stands for _____ (a) Panel Data Analysis (b) Panel Data analyst (c) Panel Data (d) Panel Dashboard
Ans	(c) Panel Data
Q12	Pandas Series can have _____ data types (a) float (b) integer (c) String (d) All of the above
Ans	(d) All of the above
Q13	A _____ is a collection of data values and operations that can be applied to that data. (a) Data Structure (b) Data Frame (c) Table (d) None of the above
Ans	(a) Data Structure

Q14	When you print/display any series then the left most column is showing _____ value. (a)Data (b)Index (c) Value (d) None of the above
Ans	(b) Index
Q15	When we create a series from dictionary then the keys of dictionary become _____ (a) Index of the series (b) Value of the series (c) Caption of the series (d) None of the series
Ans	(a) Index of the series
05 Assertion and reason Based question (1 Mark)	
Q1.	Assertion (A):- To use the Pandas library in a Python program, one must import it. Reasoning (R): - The only alias name that can be used with the Pandas library is pd. i. Both A and R are true and R is the correct explanation for A ii. Both A and R are true and R is not the correct explanation for A iii. A is True but R is False iv. A is false but R is True
Ans	iii. A is True but R is False
Q2.	Assertion (A) : Pandas is an open-source Python library which offers high performance,easy-to-use data structures and data analysis tools. Reason (R) : Professionals and developers are using the panda's library in data scienceand machine learning. i. Both A and R are true and R is the correct explanation for A ii. Both A and R are true and R is not the correct explanation for A iii. A is True but R is False iv. A is false but R is True
Ans	i. Both A and R are true and R is the correct explanation for A
Q3.	Assertion (A):- A series object can be created by calling the Series() method. Reason (R):- A Series cannot be created from a Dataframe but a Dataframe can be created from a Series. i. Both A and R are true and R is the correct explanation for A ii. Both A and R are true and R is not the correct explanation for A iii. A is True but R is False iv. A is false but R is True
Ans	ii. Both A and R are true and R is not the correct explanation for A
Q4.	Assertion (A):We cannot access more than one element of Series without slicing . Reason (R):More than one element of series can be accessed using a list of positional index

	or labeled index. i. Both A and R are true and R is the correct explanation for A ii. Both A and R are true and R is not the correct explanation for A iii. A is True but R is False iv. A is false but R is True
Ans	iv. A is false but R is True
Q5.	Assertion (A): We can add two series objects using addition operator(+) or calling explicit function add() . Reason (R): While adding two series objects index matching is implemented and missing values are filled with NaN by default. i. Both A and R are true and R is the correct explanation for A ii. Both A and R are true and R is not the correct explanation for A iii. A is True but R is False iv. A is false but R is True
Ans	i. Both A and R are true and R is the correct explanation for A
05 Short Knowledge/Understanding/Application Based Questions (2 Marks)	
Q1.	Write a python program to create a series object, country using a list that stores the capital of each country. Assume four countries to be used as index of the series object are India, UK, Denmark and Thailand having their capitals as New Delhi, London, Copenhagen, and Bangkok respectively.
Ans	import pandas as pd country=pd.Series(['New Delhi','London','Copenhagen','Bangkok'], index=['India','UK','Denmark','Thailand']) print(country) **output:** India New Delhi UK London Denmark Copenhagen Thailand Bangkok dtype: object
Q2.	>>>import pandas as pd >>>A=pd.Series(data=[35,45,55,40]) >>>print(A>45)
Ans	0 False 1 False 2 True 3 False dtype: bool
Q3.	Given two series S1 and S2: S1 S2

	A 39 A 10 B 41 B 10 C 42 D 14 D 44 F 11 Find the output for following python pandas statements? a. S1[: 2]*100 b. S1 / S2
Ans	a. A 3900 B 4100 dtype: int64 b. A 3.900000 B 4.100000 C NaN D 3.142857 F NaN dtype: float64
Q4.	import pandas as pd ser1=pd.Series([10,20,30],index=['a','b','c']) ser2=pd.Series([2.5,5.5,8.5],index=['a','c','d']) print(ser1//ser2) print(ser1+ser2>=15)
Ans	a 4.0 b NaN c 5.0 d NaN dtype: float64 a False b False c True d False dtype: bool
Q5.	Consider the following code and predict the output: import pandas as pd s1=pd.Series(range(1,8,2), index=list("abcd")) print(s1[::2])

Ans	a 1 c 5 dtype: int64
05 Short Knowledge/Understanding/Application Based Questions (3 Marks)	
Q1.	What is a Series and how is it different from a 1-Darray and a list?
Ans	A Series is a one-dimensional array having a sequence of values of any data type (int, float, list, string, etc). **Series vs 1-D array** Series can have default as well as predefined index labels whereas a numpy 1-d array has only default indexes Series can contain values of any datatype whereas arrays can contain elements of the same data type **Series vs List** Series can have default as well as predefined index labels whereas a list has only default indexes
Q2.	Create a series EngAlph, having 26 elements with the alphabetsas values and default index values.Display the alphabets 'e' to 'p' from the SeriesEngAlph.
Ans	import pandas as pd EngAlph=pd.Series(['a','b','c','d','e','f','g','h','i','j','k','l','m','n','o', 'p', 'q', 'r', 's', 't', 'u' ,'v', 'w', 'x','y','z']) print(EngAlph) print(EngAlph[4:16]) OR print(EngAlph.iloc[4:16])
Q3.	Create a series Friends, from a dictionary having roll numbers offive of your friends as data and their first nameas keys. Display the 3rd and 2nd value of the Series Friends, in that order.
Ans	import pandas as pd d={"Samir":1,"Manisha":2,"Dhara":3,"Shreya":4,"Kusum":5} friends=pd.Series(d) print(friends) print(friends[2:0:-1])
Q4.	Create a series MonthDays, from a numpy array having thenumber of days in the 12 months of a year. Thelabels should be the month numbers from 1 to 12. (a) Display the names of the months 3 through 7from the Series MonthDays. (b) Display the Series MonthDays in reverse order

Ans	import pandas as pd import numpy as np Month=np.array([31,28,31,30,31,30,31,31,30,31,30,31]) Monthdays=pd.Series(Month,index=np.arange(1,13)) print(Monthdays) (a) print(Monthdays[2:7]) (b) print(Monthdays[::-1])
Q5.	Create a series serieseven from an ndarray which contains all even numbers up to 20. (a) Display the first 8 values in the series serieseven. (b) Display the last 5 values in the series serieseven. (c) Display the total number of values present in the series serieseven.
Ans	import pandas as pd import numpy as np areven=np.arange(2,21,2) serieseven=pd.Series(areven) print(serieseven) (a) print("First 8 values:\n",serieseven.head(8)) (b) print("Last 5 values:\n",serieseven.tail()) (c) print("Total number of values:",len(serieseven))
05 Short Knowledge/Understanding/Application Based Questions (4 Marks)	
Q1.	(i) Create a series vowel1 having 5 elementswith index labels ‘a’, ‘e’, ‘i’, ‘o’ and ‘u’ having values[3, 4, 7, 5, 9] respectively. Create another series vowel2 having 5 elementswith index labels ‘a’, ‘e’, ‘i’, ‘o’ and ‘u’ having values[2,5,6,3,8] respectively. (ii) Subtract, Multiply and Divide vowel1 by vowel2. (iii) Set all the values of vowel1 to 0 and print vowel1.
Ans	(i) import pandas as pd vowel1=pd.Series([3,4,7,5,9],index=['a','e','i','o','u']) vowel2=pd.Series([2,5,6,3,8],index=['a','e','i','o','u']) print(vowel1) print(vowel2) (ii) print(vowel1-vowel2) print(vowel1*vowel2) print(vowel1/vowel2) (iii)

	vowel1[:]=0 print(vowel1)
Q2.	Create a series Hum_stream which contains 5 subjects as indexes and their average marks as values given as follows: indexes -> English, History, Geography, Economics, IP avg -> 78.85, 79.56, 81.26, 74.50, 77.58 (i) Add a subject PolSc with average marks 85.45 (ii) Display the average marks of first and last subject (iii) Display the first 5 subjects' average marks.
Ans	import pandas as pd Hum_stream=pd.Series([78.85, 79.56, 81.26, 74.50, 77.58], index=['English', 'History', 'Geography','Economics','IP']) print(Hum_stream) (i) Hum_stream['PolSc']=77.58 print(Hum_stream) (ii) print(Hum_stream[0],Hum_stream[-1]) (iii) print(Hum_stream.head())
Q3.	Create a Series weekdays, which contains name of the days as values starts from Monday. The index of the series will be number 1 to 7. (i) Display the alternate days start from Monday. (ii) Modify the Sunday to 'Weekend' (iii) Display all the days except last element i.e. Weekend
Ans	import pandas as pd import numpy as np weekdays=pd.Series(['Monday','Tuesday','Wednesday','Thursday', 'Friday','Saturday','Sunday'],index=np.arange(1,8)) print(weekdays) print(weekdays[0::2]) weekdays[7]='Weekend' print(weekdays) print(weekdays[0:len(weekdays)-1])

Q4.	What will be the output of the following code:
	>>>import pandas as pd
	>>>rollno=[1,2,3,4,5,6]
	>>>marks=[23,86,74,11,98,75]
	>>>s=pd.Series(marks, index = rollno)
	>>>print(s[s>75])
	>>> print(s[1:6:2])
	>>> print(s[::-2])
	>>> print(s//2)

Ans	2 86 5 98 dtype: int64 2 86 4 11 6 75 dtype: int64 6 75 4 11 2 86 dtype: int64 1 11 2 43 3 37 4 5 5 49 6 37 dtype: int64
Q5.	Consider the following Series object, 'company' and its profit in Crores: TCS 350 Reliance 200 L&T 800 Wipro 150 Write python statements to perform the following: (a) To display the name of the companies having profit>250. (b) To display the profit of 'Reliance'. (c) To increase the profit by 10% for all companies. (d) To add a company 'BSNL' with profit of 100 Crores.
Ans	import pandas as pd profit=[350,200,800,150] idx=['TCS','Reliance','L&T','Wipro'] company=pd.Series(profit,index=idx) print(company) print('-------') print(company[company>250]) print('-------') print(company['Reliance']) print('-------') company=company+.10*company print(company) print('-------') company['BSNL']= 100

print(company)

05 Case Based Questions (5 Marks)

Q1. Raman is working for a school 'DPS Gurgaon'. He wants to store number of students in class 11 and 12 in three streams ('Science', 'Commerce' and 'Humanities') by using two series objects s11 and s12 and wants to perform the following operations.

Class	Science	Commerce	Humanities
11	150	120	80
12	148	115	72

Help him to do the following:
(a) To find total number of students in class 11 streamwise.
(b) To reduce5 number of students in class 11 science.
(c) To find the total number of students class 12 streamwise.
(d) To find the total number of students class 11.
(e) To find the total number of students in the school.

Ans

```
import pandas as pd
import numpy as np
s1=pd.Series([11,150,120,80],index=['Class','Science','Commerce','Humanities'])
s2=pd.Series([12,148,115,72],index=['Class','Science','Commerce','Humanities'])
print(s1)
print(s2)
##(a) To find total number of students in class 11 streamwise.
print('Science:',s1['Science'],'Commerce:',s1['Commerce'],'Humanities:',s1['Humanities'])
##(b) To reduce 5 number of students in class 11 science.
s1['Science']=s1['Science']-5
print('Science:',s1['Science'])
##(c) To find the total number of students class 12 streamwise.
print('Science:',s2['Science'],'Commerce:',s2['Commerce'],'Humanities:',s2['Humanities'])
##(d) To find the total number of students class 11.
print("Total no. of students Class 11:",s1['Science']+s1['Commerce']+s1['Humanities'])
##(e) To find the total number of students in the school.
tot=s1['Science']+s1['Commerce']+s1['Humanities']+s2['Science']+s2['Commerce']+s2['Humanities'
]
print('The total number of students in the school:',tot)
```

Q2. Mr. Ankit is working in an organisation as data analyst. He wants to create the following series 'Employee' with employee name as indices and their salaries as values. Help him to do that.

John	35000
Lucy	43000
Albert	29000
Ajay	37000
Manisha	41000

He also wants to perform the following operations on the series 'Employee'. Help him to do that.
(a) To display the salary of the employee 'John'.
(b) To increase the salary of employee 'Albert' by 10%
(c) To display the total salary paid by the organization.
(d) To display the salary of employees 'Albert' and 'Manisha'.
(e) To display the names and salaries of employees who are getting salary more than 35000.

Ans

```
import pandas as pd
Employee=pd.Series([35000,43000,29000,37000,41000],
```

```
                index=['John','Lucy','Albert','Ajay','Manisha'])
print(Employee)
##(a) To display the salary of the employee ‘John’
print("Salary of John:",Employee['John'])
##(b) To increase the salary of employee ‘Albert’ by 10%.
Employee['Albert']=Employee['Albert']*1.10
print(Employee)
##(c) The total salary paid by the organization.
print("Total salary Paid:",sum(Employee))
##(d) The salary of employees ‘Albert’ and ‘Manisha’.
print(Employee[['Albert','Manisha']])
print('------')
#(e) The names and salaries of employees who are getting salary more than 35000.
print(Employee[Employee>35000])
```

Name of Chapter: Pandas Dataframe

(Ref: Various websites and blogs on IP)

Topics Covered: Data Frames: creation - from dictionary of Series, list of dictionaries, Text/CSV files; display; iteration; Operations on rows and columns: add, select, delete, rename; Head and Tail functions; Indexing using Labels, Boolean Indexing , importing/Exporting Data between CSV files and Data Frames.

Key Points

DATAFRAME:It is a two-dimensional labelled data structure with columns of potentially different data types. It represents the data in the form of rows and columns. It is similar to a spreadsheet or an SQL table, or a dictionary of series objects and is one of the most used objects in Pandas.

Dataframe Structure:

COLUMN NAME →	ARTISTNAME	AWARDS	NO. OF AWARDS
0	TAYLOR SWIFT	ALBUM OF THE YEAR	3
1	BRUNO MARS	SONG OF THE YEAR	2
2	ED SHEERAN	RECORD OF THE YEAR	2

INDEX DATA

¤ Properties of a Dataframe:

1) A Dataframe has two indexes or we can say that two axes- • Row index (axis=0) • Column index (axis=1)

2) Conceptually it is like a spreadsheet where each value is identifiable with the combination of row index and column index. The row index is known as index and the column index is known as column-name.

3) A Dataframe contains heterogenous data, is size mutable and data mutable as well.

❖A Dataframe can be created using any of the following:

1) Series 2) Lists 3) Dictionaries 4) A numpy 2D array 5) Text files 6) CSV files

1) Creating an empty DataFrame- import pandas as pd df=pd.DataFrame()
print(df)

2) Creating DataFrame from Series- import pandas as pd s=pd.Series([1,2,3,4])
df=pd.DataFrame(s)
print(df)

3) Creating DataFrame from Dictionary ofSeries-
import pandas as pd name=pd.Series(['Shyam','Disha')
stream=pd.Series(['Science','Commerce'])
Dict={'Name':name,'Stream':stream} df1=pd.DataFrame(Dict)

```
print(df1)
```

4) Creating DataFrame from List ofDictionaries-

```
import pandas as pd IP=[{"Name":"Raj","Marks":33},
    {"Name":"Priya","Marks":29}] df=pd.DataFrame(IP)
print(df)
```

¤ Iteration on Rows and Columns:
In case we need to access any record or data from a Dataframe row-wise or column-wise, iteration can be used. Pandas provides us with 2 functions to perform iterations:
1) iterrows(): It is used to access the Dataframe row-wise.
2) interitem(): It is used to access the Dataframe column-wise.

i) iterrows( )- Used to access data row-wise.

EXAMPLE:

```
import pandas as pd
1=[{'Name':'Riya','Surname':'Verma'},{'Name':'Di
a','Surname':'Sen'}] D=pd.DataFrame(1)
print(D)
for(row_index,row_value)in D.iterrows(): print('\n Row
    Index is::',row_index) print('Row Value is::')
    print(row_value)
```

OUTPUT:

```
  Name    Surname
0 Riya    Verma
1 Dia     Sen

Row Index is::0

Row Value is::
Name      Riya Surname Verma
Name: 0, dtype:object
```

ii)iteritems( ) - Used to access data column wise.

EXAMPLE:

```
import pandas as pd
1=[{'Name':'Riya','Surname':'Verma'}
   {'Name':'Dia','Surname':'Sen'}] D=pd.DataFrame(1)
print(D)
for(col_name,col_value)in D.iteritems():
    print('\n')
    print ('Column Name is::',col_name)
    print('S column Values are::')
    print(col_value)
```

OUTPUT:

```
   Name     Surname
0  Riya     Verma
1  Dia      Sen

Column Name is:: Surname Column
Values are::
0  Verma
1  Sen
Name: Surname,dtype: object
```

¤ Different operations in Dataframe:
1) To access the data in the columns, we can mention the column name as subscript.
For example: “df[empid]” – This can also be done by using “df.empid”.
2) To access multiple columns, we can write it as- “df[[col1,col2…..]]”.
3) We can add columns of a Dataframe too:

```
df = pd.DataFrame({"D": [1, 2, 3], "E": [4, 5, 6]})
  f = [7,8,9]
df[‘F'] = f
```

In the same manner rows can be added as well.

Select operation inDataFrame-

- To access the column data,we can mention the column name assubscript.

*i)*To access singlecolumn:
Syntax-df ['col_name'] or, df.col_name
Eg-df ['rollno'] or, df.rollno

*ii)*To access multiplecolumns:
Syntax-df[[col1,col2,. .....]]
Eg-df [['rollno','sname']]

Adding/Modifying a Column andRow-

☆Columns in a dataframe can be used in multiple ways.Assigning a value to a column:

◇will modify it,if the column alreadyexists.

◇will add a new column,if it does not existalready.

□To change or add a new column,syntax:

df.columns=<new value> Or,
df[new col_name]=<new value>

Eg-

```
df.columns=['List1']
    Or,
df['List2']=10
```

Adding two DataFrames ,syntax:

df[new col_name]=df[col1]+df [col2]

Eg-

df ['List3']=df ['List1']+df ['List2']

Note: we can use append () function to add two DataFrames.

Syntax-

df1.append(df2)

□**Adding/Modifying a Row:**

We can change or add rows in Dataframe using at or loc attributes.

To change or add a row.syntax:

df.at [row_name,:]=<new value> Or,

df.loc [row_name,:]=<new value>

Eg-

df.at ['Bangalore',:]=1000 And,

df.loc ['Mohali',:]=[45200,56,211]

Note:

•while adding a row ,we have to make sure that the sequence containing values for different columns has values for all the columns,otherwise it will raise ValueError.

We can delete the columns and rows from a Dataframe by using any of the following:

i) del ii) pop() iii) drop()

i) del

Syntax-del df [col_name]

Eg-del df ['List 3']

ii) pop ()

Syntax-df.pop (col_name)

Eg- df.pop ('List2')

iii) drop()

Syntax for deleting the data column wise with example:

df1=df.drop('List2',axis=1)

Syntax for deleting the data row wise with example:

df2=df.drop (index=[2,3],axis=0)

Renaming Rows/Columns:

Syntax-

i)To renamerows-

Syntax- df.rename(index={<names dictionary>},inplace=False)

Eg-

Df.rename (index={'Sec A':'A' , 'Sec B':'B'})

ii) To renamecolumns-

Syntax- df.rename(columns={names dictionary},inplace=False)

Eg- Df.rename(columns={'Rollno':'Rno'})

Note:

i) For both **index** and **columns** arguments ,specify the names-change dictionary containing original names and the new names and the new names in a firm like {**old name:newname}**

ii) Specify **inplace**argument as **True** ,if we want to rename rows/columns in the samedataframe.

Accessing a Dataframe:

We can access a Dataframe through loc() and iloc() method or indexing using the following functions:

Pandas provides us with loc() and iloc() methods to access the subset from a Dataframe using row/column:

1) loc(): It is used to access a group of rows and columns- .loc[: , :]

2) iloc(): It is used to access a group of rows and columns based on numeric index value. .iloc[: , :]

• The above two syntaxes are generally used to access single/multiple rows/columns. There are other syntaxes as well which are used for accessing a particular type of subset such as single rows only, multiple rows only, etc.

1) Accessing the DataFrame through loc() and iloc() method - i)**loc()-**label based. Used to access group of rows and columns.**Syntax-**df.loc [Start Row : End Row,StartColumn:EndColumn]

Note-if we pass : in row or column part ,it means it will print the entire rows or columns respectively.

Eg-

We are using DataFrame dtf5:

	Population	Hospitals	Schools
Delhi	10927986	189	7916
Mumbai	12691836	200	8508
Kolkata	4631392	149	7226
Chennai	4328063	157	7617

```
>>>dtf5.loc['Delhi',:]
Population     10927986
Hospitals            189
Schools             7916

>>>dtf5.loc['Mumbai' : 'Chennai',:] Population
            Hospitals Schools
Mumbai   12691836      208      8508
Kolkata   4631392      149      7226
Chennai   4328063      157      7612
```

```
>>>dtf5.loc[:,'Population' : 'Hospitals']
            Population    Hospitals
Delhi       10927986      189
Mumbai      12691836      208
Kolkata     4631392       149
Chennai     4328063       157

>>>dtf5.loc['Delhi':
'Mumbai','Population': 'Hospitals']
            Population    Hospitals
Delhi       10927986      189
Mumbai      12691836      208
```

ii)**iloc()-**integer based.Used to access a group of rows and columns based on numeric index value.

Syntax-df.iloc [Start Rowindexes : End Row index, Start Columnindex : End Columnindex]

Note-Here,end result is excluded.

Eg-

We are using DataFrame dtf5: Population

```
          HospitalsSchools
Delhi     10927986    189       7916
Mumbai12691836      200       8508
Kolkata4631392      149       7226
Chennai4328063      157       7617

>>>dtf5.iloc[0:2,1:2]
          Hospitals
Delhi         189
 Mumbai       200

>>>dtf5.iloc [:,0:2]
          Population  Hospitals
Delhi     10927986    189
Mumbai    12691836    200
Kolkata   4631392     149
Chennai   4328063     157
```

¤ head() and tail() Method:

The method of head() gives the first 5 rows by default and the method tail() returns the last 5 rows by default. The syntax head(5) and tail(5) work the same and can be customizable as desired.

head() and tail() method-

i)**head()** -used to access first 5 rows

head(3)-used to access first 3rows.

ii)**tail()-** used to access last 5rows.

tail(3)-used to access last 3rows.

¤ Boolean Indexing in Dataframe: Boolean indexing, as the name suggests, means having Boolean values [(True or false) or (1 or 0) sometimes] as indexes of a Dataframe. While creating a Dataframe with Boolean indexing True and False, we make sure that True and False are not enclosed in quotes (i.e., like 'true' or 'false'), otherwise it will give us error(KeyError) while accessing data Boolean indexes using .loc, because 'True' and 'False' are string values, not Boolean values.

Boolean Indexing in DataFrame-

Used to select the data from the DataFrame using Boolean operators,i.e.True and False.

```
import pandas as pd
dict = {'name':["Deep","Rahul","Priya","Vinod"],
        'age': ["28", "39", "34", "36"]}
info = pd.DataFrame(dict, index = [True, True, False,True])
print(info)
print(info.loc[True])
print (info.iloc [2])
```

Output-

```
        name    age
True    Deep    28
True    Rahul   39
False   Priya   34
True    Vinod   36
        name    age
True    Deep    28
True    Rahul   39
True    Vinod   36
name    Priya
age     34
```

Some Imp.DataFrameAttributes:

I)index- display the index of theDataFrame.

II)columns- display the column labels of theDataFrame.

III)size-return an int representing the number of elements in this object. **IV)shape-** return a tuple representing the dimensionality of the DataFrame.

V)empty- indicate whether DataFrame isempty.

VI)ndim- return an int representing arraydimensions.

VII)T-to transpose index andcolumns.

(Transpose means to interchange the order of rows and columns in place of each other.)

Eg-

We take DataFrame dfn:

	Marketing	Sales
Age	25	24
Name	Neha	Rohit
Sex	Female	Male

```
>>>dfn.index
Index (['Age',Name','Sex'],dtype='object')

>>>dfn.columns
Index (['Marketing','Sales'],dtype='object')

>>>dfn.size                6

>>>dfn.shape(3,2)

>>>dfn.emptyFalse

>>>dfn.ndim2


>>>dfn.T
```

	Age	Name	Sex
Marketing	25	Neha	Female
Sales	24	Rohit	Male

15 Objective Question (1 Mark)

Q1.	Complete the following code –
	_____ #missing statement
	D = {'code' : [102 , 104, 105], 'ename' : ['Arun', 'Geet', 'Amy'] }
	df1 = pp.DataFrame(D)
	print(df1)
	a) import pandas
	b) import pandas as pp
	c) import Pandas as pp
	d) import pandas as pd

Ans	b) import pandas as pp
Q2.	Missing data in a Dataframe object is represented through – a) NULL b) None c) NaN d) <empty>
Ans	c) NaN
Q3.	The function to create a dataframe from a CSV file is – a) to_csv() b)load_csv() c) fetch_csv() d) read_csv()

Ans	d) read_csv()
Q4.	Method or function to add a new row in a data frame is: a. .loc() b. .iloc() c. join d. add()
Ans	a. .loc()
Q5.	While accessing the column from the data frame, we can specify the column name. In case column does not exist, which type of error it will raise: a. Key Error b. Syntax Error c. Name Error d. Runtime Error
Ans	a. Key Error
Q6.	What is a correct syntax to return the values of first row of a Pandas DataFrame? Assuming the name of the DataFrame is dfRent. a. dfRent[0] b. dfRent.loc[1] c. dfRent.loc[0] d. dfRent.iloc[1]
Ans	c. dfRent.loc[0]
Q7.	Difference between loc() and iloc().: a. Both are Label indexed based functions. b. Both are Integer position-based functions. c. loc() is label based function and iloc() integer position based function. d. loc() is integer position based function and iloc() index position based function.
Ans	c. loc() is label based function and iloc() integer position based function.
Q8.	Which command will be used to delete 3 and 5 rows of the data frame. Assuming the data frame name as DF. a. DF.drop([2,4],axis=0) b. DF.drop([2,4],axis=1) c. DF.drop([3,5],axis=1) d. DF.drop([3,5])
Ans	a. DF.drop([2,4],axis=0)
Q9.	Which attribute is not used with DataFrame? a.sizeb.type c,emptyd.column
Ans	d.column
Q10.	When we create a Dataframe from a list of Dictionaries the columns labels are formed by the a.union of the keys of the dictionaries b.intersection of the keys of the dictionaries c.union of the values of the dictionaries d.intersection of the values of the dictionaries
Ans	a.union of the keys of the dictionaries
Q11	The data of any CSV file can be shown in which of the following software? a.MS Word b.Notepad c,Spreadsheet d.All of the above
Ans	d.All of the above
Q12	Statement that displays first 5 rows of a DataFramedf: a.df[:5]b.df.head() c.df.iloc[:5] d.All of these
Ans	d.All of these
Q13	Choose the correct function to rename city columns to location using rename() function: a. df.rename(columns={'City':'Location'}) b. df.rename(columns={'City'='Location'}) c. df.rename('City'='Location') d. df.rename(df.columns('City','Location'))

Ans	a. df.rename(columns={'City':'Location'})
Q14	Which attribute is used to transpose the Data Frame ? a) Trans b)Transpose c) T d)Tp
Ans	c) T
Q15	Considering above DataFramedf which python command use to display all the records in the reverse order. a. print(df[::1]) b. print(df.iloc[::-1]) c. print(df[-1:]+df[:-1]) d. print(df.reverse())
Ans	b. print(df.iloc[::-1])
05 Assertion and reason Based question (1 Mark)	
Q1.	**Assertion** – DataFrame is a two-dimensional Pandas structure, with ordered collections of columns that can store data of different types. **Reason** - Dataframe is an array-like structure with two indices or axes – row index (axis = 0) and column index (axis=1). Dataframe is value-mutable as well as size- mutable with heterogeneous data. a) Assertion is True & Reason is correct explanation of Assertion b) Assertion is True, but Reason is partially True c) Assertion is True but Reason is False d) Both Assertion and Reason are False
Ans	a) Assertion is True & Reason is correct explanation of Assertion
Q2.	**Assertion** – Two basic data structure in Python are: Series and Dataframe. But both are different from each other. **Reason** - Series stores heterogenous data while Dataframe stores homogenous data. a)Assertion is True & Reason is correct explanation of Assertion b)Assertion is True, but Reason is partially True c)Assertion is True but Reason is False d)Both Assertion and Reason are False
Ans	a) Assertion is True but Reason is False
Q3.	Assertion (A):- DataFrame has both a row and column index. Reasoning (R): - A DataFrame is a two-dimensional labelled data structure like a table of MySQL. a)Assertion is True & Reason is correct explanation of Assertion b)Assertion is True, but Reason is partially True c)Assertion is True but Reason is False d)Both Assertion and Reason are False
Ans	a) Both A and R are true and R is the correct explanation for A
Q4.	Assertion: It is not possible to add a new row to an empty DataFrame, with no columns defined. Reason: DataFrame does not support adding of new rows a) Both the Assertion and the Reason are correct and the Reason is the correct explanation of the Assertion.

	b) The Assertion and the Reason are correct but the Reason is not the correct explanation of the Assertion. c) Assertion is true but the Reason is false. d) The statement of the Assertion is false but the Reason is True.
Ans	c) Assertion is true but the Reason is false.
Q5.	Statement 1 :-Labeled indexing uses rows and columns title to select data in the DataFrame Statement 2 :- In boolean indexing, we will select rows or columns based on the actual values of the data in the DataFrame a. Statement 1 is valid but statement 2 is invalid b. Statement 1 is invalid but statement 2 is valid c. Both statement are valid d. Both statements are invalid
Ans	c.Both statements are valid.
05 Short Knowledge/Understanding/Application Based Questions (2 Marks)	
Q1.	Write a python code to create a dataframe as per given structure in the figure.

	Country	Rank
0	Russia	121
1	Colombia	40
2	Chile	100
3	Equador	130
4	Nigeria	11

Ans	import pandas as pd data = pd.DataFrame({'Country': ['Russia','Colombia','Chile','Equador','Nigeria'], 'Rank':[121,40,100,130,11]}) print(data)
Q2.	The python code written below has syntactical errors. Rewrite the correct code and underline the corrections made. Import pandas as pd df ={"IP":["Programming","MySQL","Networking"],"Marks":[30,30,10]} df= Pd.dataframe(df) Print(df)
Ans	**import** pandas as pd df ={"IP":["Programming","MySQL","Networking"],"Marks":[30,30,10]} **df= pd.DataFrame(df)** **print(df)**
Q3.	Carefully observe the following code: import pandas as pd

	Year1={'Q1':5000,'Q2':8000,'Q3':12000,'Q4': 18000} Year2={'A' :13000,'B':14000,'C':12000} totSales={1:Year1,2:Year2} df=pd.DataFrame(totSales) print(df) Answer the following: i. List the index of the DataFramedf ii. List the column names of DataFramedf.
Ans	i. The index labels of df will include Q1,Q2,Q3,Q4,A,B,C ii. The column names of df will be: 1,2
Q4.	Consider the given DataFrame 'Stock': **Name Price** **0 Nancy Drew 150** **1 Hardy boys 180** **2 Diary of a wimpy kid 225** **3 Harry Potter 500** Write suitable Python statements for the following : i. Add a column called Special_Price with the following data: [135,150,200,440]. ii. Add a new book named 'The Secret' having price 800.
Ans	i. Stock['Special_Price']=[135,150,200,400] ii. Stock.loc['4']=['The Secret',800]
Q5.	What the following statements are doing? (I) df['city']=['Gwalior','Indore','Agra','Dewas','Gwalior','Indore'] (II) df.loc[2, :]
Ans	(i) creating a new column city with new data (ii) getting all columns of row index 2

05 Short Knowledge/Understanding/Application Based Questions (3 Marks)	
Q1.	Write a Python code to create a DataFrame with appropriate column headings from the list given below: [[1001,'IND-AUS','2022-10-17'], [1002,'IND-PAK','2022-10-23'], [1003,'IND-SA' , '2022-10-30], [1004,'IND-NZ','2022-11-18']]
Ans	import pandas as pd data=[[1001,'IND-AUS','2022-10-17'], [1002,'IND-PAK','2022-10-23'], [1003,'IND-SA' , '2022-10-30], [1004,'IND-NZ','2022-11-18']] df=pd.DataFrame (data, columns = ['MatchID', 'TEAMS', 'DATE']) print(df)
Q2.	Consider the given DataFrame 'Items':

	Name Price Quantity 0 CPU 7750 15 1 Watch 475 50 2 Key Board 225 25 3 Mouse 150 20 Write suitable Python statements for the following: i) Add a column called Sale Price which is 10% decreased value of Price ii) Add a new item named "Printer" having price 8000 and Quantity as 10. iii) Remove the column Quantity
Ans	i) Items['Sale_Price']=0.90 * Items['Price'] ii) Items.loc['4']=["Printer", 8000, 10] iii) Items=Items.drop('Quantity', axis=1)
Q3.	DataFrame 'STU_DF': rollno name marks 0 115 Pavni 97.5 1 236 Rishi 98.0 2 307 Preet 98.5 3 422 Paul 98.0 Perform the following operations on the DataFramestuDF: i. Add a new row in dataframe STU_DF with values [444,'karan',88.0] ii. Print no of rows and columns in dataframe STU_DF iii. Delete row for rollno 307.
Ans	i. STU_DF.loc[4,:]= [444,'karan',88.0] or STU_DF.loc[4]= [444,'karan',88.0] ii. print(STU_DF.shape) iii. STU_DF.drop(2, axis = 0)

Q4. Write a Python code to create a DataFrame 'Df' using dictionary of lists for the following data.

	Arnab	Ramit	Samridhi
Maths	90	92	89
Science	91	81	91
Hindi	97	96	88

Ans

```
import pandas as pd
D={'Arnab':[90,91,97],'Ramit':[92,81,96],'Samridhi':[89,91,88] }
Df=pd.DataFrame(D,index=['Maths','Science','Hindi'])
print(Df)
```

Q5. Consider the following dataframendf as shown below :

	Col1	Col2	Col3	Res
T1	62.893165	100.0	60.00	True
T2	94.734483	100.0	59.22	True
T3	49.090140	100.0	46.04	False
T4	38.487265	85.4	58.60	False

What will be the output produced by following statements :-

a. print(ndf.loc [: , 'Col3' :])

b. print(ndf.iloc[2 : , : 3])

c. print(ndf.iloc [1:3 , 2:3])

Ans

a.

	Col3	Res
T1	60.00	True
T2	59.22	True
T3	46.04	False
T4	58.62	False

b.

	Col1	Col2	Col3
T3	49.090140	100.0	46.04
T4	38.487265	85.4	58.62

c. Col3

T2 59.22

T3 46.04

05 Short Knowledge/Understanding/Application Based Questions (4 Marks)

Q1. Mr. Kapoor, a data analyst has designed the dataframe DF that contains data about Attendance and number of classes of a week as shown below. Answer the following questions:

	No_of_classes	Atten
Monday	5	15
Tuesday	8	24
Wednesday	4	20
Thursday	5	10
Friday	4	12
Saturday	8	16

A. Predict the output of the following python statement:

a. print(DF[3:])

b. print(DF.index)

B. Write Python statement to display the data of 'No_of_classes' column of indexes 'Tuesday' to 'Thursday'

OR (for option B only)

Write python statement to calculate No_of_classes * Atten and display it as Total attendance in a day.

Ans A. (a)

	Days	No_of_classes
3	Thursday	2
4	Friday	4
5	Saturday	0

(b)

```
Index(['Monday', 'Tuesday', 'Wednesday', 'Thursday', 'Friday', 'Saturday'], dtype='object')
```

B. print(df.loc['Tuesday': 'Thursday', 'No_of_classes'])

OR

```
print("Total Attendance in a day:")
print(df['Atten'] * df['No_of_classes'])
```

Q2. Mr. Roshan, a data analyst has designed the DataFramedf that contains data about sales made by 4 salesmans in two half yearlys as shown below. Answer the following questions:

	First	Second
Salesman1	23000	18000
Salesman2	11000	15000
Salesman3	60000	40000
Salesman4	35000	12000

A. Predict the output of the following python statement:

i. df.shape

ii. df[1:3]

B. Write Python statement to display the data of Second half yearly column of indexes Salesman1 to Salesman3

OR (Option for part iii only)

Write Python statement to compute and display the sum of data of First column and Second column of the above given DataFrame.

Ans

A i. (4, 2)

i. df[1:3]

```
            First   Second
Salesman2  11000   15000
Salesman3  60000   40000
```

B. print(df.loc['Salesman1': 'Salesman3', 'Second'])

OR

print(df.First+df.Second)

Q3. Mr. Sharma, a data analyst has designed the df that contains data about Cyber Olympiad details with 'Cyb1', 'Cyb2', 'Cyb3', 'Cyb4', 'Cyb5' as indexes shown below.

Answer the following questions:

	School	Total_students	Topper	Runnerup
Cyb1	KVS	50	46	4
Cyb2	NVS	45	35	10

Cyb3	DPS	30	20	10
Cyb4	MPS	25	18	7
Cyb5	GPS	37	28	9
Cyb6	BPS	33	25	8

A. Predict the output of the following –

i. df.shape

ii. df.[2:4]

B. Write python statement to display the data of the topper column of index Cyb2 to Cyb5.

OR

Write Python statement to compute and display the difference of data of Total_students column and Runnerup column of the above given DataFrame.

Ans

A. Output:
i. (5,4)

ii.

	School	tot_students	Topper	Runnerup
Cyb2	GPS	20	18	2
Cyb4	MPS	18	10	8

B. Python statement: print(df.loc['Cyb2': 'Cyb5','Topper'])

OR

A. print(df.Total_students-df.Runnerup)

Q4. Consider the following Data Frame sports:

	ID	NAME	GENDER
SD1	1	ARUNA	F
SD2	2	RAHUL	M
SD3	3	SAKSHI	F
SD4	4	RAMAN	M

(A)

i. Predict the output of the following python statement: sports. Shape
ii. Write command to Add a new row with values(5, SAM, M):

(B) Write python statement to delete column Gender.

	OR (Option for part iii only) Write python statement to delete the row with index SD3.
Ans	A. Output i. (4,3) ii Sports.loc['SD5']=[5, 'RAM', 'M'] B. del Sports['GENDER'] OR Sports.drop(index= 'SD3')) **2 mark for correct python statement**
Q5.	Mr. Ankit is working in an organization as data analyst. He uses Python Pandas and Matplotlib for the same. He got a dataset of the passengers for the year 2010 to 2012 for January, March and December. His manager wants certain information from him, but he is facing some problems. Help him by answering few questions given below: A. Predict the output of the following python statement: i. df.tail(3) ii. df[df. Passengers>50] iii.Write Python statement to display the data of year column of indexes 1 to 3. **OR (Option for part iii only)** Write the Python code to rename the name of the column name is 'mon_name'in place of "Month" in the above Dataframe.
Ans .	A) **i) 2 2012 Jan 35** **3 2010 Dec 55** **4 2012 Dec 65** **ii) 55 65** **B**)Python statement: print(df.loc[1: 3, 'year'])

	Year	Month	Passengers
0	2010	Jan	25
1	2010	Mar	50
2	2012	Jan	35
3	2010	Dec	55
4	2012	Dec	65

OR

df.rename(columns={'Month':'mon_name')

05 Case Based Questions (5 Marks)

Q1. Mr. Ankit is working in an organisation as data analyst. He uses Python Pandas and Matplotlib for the same. He got a dataset of the passengers for the year 2010 to 2012 for January, March and December. His manager wants certain information from him, but he is facing some problems. Help him by answering few questions given below:

	Year	Month	Passengers
0	2010	Jan	25
1	2010	Mar	50
2	2012	Jan	35
3	2010	Dec	55
4	2012	Dec	65

Code to create the above data frame:

import pandas as _____ #Statement 1
data={"Year":[2010,2010,2012,2010,2012],"Month":["Jan","Mar","Jan","Dec","Dec"]
,"Passengers":[25,50,35,55,65]}

df=pd._____(data) #Statement 2 print(df)

i. Choose the right code from the following for statement 1.

i. pd ii. df iii. data iv. p

ii. Choose the right code from the following for the statement 2.

i. Dataframe ii. DataFrame iii. Series iv. Dictionary

iii. Choose the correct statement/ method for the required output: (5,3)

i. df.indexii. df.shape()
iii. df.shapeiv. df.size

iv. He wants to print the details of "January" month along with the number of passengers, Identify the correct statement:

	Month	Passengers
0	Jan	25
2	Jan	35

i. df.loc[['Month','Passengers']][df['Month']=='Jan']
ii. df[['Month','Passengers']][df['Month']=='Jan']
iii. df.iloc[['Month','Passengers']][df['Month']=='Jan']
iv. df(['Month','Passengers']][df['Month']=='Jan')

v. Mr. Ankit wants to change the index of the Data Frame and the output for the same is given below. Identify the correct statement to change the index..Mr. Ankit wants to change the index of the Data Frame and the output for the same is given below. Identify the correct statement to change the index.

	Year	Month	Passengers
Air India	2010	Jan	25
Indigo	2010	Mar	50
Spicejet	2012	Jan	35
Jet	2010	Dec	55
Emirates	2012	Dec	65

i. df.index[]=["Air India","Indigo","Spicejet","Jet","Emirates"]
ii. df.index["Air India","Indigo","Spicejet","Jet","Emirates"]
iii. df.index=["Air India","Indigo","Spicejet","Jet","Emirates"]
iv. df.index()=["Air India","Indigo","Spicejet","Jet","Emirates"]

Ans

Answer: (i) pd

Answer: (ii) DataFrame

Answer: (iii) df.shape

Answer: (iv) df[['Month','Passengers']][df['Month']=='Jan']

Answer: (v) df.index=["Air India","Indigo","Spicejet","Jet","Emirates"]

Q2. Sanyukta is the event incharge in a school. One of her students gave her a suggestion to use Python Pandas andMatplotlib for analysing and visualising the data, respectively. She has created a Data frame “SportsDay” to keeptrack of the number of First, Second and Third prizes

won by different houses in various events.

	House	First	Second	Third
0	Chenab	5	7	6
1	Ganges	10	5	4
2	Jamuna	8	13	15
3	Jhelum	12	9	12
4	Ravi	5	11	10
5	Satluj	10	5	3

i. Display the house names where the number of Second Prizes are in the range of 12 to 20.

a. df['Name'][(df['Second']>=12) and (df['Second']<=20)]

b. df[Name][(df['Second']>=12) & (df['Second']<=20)]

c. df['Name'][(df['Second']>=12) & (df['Second']<=20)]

d. df[(df['Second']>=12) & (df['Second']<=20

ii. Display all the records in the reverse order.

a. print(df[::1])

b. print(df.iloc[::-1])

c. print(df[-1:]+df[:-1])

d. print(df.reverse())

iii. Display the bottom 3 records.

a. df.last(3)

b. df.bottom(3)

c. df.next(3)

d. df.tail(3)

iv. Choose the correct output for the given statements: x=df.columns[:1] print(x)

a. 0 b. Name

c. First d. Error

v. Which command will give the output 24:

a. print(df.size) b. print(df.shape)

	c. print(df.index) d. print(df.axes)
Ans	Answer: c. df['Name'][(df['Second']>=12) & (df['Second']<=20)] Answer: b. print(df.iloc[::-1]) Answer: d. df.tail(3) Answer: b. Name Answer: a. print(df.size)
Q3.	Naman has created the following dataframe "Climate" to record the data about climatic conditions of four years. i. Which of the following code snippets will return the MaxTemp and Rainfall for year 2018 and 2019? a. Climate[['MaxTemp','Rainfall']][1:3] b. Climate['MaxTemp', 'Rainfall'][1:3] c. Climate.iloc[1:3] d. Climate.iloc[1:3,1:2] ii. Display the temperature difference between MaxTemp and MinTemp for all the rows in the dataframe Climate. a. Climate=Climate["MaxTemp"]-Climate["MinTemp"] b. print(Climate["maxt"]-Climate["mint"]) c. print(Climate["MaxTemp"]-Climate["MinTemp"]) d. print(Climate.Climate["MaxTemp"]-Climate["MinTemp"]) iii. To display 2 rows from the top in the dataframe, which of the following statement is correct: a. print (Climate.head()=2) b. print (Climate.head(n==2)) c. print (Climate.head(range(2))) d. print (Climate.head(2))

Year	MaxTemp	MinTemp	Rainfall
2017	32	20	123
2018	33	22	140
2019	35	21	135
2020	34	23	160

	. iv. Which of the following statement/s will give the exact number of values in each column of the dataframe? a. print(Climate.count()) b. print(Climate.count(0)) c. print(Climate.count) d. print(Climate.count(axis='index')) Choose the correct option: a) both (a) and (b) b) only (b) c) (a), (b) and (c) (d) (a), (b) and (d) v. To display 2 rows from the bottom in the dataframe, which of the following statement is correct: a. print (Climate.tail()=2) b. print (Climate.tail(n==2)) c. print(Climate.tail(range(2))) d. print (Climate.tail(2))
Ans	Answer: (c) Climate.iloc[1:3] Answer : (c) print(Climate["MaxTemp"]-Climate["MinTemp"]) Answer: (d) print (Climate.head(2)) Answer: (d) (a), (b) and (d) Answer: (d) print (Climate.tail(2))
Q4.	HR Department of ABCTech has created following dataframe to store data about salaries and bonus paid to their employees. import pandas as pd import numpy as np

d1={'EName':['Kavita', 'Sudha', 'Garima']],'Sal':[50000,60000,55000],

'Bonus':[3000,4000,5000]}

Df1=pd.DataFrame(d1)

Choose the python statement using suitable functions for the following tasks:

i. Display the columns Sal and Bonus

a. df1 [:Sal :Bonus]

b. df1.loc(['Sal','Bonus'])

c. df1.iloc(['Sal','Bonus'])

d.df1[['Sal','Bonus']]

ii. Display the details of employee Kavita.

a. df1[df1.EName='Kavita']

b. df1.loc[df1.EName=='Kavita']

c. df1.iloc[df1.EName=='Kavita']

d. df1[EName='Kavita']

iii. Display the details of the last employee.

a. Df1.tail(1) b. Df1.tail (-1)

c. Df1.head(n=1) d. Df1.head()

iv. Add a new column named 'Email' with the value "abc@gmail.com".

a. Df1['Email']= 'abc@gmail.com'

b. Df1[Email]='abc@gmail.com'

c. Df1.loc['Email']='abc@gmail.com'

d. Df1('Email')='abc@gmail.com'

v. Write a python statement to print the details of employees having Sal more than 50000

a. df1.Sal>=5000

b. df1[df1.Sal>=5000]

c. df1[df1.'Sal'>=5000]

	d.df1.iloc[df1.Sal>=5000]
Ans	Answer: (b) df1.loc(['Sal','Bonus']) Answer (b) df1.loc[df1.EName=='Kavita'] Answer: (a) Df1.tail(1) Answer : (a) Df1['Email']= 'abc@gmail.com Answer: (b) df1[df1.Sal>=5000]
Q5.	Zeenat has created the following data frame dataframe1 to keep track of data Rollno, Name, Marks1 and Marks2 for various students of her class where row indexes are taken as the default values:

Rollno	Name	Marks1	Marks2
1	Swapnil Sharma	30	50
2	Raj Batra	75	45
3	Bhoomi Singh	82	95
4	Jay Gupta	90	95

i. Which among the following option will give 90, 95 as output

a) print(max(dataframe1['Marks1','Marks2'])

b) print((dataframe1.Marks1.max(),(dataframe1.Marks2.max())))

c) print(max(dataframe1['Marks1'])

d) print(max(dataframe1['Marks2'])

ii. She needs to know the marks scored by Rollno 2. Help her to identify the correct set of statement/s from the given options:

a. print(dataframe1[dataframe1['Rollno']= =2])

b. print(dataframe1['Rollno']= =2)

c. print(dataframe1[dataframe1. Rollno = =2])

d. print(dataframe1[dataframe1['Rollno']])

iii. Which of the following statement/s will delete the 3rd column?

a. del dataframe1['Marks1']

b. dataframe1.pop('Marks1')

	c. drop dataframe1['Marks1'] d. pop dataframe1['Marks1'] Choose the correct option: a) both (a) and (b) b) only (b) c) (a), (b) and (c) d) (a), (b) and (d) iv. Which of the following command will display the total number of elements in the dataframe? a. print(dataframe1.shape) b. print(dataframe1.num) c. print(dataframe1.size) d. print(dataframe1.elements) v. Now she wants to add a new column Marks3 with relevant data. Help her choose the command to perform this task. a. dataframe1.column=[45,52,90,95] b. dataframe1 ['Marks3']= [45,52,90,95] c. dataframe1.loc['Marks3']= [45,52,90,95] d. Both (b) and (c) are correct
Ans	Answer: (b) print((dataframe1.Marks1.max(),(dataframe1.Marks2.max()))) Answer : (c) print(dataframe1[dataframe1. Rollno = =2]) Answer : (a) both (a) and (b) Answer (c) print(dataframe1.size) Answer: (b) dataframe1 ['Marks3']= [45,52,90,95]

Name of Chapter: Data Visualization using Pyplot

Topics Covered: Purpose of plotting; drawing and saving following types of plots using Matplotlib – line plot, bar graph, histogram Customizing plots: adding label, title, and legend in plots.

Key Points

Data Visualization means representing the data in a graphical format which is easier to understand. For Data Visualization in Python we are using the Matplotlib library

Matplotlib

Matplotlib is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms. Matplotlib can be used in Python scripts, the Python and IPython shells, the Jupyter notebook, web application servers, and different graphical user interface toolkits.

Types of plots/charts (to be studied as per syllabus):

Some examples of charts are:

1. Line plots
2. Bar plot
3. Histograms

Working with matplotlib

For working with matplotlib usually we use the following import command:
import matplotlib.pyplot as plt

and frequently we need numpy for creating datasets, so numpy is also imported as follows:
import numpy as np

Basic Steps involved in drawing any plot

1. **Identify the data you want to represent on the plot.**
   For plots such as line graph it means identify the values that will be represented in the X-axis as well as Y-axis. For pie-charts, histograms etc. there will usually be only one dataset.
2. **Identify the structure of the plot you want**
   The next step is identifying which plot will be suitable to represent the data accurately. It can be line plot, bar plot, histogram etc. Also consider whether you want many sets of data to be represented in the same plot or to show different plots for different sets of data.
3. **Setup the different parameters of the plot**
   Each plot has different components such as the xticks, yticks, the shape/colour of markers/plots, legend etc. Set the parameters of the plot.
4. **Draw the plot.**

Line plots:

A line chart or line plot or line graph or curve chart is a type of chart which displays information as a series of data points called 'markers' connected by straight line segments.

A line chart is often used to visualize a trend in data over intervals of time – a time series – thus the line is often drawn chronologically.

1. For drawing any plot usually the matplotlib.pyplot is imported as plt
2. The plot function plt.plot() is used to draw line graphs drawing lines between any two successive values of x and y.
3. The plot function accepts two datasets, the first one is a list of x-coordinates and the second a list of corresponding y-coordinates. The number of values in both the x and y lists must be

same.

4. The plt.plot(x,y) is used to draw the graph and the plt.show() function is used to display the plot on the screen.

The various parameters that can be set are:

1. plt.xlabel('time') -xlabel sets the x-axis label
2. plt.ylabel('speed') -ylabel sets the y-axis label
3. plt.yticks([5,7,10]) -yticks sets the tick marks that appear on y-axis
4. plt.xticks([1,3,4],['abc','def','ghi']) -xticks sets the ticks to appear on the x-axis at points [1,3,4] the second parameter changes the corresponding labels to ['abc','def','ghi'].
5. plt.grid() - displays the gridlines
6. plt.legend() - displays the legend using the labels for the corresponding plots. legend is drawn only after the plot() is called since it takes the labels from plot function.
7. Besides the plot function can accept a third parameter the format string - *plt.plot(x,y,'>--c', label='car 1')*
   The format string (fmt) has the following specification: fmt = '[marker][line][color]'
   All the values are optional and the possible values for marker, line and color are shown below:

Marker	
character	description
'.'	point marker
','	pixel marker
'o'	circle marker
'v'	triangle_down marker
'^'	triangle_up marker
'<'	triangle_left marker
'>'	triangle_right marker
'1'	tri_down marker
'2'	tri_up marker
'3'	tri_left marker
'4'	tri_right marker
's'	square marker
'p'	pentagon marker
'*'	star marker
'h'	hexagon1 marker
'H'	hexagon2 marker
'+'	plus marker
'x'	x marker
'D'	diamond marker
'd'	thin_diamond marker
'|'	vline marker
'_'	hline marker

Line Style	
character	description
'-'	solid line style
'--'	dashed line style
'-.'	dash-dot line style
':'	dotted line style

Color	
character	color
'b'	blue
'g'	green
'r'	red
'c'	cyan
'm'	magenta
'y'	yellow
'k'	black
'w'	white

Example of format strings:

'b' # blue markers with default shape
'or' # red circles
'-g' # green solid line
'--' # dashed line with default color
'^k:' # black triangle_up markers connected by a dotted line

Bar plot:

A bar chart or bar graph is a chart or graph that presents categorical data with rectangular bars with heights or lengths proportional to the values that they represent. The bars can be plotted vertically or horizontally.

A bar graph shows comparisons among discrete categories. One axis of the chart shows the specific categories being compared, and the other axis represents a measured value.

Plotting a Bar Graph

The syntax for plotting a Bar Graph is:

plt.bar(*x, height, width=0.8, bottom=None, align='center', data=None*)

where:

x : sequence of scalars which form the x coordinates of the bars
height: sequence of scalars which form the heights of the bars.
width: scalar or array-like, optional which are the width(s) of the bars (default: 0.8)
bottom: scalar or array-like, optional. The y coordinate(s) of the bars bases (default: 0)
align: {'center', 'edge'}, optional, default: 'center'. It shows the alignment of the bars to the x coordinates:
'center': Center the base on the x positions.
'edge': Align the left edges of the bars with the x positions.
data: If the source of the data is another matrix like structure such as a DataFrame then the name of the object is mentioned here.

Horizontal Bar plot

To plot a horizontal bar plot use the barh() with the following syntax:

plt.barh(y, *width, height=0.8, align='center', data=None*)

where:
y : sequence of scalars which form the y coordinates of the bars
width: scalar or array-like, which are the width(s) of the bars on the x-axis
height: sequence of scalars which form the heights of the bars (default: 0.8)
align: {'center', 'edge'}, optional, default: 'center'. It shows the alignment of the bars to the y coordinates:
'center': Center the base on the x positions.
'edge': Align the left edges of the bars with the x positions.
data: If the source of the data is another matrix like structure such as a DataFrame then the name of the object is mentioned here.

Histogram

Histogram is a graphical display of data using bars of different heights to group numbers into ranges. The height of each bar shows how many of the data fall in that particular range.

A histogram is an accurate representation of the distribution of numerical data. It differs from a bar graph, in the sense that a bar graph relates two variables, but a histogram relates only one variable.

How to draw a Histogram

Example 1: For the dataset containing CGPA of 15 students shown below draw the histogram for bin size 10:

6.1, 4.12, 8.2, 6.4, 3.6, 9.2, 5.5, 8.4, 6.2, 9.8, 5.3, 3.9, 8.1, 6.1, 2.7

Step 1: Calculate the range of the data set
range = largest value - smallest value = 9.8 - 2.7 = 7.1

Step 2: Divide the range by the number of groups you want and then round up.
For example we want to divide the data set into 10 groups **(in python if bin size is not mentioned then 10 is taken as the default bin size)**, and then the width of each group is found by
class-width = range / number of groups = 7.1 / 10 =
0.71 Therefore **class width = 0.71**

Step 3: Use the class width to create your groups
The smallest value is 2.7 and class-width is 0.71, so **first class or first bin** is from 2.7 to (2.7 + 0.71) i.e.**from 2.7 to 3.41.**
The **second class or second bin** is from 3.41 to (3.41 +0.71) i.e. second bin is **3.41 to 4.12** and so on…

Drawing histogram using hist() function of pyplot

The hist() function can be used to draw a histogram. It accepts only a single dimensional 1D array or a list to draw the histogram. The other properties of plot object such as setting xlabels, ylabels, xticks, yticks etc. remain the same as line/bar plots.

In its simplest form histogram is drawn using the command:

plt.hist(data, bins=10)

where-

data - is the list or 1D array containing the data on which histogram is to be created

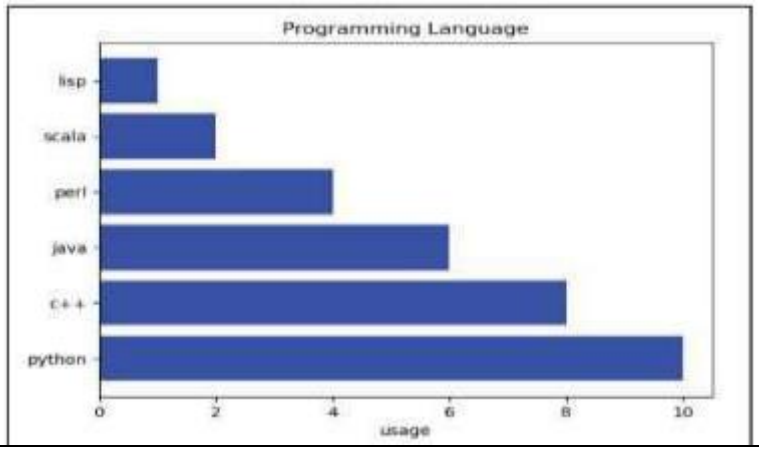
bins - it can either be a number or a list. If it is a single number it denotes the number of intervals of the histogram we want. If bins parameters is a list, then the elements of the list are the bin edges. The number of bin edges must be one greater than the number of intervals needed for the histogram. If bins parameter is not passed a default value of 10 is taken.

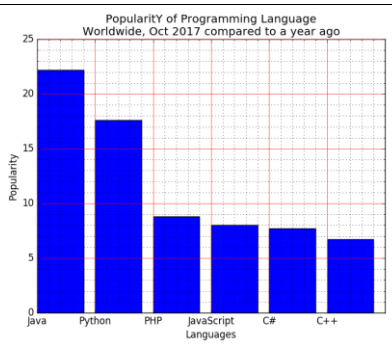
15 Objective Question (1 Mark)	
Q1.	**What is Matplotlib?** a) programming language b) A data visualization library c) A database management system d) An operating system
Ans	**b)**
Q2.	**What is the purpose of Matplotlib's pyplot module?** a)To create data visualizations b) To manage data storage c) To manipulate data frames d) To install third-party packages
Ans	**a)**
Q3.	**How can you add a title to a Matplotlib plot?** a) By using the title() function b) By using the label() function c) By using the text() function d) By using the legend() function
Ans	a
Q4.	**How can you save a Matplotlib plot as an image file?** a)By using the save() function b) By using the export() function c) By using the savefig() function d) By using the exportfig() function
Ans	c
Q5.	**What is the default color for Matplotlib plots?** a) Red b) Blue c) Green d) Black
Ans	b
Q6.	**How can you change the color of a Matplotlib plot?**

	a) By using the color() function b) By using the hue() function c) By using the palette() function d) By specifying the color parameter in the plot() function
Ans	d
Q7.	**What is the purpose of the legend() function in Matplotlib?** a) To label the x and y axes of a plot b) To add a title to a plot c) To add annotations to a plot d) To label different lines or markers on a plot
Ans	d
Q8.	**What is the purpose of the axis() function in Matplotlib?** a) To add grid lines to a plot b) To change the x and y limits of a plot c) To change the color of a plot d) To label different lines or markers on a plot
Ans	b
Q9.	**Which argument keyword can be used to emphasize each point with a specified marker in plotting?** a)marker_points b)marker c)ring d)types
Ans	b
Q10.	Which type of chart shows the relationship between a numerical variable and categorical variable? a. line b. bar c. pie d. x-yplot
Ans	b
Q11	Using pyplot matplotlib,_____can be used to count how many values fall into each interval. a. Histogram b. Pyplot c. Barchart d. Piechart
Ans	a
Q12	The_____argument of legend() provides the location of legends. a. loc b. Toc c. Goc d. None of these

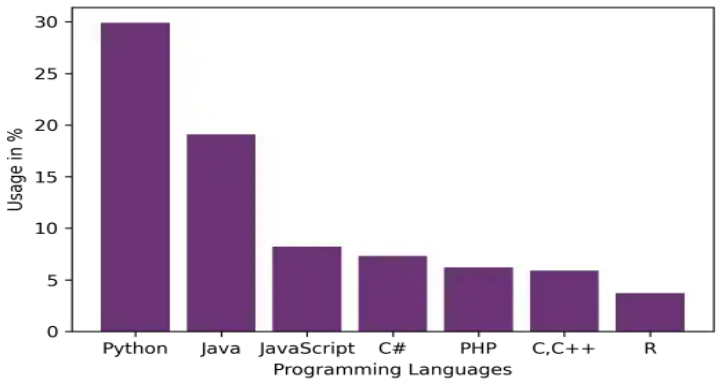
Ans	a
Q13	The _____ argument of hist() is set to create a horizontal histogram. a. landscape b. portrait c. documentation d. orientation
Ans	d
Q14	Which argument of bar() lets you set the thickness of bar? a. thick b. thickness c. width d. barwidth
Ans	c
Q15	A_____graph is a type of chart which displays information as a series of data points connected by straight line segment. a. line b. bar c. pie d. boxplot
Ans	a
05 Assertion and reason Based question (1 Mark)	
Q1.	Assertion(A): A histogram is a plot that shows the underlying frequency distribution of a set of continuous data. Reason(R): Pyplot interface is a collection of methods with in matplotlib library of python.
Ans	b
Q2.	Assertion(A): Pyplot plot() function is used to create line charts. Reason(R): Pyplot'sbarh() function is used to create horizontal bar graph.
Ans	b
Q3.	Assertion(A): The data point plotted on a graph are called markers. Reason(R):The width argument of plot() specifies the width of the line.
Ans	c
Q4.	Assertion(A):The line style argument of plot specifies the style of the line. Reason(R):The line argument of bar() specifies the bar width.
Ans	c
Q5.	Assertion(A):Pyplot is a collection of methods with in matplotlib library which allows user to construct 2D plots easily and interactively. Reason(R):A histogram is a statistical tool used to summaries discrete or continuous data. It provides a visual interpretation of numerical data by showing the number of data points that fall within a specified range of values (called bins).
Ans	b
05 Short Knowledge/Understanding/Application Based Questions (2 Marks)	
Q1.	What is data visualization?

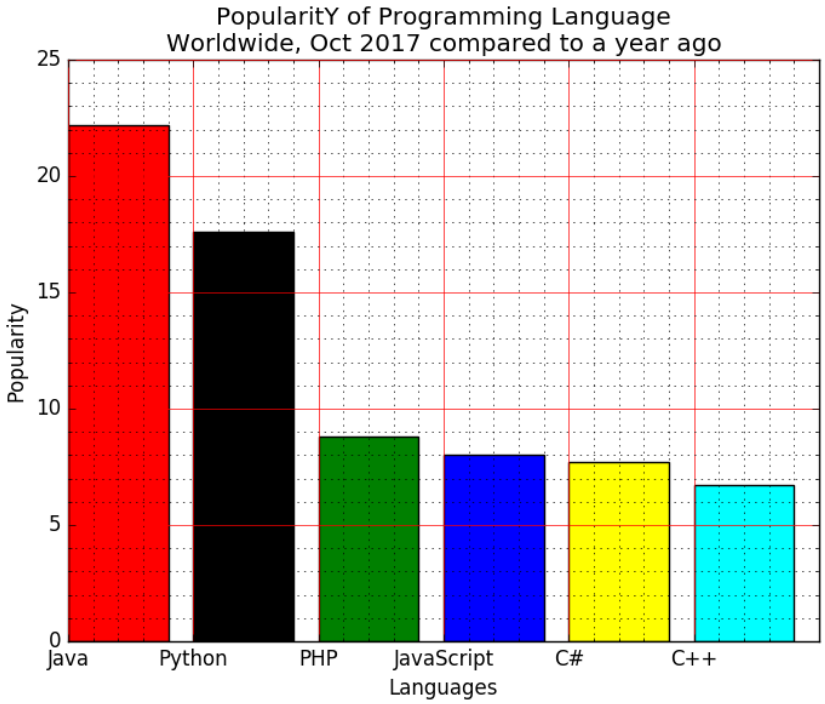
Ans	Data visualization is a general term that describes any effort to help people understand the significance of data by placing it in a visual context. In simple words, Data visualization is the process of displaying data/information in graphical charts, figures and bars.
Q2.	What is the significance of data visualization ?
Ans	Patterns, trends and correlations that might go undetected in text-based data can be exposed and recognized easier with data visualization techniques or tools such as line chart, bar chart, pie chart, histogram, scatter chart etc. Thus with data visualization tools, information can be processed in efficient manner and hence better decisions can be made.
Q3.	What is pyplot? Is it a Python library?
Ans	The pyplot is one of the interfaces of matplotlib library of Python. This interface offers simple MATLAB style functions that can be used for plotting various types of charts using underlying matplotlib library's functionality. Pyplot is an interface, i.e., a collection of methods for accessing and using underlying functionality of a library, not a library. The matplotlib library has may other interfaces too, along with pyplot interface.
Q4.	**What is the purpose of a legend?**
Ans	A legend is an area describing the elements of the graph. In the matplotlib library, there's a function called legend() which is used to Place a legend on the axes. The attribute Loc in legend() is used to specify the location of the legend. Default value of loc is loc="best" (upper left). The strings 'upper left', 'upper right', 'lower left', 'lower right' place the legend at the corresponding corner of the axes/figure.
Q5.	**What is the difference between a bar plot and a histogram in Matplotlib?**
Ans	1. Histogram refers to a graphical representation; that displays data by way of bars to show the frequency of numerical data. A bar graph is a pictorial representation of data that uses bars to compare different categories of data. 2. A histogram represents the frequency distribution of continuous variables. Conversely, a bar graph is a diagrammatic comparison of discrete variables. 3. Histogram presents numerical data whereas bar graph shows categorical data. 4. The histogram is drawn in such a way that there is no gap between the bars. On the other hand, there is proper spacing between bars in a bar graph that indicates discontinuity. 5. Items of the histogram are numbers, which are categorised together, to represent ranges of data. As opposed to the bar graph, items are considered as individual entities. 6. In the case of a bar graph, it is quite common to rearrange the blocks, from highest to lowest. But with histogram, this cannot be done, as they are shown in the sequences of classes. 7. The width of rectangular blocks in a histogram may or may not be same while the width of the bars in a bar graph is always same.
05 Short Knowledge/Understanding/Application Based Questions (3 Marks)	
Q1.	**Write a Python program to display a bar chart of the number of students in a school.** Sample data: Class: I,II,III,IV,V,VI,VII,VIII,IX,X Strength: 38,30,45,49,37,53,48,44,36,46
Ans	import matplotlib.pyplot as plt a = ['I','II','III','IV','V','VI','VII','VIII','IX','X'] b = [38,30,45,49,37,53,48,44,36,46]

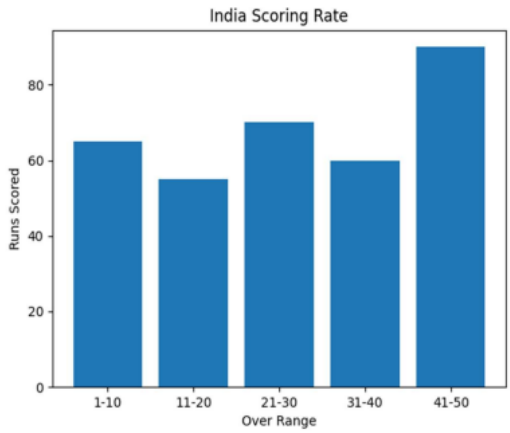
	plt.bar(a,b) plt.xlabel() plt.show()
Q2.	**Write a Python program to plot the given bar graph to depict the popularity of various programming languages. Label the graph with x-axis, y-axis, y-ticks and title.** Data : Programming languages: Python, C++, Java, Perl, Scala, Lisp Usage= 10,8,6,4,2,1 
Ans	import matplotlib.pyplot as plt a = [“Python”,” C++”, “Java”, “Perl”, “Scala”,” Lisp”] b = [10,8,6,4,2,1] plt.bar(a,b) plt.xlabel(‘usage) plt.show()
Q3.	What changes will you make to the code so that the bars are visible for all four points? But do keep in mind that the x-axis must begin from the point -3. import matplotlib.pyplot as plt a = [3, 6, 9, 12] b = [30, 48, 54, 48] plt.xlim(-3, 5) plt.bar(a,b) plt.show()
Ans	import matplotlib.pyplot as plt a = [3, 6, 9, 12] b = [30, 48, 54, 48] plt.xlim(-3, 15) plt.bar(a,b) plt.show()
Q4.	Write a Python programming to display a bar chart of the popularity of programming Languages.

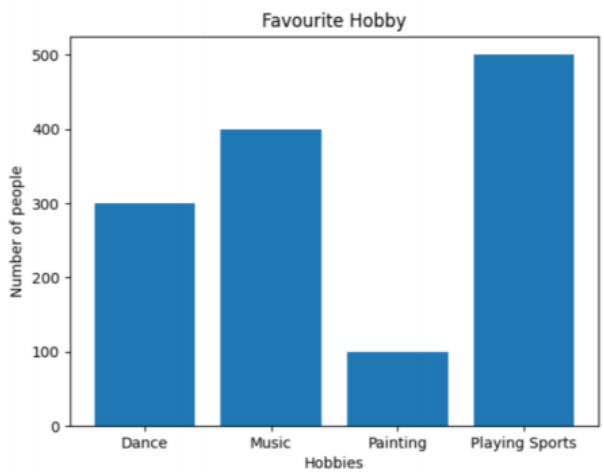
	
Ans	import matplotlib.pyplotas plt x =['Java','Python','PHP','JavaScript','C#','C++'] popularity =[22.2,17.6,8.8,8,7.7,6.7] plt.bar(x_pos,popularity,color='blue') plt.xlabel("Languages") plt.ylabel("Popularity") plt.title("Popularity of Programming Language\n"+"Worldwide, Oct 2017 compared to a year ago") plt.show()

Q5.	Write the output of following code.
	import matplotlib.pyplot as pyplot
	# Manual data setup
	labels = ('Python', 'Java', 'JavaScript', 'C#', 'PHP', 'C,C++', 'R')
	index = (1, 2, 3, 4, 5, 6, 7) # provides locations on
	x axis sizes = [29.9, 19.1, 8.2, 7.3, 6.2, 5.9, 3.7]
	# bar chart setup
	pyplot.bar(index, sizes, color="#6c3376", tick_label=labels)
	# layout configuration
	pyplot.ylabel('Usage in %')
	pyplot.xlabel('Programming Languages')
	# Save the chart file
	pyplot.savefig('filename.png', dpi=300)
	# Print the chart
	pyplot.show()

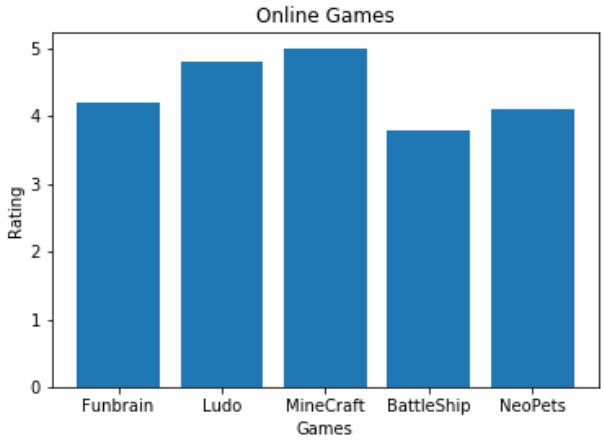
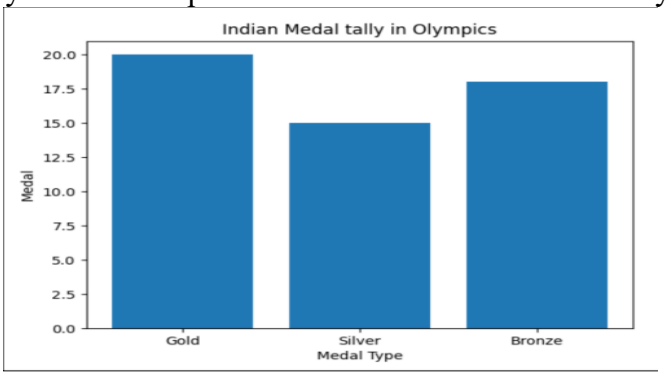
Ans	
05 Short Knowledge/Understanding/Application Based Questions (4 Marks)	
Q1.	Write a Python programming to display a bar chart of the popularity of programming

	Languages. Use different color for each bar. 
Ans	importmatplotlib.pyplotasplt x =['Java','Python','PHP','JavaScript','C#','C++'] popularity =[22.2,17.6,8.8,8,7.7,6.7] plt.bar(x_pos, popularity,color=['red','black','green','blue','yellow','cyan']) plt.xlabel("Languages") plt.ylabel("Popularity") plt.title("PopularitY of Programming Language\n"+"Worldwide, Oct 2017 compared to a year ago") plt.show()
Q2.	**A)**Mr. Anoy wants to plot a bar chart for the given set of values of Class on X axis and number of students in the class on Y axis. Write the appropriate python statement in order to complete the code to perform the following operations:- i. To plot the bar graph in statement 1. ii. To display the graph in statement 2. import matplotlib.pyplot as plt Class=[‘IX’, ‘X’, ‘XI’, ‘XII’] Std=[38, 42, 47, 40] _____ Statement 1 _____ Statement 2 B) Write Python code to plot a bar chart for India Scoring Rate as shown below:

<table>
<tr><td></td><td> 
</td></tr>
<tr><td>Ans</td><td>A)
plt.plot(Class,Std)Statement 1
plt.show() Statement 2

B)
import matplotlib.pyplot as plt
X=[‘1-10,’11-20’,’21-30’,31-40’,’41-50’]
Y=[65,55,70,60,90]
plt.bar(X,Y)
plt.xlabel("Over Range")
plt.ylabel("Runs Scored")
plt.title("Indian Scoring Rating")
plt.show()</td></tr>
<tr><td>Q3.</td><td>Write suitable Python code to create 'Favourite Hobby' Bar Chart as shown below:
 

Also give suitable python statement to save this chart.</td></tr>
<tr><td>Ans</td><td>import matplotlib.pyplot as plt
hobby = ('Dance', 'Music', 'Painting', 'Playing Sports')
users = [300,400,100,500]
plt.bar(hobby, users)
plt.title("Favourite Hobby")
plt.ylabel("Number of people")
plt.xlabel("Hobbies")</td></tr>
</table>

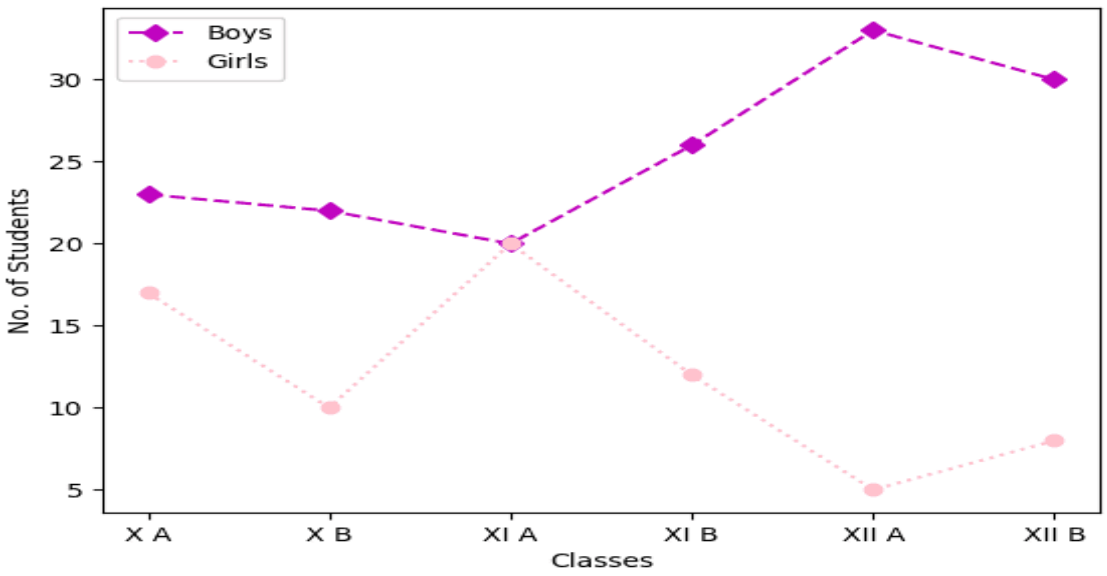
	plt.show() plt.savefig("hobbies.jpg")
Q4.	Suraj is working in a game development industry and he wants to compare the given chart on the basis of the rating of the various games available on the play store. Write a Python code to get the following bar chart. Save the chart in png format. 
Ans	import matplotlib.pyplot as plt Games=["Funbrain","Ludo","MineCraft","BattleShip","NeoPets"] Rating=[4.2,4.8,5.0,3.8,4.1] plt.bar(Games,Rating) plt.xlabel("Games") plt.ylabel("Rating") plt.title("Online Games") plt.savefig("OnlineGames.png") plt.show()
Q5.	Write Python code to plot a bar chart for India's medal tally as shown below:  Also give suitable python statement to save this chart.
Ans	import matplotlib.pyplot as plt X=['Gold,'Silver','Bronze'] Y=[20,15,17.5] plt.bar(X,Y) plt.xlabel("Medal Type") plt.ylabel("Medal ")

```
plt.title("Indian Medal Tally in Olampics")
plt.savefig("game.png")
plt.show()
```

05 Case Based Questions (5 Marks)

Q1. **Ms.Ekta is a coordinator in the senior section school. She represented data on number of students who passed the exam on line chart as follows:**

She has written the following code but not getting the desired output. Help her by correcting her code.



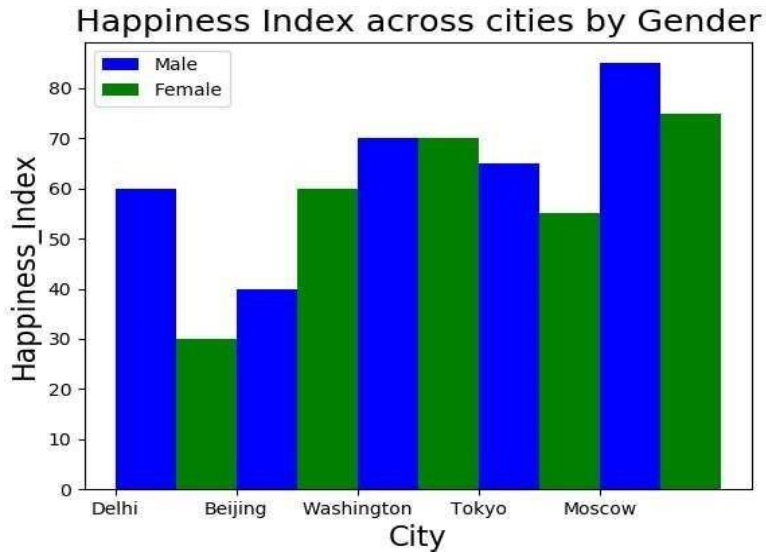

```
import matplotlib.pyplot as plt
classes=["X A","X B","XI A","XI B","XII A","XII B"]
no_of_boys=[23,22,20,26,33,30]
no_of_girls=[17,10,20,12,5,8]
plt.line(classes,no_of_boys) #Statement 1
plt.line(classes,no_of_girls) #Statement 2
plt.xtitle("No of Stduents") #Statement 3
plt.ytitle("Classes") #Statement 4
```

i) What will be the correct code for Statement 1 and Statement 2?
ii) What is the correct function name for Statement 3 and Statement 4?
iii) Write a method and parameter required to display legends?
iv) Write the code for giving the graph 'Classroom visualisation' as title.
v) Write to save the figure as image.

Ans
i)plt.plot
plt.plot
ii)plt.xlabel
plt.ylabel
iii)plt.legend()
iv)plt.title('Classroom visualisation')
v)plt.savefig('d:/pypl/ .jpg')

Q2. Gaurav has written a Python Code to create a bar plot as given below using the following data :

City	Happiness_IndexMale	Happiness_IndexFemale
Delhi	60	30
Beijing	40	60
Washington	70	70
Tokyo	65	55
Moscow	85	75



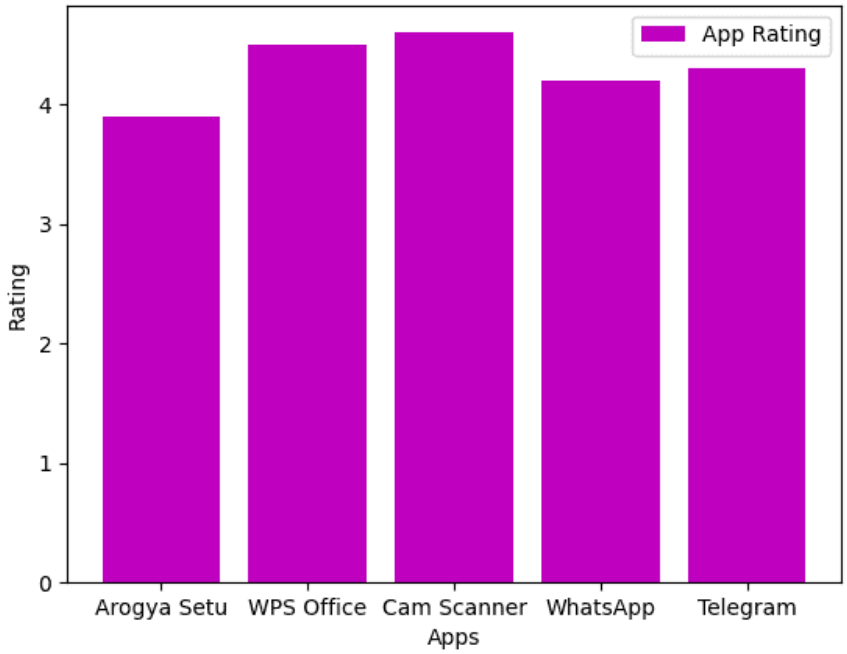
import _____as _____#Statement 1
City=['Delhi','Beijing','Washington','Tokyo','Moscow']
Gender=['Male','Female'] Happiness_Index_Male=[60,40,70,65,85]
Happiness_Index_Female=[30,60,70,55,75]
plt.bar([0.25,1.25,2.25,3.25,4.25],Happiness_Index_Male,color='blue',label="
Male",width=.5)
plt._____ ([.75,1.75,2.75,3.75,4.75],Happiness_Index_Female,color='Green',w
idth=.5,label="Female") #Statement 2
pos=range(len(City)) print(pos)
plt.xticks(pos,City,fontsize=10) plt.xlabel('City', fontsize=16)
plt.ylabel('Happiness_Index', fontsize=16)
_____ #Statement 3
_____ #Statement 4
_____ #Statement 5

i. Identify the suitable code to be used in the blank space in line marked as Statement1.
a.matplotlib as plt
b.numpy as np
c.pandas as pd
d.matplotlib.pyplot as plt

	ii. What is the name of the function to plot the required bar graph in the line marked as Statement 2 a. hist() b. pie() c. bar() d. scatter() iii. Fill in the blank in statement 3 to set Chart Title as “Happiness Index acrosscities by gender “and font size as 18. a. plt.xtitle(“Happiness Index across cities by gender”,fontsize=18) b. plt.title(“Happiness Index across cities by gender”,fontsize=18) c. plt.ytitle(“Happiness Index across cities by gender”,fontsize=18) d. plt.show(“Happiness Index across cities by gender”,fontsize=18) iv. Identify the suitable code for line marked as Statement 4 to display the legends as shown in the plot. a. plt.showlegend() b. plt.legend() c. plt.display() d. plt.show() v. Fill in the blank marked in Statement 5 to display the plot. a.plt.plot() b. plt.showplot() c.plt.display() d. plt.show()
Ans	i. d. matplotlib.pyplot as plt ii. c. bar() iii. b. plt.title(“Happiness Index across cities by gender”,fontsize=18) iv. b. plt.legend() v. d. plt.show()

Q3.	Abhradeep is working in a game development industry and he was comparing the given chart on the basis of the rating of the various games available on the play store.
	He is trying to write a code to plot the graph. Help Abhradeep to fill in the blanks of the code and get the desired output.
	
	import_____ #Statement 1
	Games=[“Subway Surfer”,”TempleRun”,”CandyCrush”,”BottleShot”,”RunnerBest”]
	Rating=[4.2,4.8,5.0,3.8,4.1]
	plt._____(Games,Rating) #Statement 2
	plt.xlabel(“Games”)
	plt._____(“Rating”) #Statement 3
	plt._____ #Statement 4
	
	1 Choose the right code from the following for statement 1.
	i. matplotlib as plt
	ii. pyplot as plt
	iii. matplotlib.pyplot as plt
	iv. matplotlib.plt as pyplot

	2. Identify the name of the function that should be used in statement 2 to plot the above graph. i. line() ii. bar() iii. hist() iv. barh() iii. Choose the correct option for the statement 3. i. title("Rating") ii. ytitle("Rating") iii. ylabel("Rating") iv. yaxis("Rating") iv. Choose the right function/method from the following for the statement 4. i. display() ii. print() iii. bar() iv. show() v. In case Abhradeep wants to change the above plot to the any other shape, which statement, should he change. i. Statement 1 ii. Statement 2 iii. Statement 3 iv. Statement 4
Ans	1. Answer. (iii) matplotlib.pyplot as plt 2. Answer:(ii) bar() 3. Answer: (iii) ylabel("Rating") 4. Answer: (iv) show() 5. Answer: Statement 2
Q4.	**Ms.Soumya is working in the mobile app development industry and he was comparing the given chart on the basis of the rating of the various apps available on the play store.**

	
	He is trying to write a code to plot the graph. Help Mr. Vijay to fill in the blanks of the code and get the desired output.
	import _____ as plt #Statement 1
	apps=["Arogya Setu","WPS Office","CamScanner","WhatsApp","Telegram"]
	ps_rating=[3.9,4.5,4.6,4.2,4.3]
	plt._____(apps,ps_rating,color='m',label=_____) #Statement 2 Statement 3
	plt.xlabel("Apps")
	plt._____("Rating") #Statement 4
	plt._____ #Statement 5
	plt._____ #Statement 6

Ans	i)matplotlib.pyplot ii) bar, App Rating iii) ylabel iv) legend(), show() v) Statement 2 should be changed. It requires plot() method to plot the data.
Q5	The heights of 10 students of eighth grade are given below: Height_cms=[145,141,142,142,143,144,141,140,143,144] Write suitable Python code to generate a histogram based on the given data, along with an appropriate chart title and both axis labels. Also give suitable python statement to save this chart.

Ans	import matplotlib.pyplot as plt #Statement 1
	Height_cms=[145,141,142,142,143,143,141,140,143,144] #Statement 2 plt.hist(Height_cms) #Statement 3
	plt.title("Height Chart") #Statement 4
	plt.xlabel("Height in cms") #Statement 5
	plt.ylabel("Number of people") #Statement 6

	plt.show() #Statement 7 (½ mark each for each correct statement 1,2,4,5,6,7) (1 mark for correct statement 3) plt.savefig("heights.jpg") (1 mark for the correct statement)
Q6.	A) What changes will you recommend to rectify the error in given code?(Note. All required libraries have been imported and are available) a = range (10, 50, 12) b = range (90, 200, 20) matplotlib.pyplot.plot(a, b) B) Correct the errors in the following code import matplotlib.pyplot as plt import numpy as np ypoints = np.array([3, 8, 1, 10]) plt.bar(ypoints, style = 'dotted') plt.show() C) Name any 2 parameters of plot function.
Ans	A) Since both the sequences being plotted must match in their shape, we can achieve this either by adding two elements to sequence a so that it has the same shape as sequence b (i.e., 6 elements) or by removing two elements from sequence b so that it matches the shape of sequence a (i.e., 4 elements). B) import matplotlib.pyplot as plt import numpy as np ypoints = np.array([3, 8, 1, 10]) plt.plot(ypoints, linestyle = 'dotted') plt.show() C) Any 2 from color, linestyle, marker, markerfacecolor, markersize

Name of Chapter: DBMS (MySQL)

SINGLE ROW FUNCTIONS

Topics Covered

- Revision of database concepts and SQL commands covered in class XI
- SQL Functions-Math Function, Text Function, Date and Time functions

Key Points

Built-In Functions

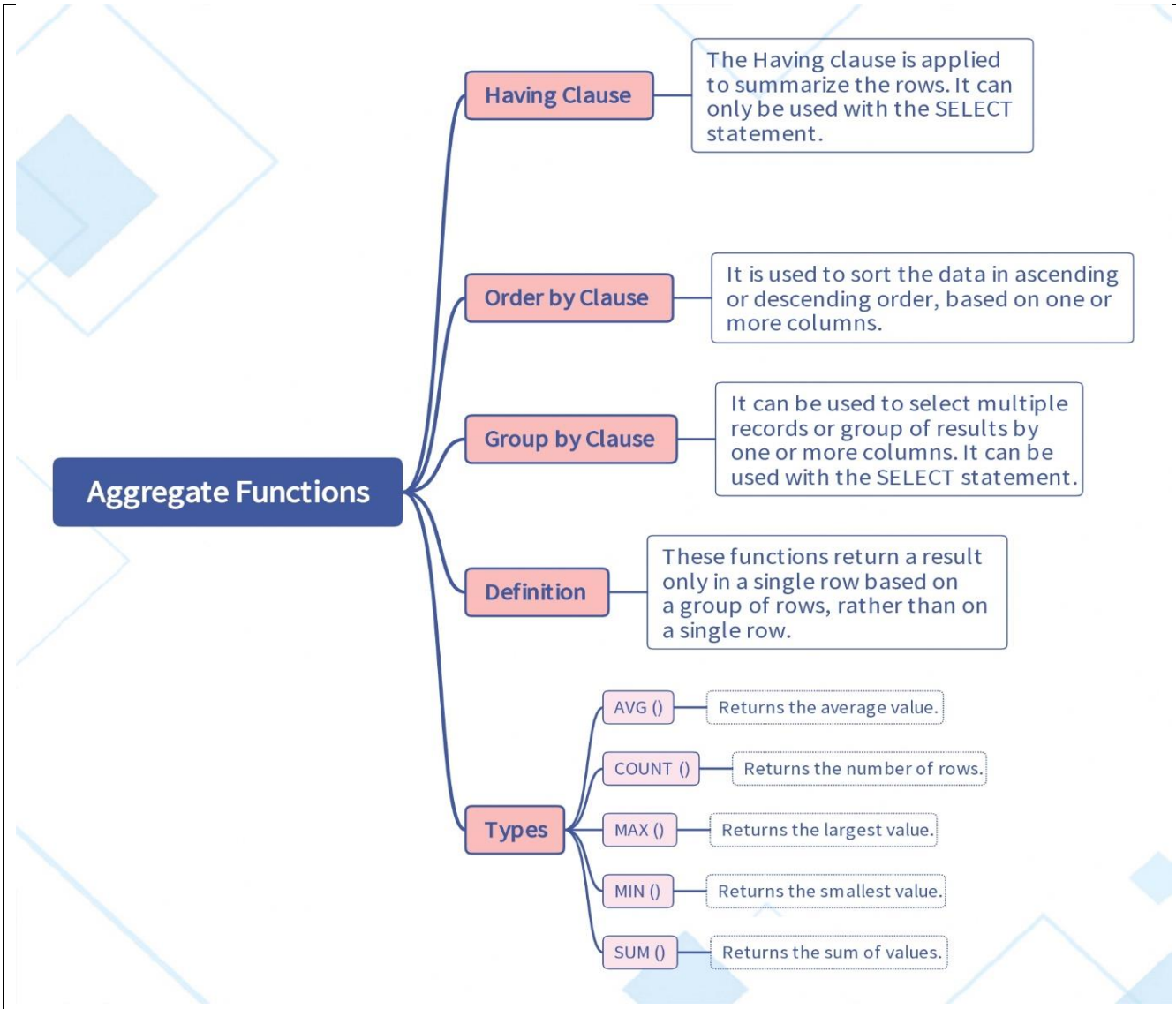
- SQL provides many built-in Functions to perform operations on data. These functions are useful while performing mathematical calculations, string concatenations, sub-strings, etc.
- According to the processing on the value of column SQL functions can be classified as below.

TWO types of SQL Functions

- **Group Functions**
  - "Functions act on set of values are known asgroup functions".
  - SUM, AVG, IMIN, MAX (AggregateFunctions)
- **Scalar Functions**
  - "Functions act on only one value at a time are known as Scalar functions".
  - Length, ASCII.

- According to the SQL Data Type functions can be classified as below.
  - **Numeric Functions:**
    - For processing Number Data type.
  - **String Functions:**
    - For processing on String Data type.
  - **Conversion Functions:**
    - For converting one Data type from one data typeto another.
  - **Date Functions:**
    - For processing Date type Data
  - **Mathematical Functions:**
    - There are various built-in functions available in MySQL for mathematical calculations. These mathematical functions accept numeric value, perform pre-defined specific operations on it and also return numeric valuein result.

Built-in Functions
Text Functions
These functions are used to extract, change, format or alter character string.
LEFT()
Returns a specified number of characters from the left.
TRIM()
Removes all leading and trailing spaces of the string.
LTRIM()
Removes all leading spaces of the string.
RTRIM()
Removes all trailing spaces of the string.
LENGTH()
Returns the length of the string argument.
UCASE()
Converts the characters of the string to upper case.
MID()
Returns a substring of the specified length, starting from specified position.
LCASE()
Converts the characters of the string to lower case.
INSTR()
Returns the position of the first occurrence of a specified substring from a given string.
Math Functions
Maths functions or mathematical functions accept numeric values perform some operations on it and also return numeric values as a result.
POWER()
Returns the power of the given value.
MOD()
This function is used to return the remainder.
ROUND()
Returns the given number rounded up or downward to the nearest whole number.
Date Functions
These functions are used to perform operations on the date type data which are stored in the database.
DATE()
Returns the name of the date of the week for a specified date.
MONTH()
Returns the month number for a date which is a value between 1 and 12 for January to December.
YEAR()
Returns the year of the date type value.
NOW()
Returns the current date and time as per the configured time zone.
DAY()
Returns the day number of a month from a given date.
MONTHNAME()
Returns the full name of the month.



Mind Map of Aggregate Functions

Some mathematical functions used in MySQL are asfollows:

(i) **POWER( ):** This function is used to get thepower of the given values.

Syntax:

POWER (m, n)

Parameter:

m: It is a base value in the calculation.

n: It is exponent value in the calculation.

This function returns m raised to the nthpower.

Example:

SELECT POWER (4, 3);

Output:

POWER (4, 3)
64

(ii) **ROUND( ):** This function is used to round up the number to the upwards or downwards

whichever is the nearest whole number.

Syntax:

ROUND (number)

If you want to get number with certainnumber of decimal places, you can also passthat number, and use following syntax.

ROUND (number, decimal place);

Example 1:

SELECT ROUND (53.327);

Output:

ROUND (53.327)
53

Example 2:

SELECT ROUND (53.327, 2);

Output:

ROUND (53.327)
53.32

(iii) **MOD( ):** This function is used to return the remainder of one number or expression by dividing it to anothernumber or expression.

Syntax:

MOD (m, n)

Parameter:

m: It is the number to be divided by n.

n: It is the number that will divide m.

Example:

SELECT MOD (26, 3);

Output:

MOD (26, 3)
2

➢ Text/ String Functions:

MySQL text functions manipulate the character string data effectively. Some text functions used in MySQL are as follows:

(i) **UCASE/UPPER function**: UCASE () or UPPER () function is used to convert thestring argument into upper case characters.

Syntax:

UCASE (str)

Or

UPPER (str)

Example:

SELECT UCASE (‘Hello’);

Output:

UCASE (‘Hello’)

HELLO

Example:

SELECT UPPER (‘World’);

Output:

UPPER (‘World’)
WORLD

(ii) **LCASE/LOWER function**: LCASE () or LOWER () function is used to convert thestring argument into lower case characters.

Syntax:

LCASE (str)

Or

LOWER (str)

Example:

SELECT LCASE (‘HELLO’);

Output:

LCASE (‘HELLO’)
hello

Example:

SELECT LOWER (‘WORLD’);

Output:

LOWER (‘WORLD’)
world

(iii) **MID( ):** This function extracts a substring from a string and returns a string with given length and position.

Syntax:

MID (str, pos, len)

Example:

SELECT MID('Python program' , 3, 5);

Output:

MID('Python program' , 3, 5)
thon

(iv) **SUBSTRING ()/ SUBSTR():** These functionsare same as MID( ) function.

(v) **LENGTH( )**:This function is used to return the length of the specified string. It returns the length in bytes. This function also includes all the blank spaces which are include in string.

Syntax:

LENGTH (str)

Example:

SELECT LENGTH ('Python');

Output:

SELECT LENGTH ('Python')
6

(vi) **LEFT( ):** This function is used to return a specified number of characters from the left ofthe string. The number of characters returnedis determined by the second argument.

Syntax:

LEFT (str, len)

Example:

SELECT LEFT ('Statement', 4);

Output:

SELECT LEFT ('Statement', 4)
Stat

(vii) **RIGHT( ):** This function is the opposite of LEFT( ) function. It is used to return a specified number of characters from the right of the string. The number of characters returned is determined by the second argument.

Syntax:

RIGHT (str, len);

Example:

SELECT RIGHT ('Statement', 4);

Output:

SELECT RIGHT ('Statement', 4)
ment

(viii) **INSTR( ):** This function takes two arguments as str (string) and sub_str (sub string) and returns the position of the first occurrence of a specified sub str from a given Str.

Syntax:

INSTR(str, sub_str);

Example:

SELECTINSTR (‘Hello World’, ‘Wor’);

Output:

SELECTINSTR (‘Hello World’, ‘Wor’)
7

(ix) **LTRIM( ):** This function takes a string argument and returns a new string with all the leading space characters removed. Spaces in the middle or trailing spaces are not removed.

Syntax:

LTRIM (str);

Example:

SELECT LTRIM('Python Program’) AS Result;

Output:

Result
Python Program

(x) **RTRIM( ):** This function takes a string argument and returns a new string with all the trailing space characters removed. Spaces in the middle or leading space are not removed.

Syntax:

RTRIM (str);

Example:

SELECT RTRIM('Python Program ') AS Result;

Output:

Result
Python Program

(xi) **TRIM( ):**This functions enables you to remove both leading and trailing white space from string.

Syntax:

TRIM (str);

Example:

SELECT TRIM('Python Program ') AS Result;

Output:

Result
Python Program

➢ Date Functions:

The date functions are used to perform some operations on date that is stored in the database. Some common date functions are as follows:

(i) **NOW( ):** This function returns the current date and time in the configured time zone as a string, or a number in the 'YYYY-MM-DD HH: MM: SS' or 'YYYYMMDDHHMMSS' format.

Syntax:

NOW();

Example:

SELECT NOW ();

Output:

SELECT NOW ()
2023-10-06 18:12:25

(ii) **DATE( ):** This function extracts the date value from a given date.

Syntax:

DATE (date);

Example:

SELECT DATE ('2023-10-06');

Output:

SELECT DATE ('2023-10-06')
2023-10-06

(iii) **MONTH( ):** This function returns the month for date, in the range I to 12 for January to December. If it returns '0' then month part of the given date contains NULL.

Syntax:

MONTH (date);

Example:

SELECT MONTH ('2023-10-06');

Output:

SELECT MONTH ('2023-10-06')
10

(iv) **MONTHNAME( ):** This function returns the full name of the month for the given date.

Syntax:

MONTHNAME (date);

Example:

SELECT MONTHNAME ('2023-10-06') AS Result;

Output:

Result
October

(v) **YEAR( ):** This function returns the year of the given date. It returns a year value in the range 1000 to 9999. If the date is zero, it returns 0.

Syntax:

YEAR (date);

Example:

SELECT YEAR ('2023-10-06 ') AS Result;

Output:

Result
2023

(vi) **DAY( ):** This function returns the day of the month of a given date. If the date argument is zero, it returns 0. In case, the date in NULL, this function returns NULL.

Syntax:

DAY (date);

Example:

SELECT DAY ('2023-10-06');

Output:

SELECT DAY ('2023-10-06')
06

(vii) DAYNAME( ): It returns the name of the day from the given date.

Syntax:

DAYNAME (date);

Example:

SELECT DAYNAME ('2023-10-06');

Output:

SELECT DAYNAME ('2023-10-06')
Friday

➢ Types of Functions

Depending on their application in one or multiplerows, SQL functions may be categorised as:

- **Single Row functions:**
  These are also known as Scalar functions. Single row functions are the one the work on single row and return one output per row. For example, length and case conversion functions are single row functions.

- **Multiple Row functions:**
  Multiple row functions are also called Aggregate functions. Multiple row functions work upon group of rows and return one result for the complete set of rows. They are also known as Group Functions.

➢ Aggregate Functions

- An aggregate function performs a calculation on one or more values and returns a single value. We often use aggregate functions with the GROUP By and HAVING clauses of the SELECT statement.
- Except for count C), aggregate functions totally ignore NULL values and considers all values in the present in a column.

Some aggregate functions are asfollows:

(i) **MAX():** This function returns the maximum value in selected columns. MAX() function ignores NULL values and considers all values in the calculation.

Syntax:

SELECT MAX (column_name) FROM Table_Name ;

(ii) **MIN ():** This function returns the minimum value in selected columns. MINO function ignores NULL values.

Syntax:

SELECT MIN(column_name) FROM Table_Name;

(iii) **AVG ( ):** This function calculates the average ofspecified column(s). It ignores NULL values.

Syntax:

SELECT AVG (column_name) FROM Table_Name ;

(iv) **SUM():** This function calculates the sum of all values in the specified columns. It accepts only the expression that evaluates to numeric values.

Syntax:

SELECT SUM (column_name) FROM Table_Name;

(v) **COUNT():** This function returns the number of rows found in a set. The COUNT (*) function returns a number of rows in a specified table or view that includes the number of duplicates and NULL values.

Syntax:

SELECT COUNT (*) FROM Table_Name ;

➢ GROUP BY Clause:

GROUP BY clause is used to group rows returned by SELECT statement into a specified rows or groups.

Syntax:

SELECT column 1, column 2, ..., Aggregate_function
(exp)
FROM Table Name
WHERE condition
GROUP BY Column Name;

➢ ORDER BY clause:

ORDER BY clause is used to sort a result set returned by a SELECT statement.
To sort a result set in ascending order, use ASC Keyword and in descending order, use DESC Keyword. The ORDER BY clause sorts the result set in ascending order by default.

Syntax:

SELECT column 1, column 2,...
FROM Table Name
ORDER BY Column Name <ASC/DESC>

➢ HAVING Clause:

HAVING clause is often used with the GROUP BY rows based on a specified condition.

Syntax:

SELECT column 1, column 2, ..., Aggregate_function

(Exp)
FROM Table Name
GROUP BY Column Name
HAVING condition;

Some Notable Points:

- **Data type:** It is an attribute that specifies thetype of data that the object can hold.
- **Expression:** It is a combination of one or more values, operators and SQL functions thatevaluate to a value.
- DISTINCT Keyword help us in removing the duplicates values from the result.
- COUNT (*) includes all the null and duplicate values.
- ALL keyword includes even duplicates. If nothing is specified, ALL is assumed as the default.
- SUM and AVG functions only work on numeric data.

15 Objective Question (1 Mark)

Q1.	Write the output of the following SQL command. SELECT ROUND (49.88); a. 49.88 b. 9.88 c. 49.0 d. 50
Ans	Option (D)
Q2.	Predict the output of the following query: SELECT MOD (9,0); a. 0 b. NULL c. NaN d. 9
Ans	Option (B)
Q3.	The avg () function in MySQL is an example of …….. a. Math Function b. Text Function c. Date Function d. Aggregate Function
Ans	Option (D)
Q4.	Write the output of the following SQL command. SELECT SUBSTR ('COMPUTER', 3, 4); a. MPUT b. PUTE c. PU d. MP
Ans	Option (A)
Q5.	The string function that returns the index of thefirst occurrence of substring is _____. a. INSERT () b. INSTR () c. INSTRING () d. INFSTR ()
Ans	Option (B)
Q6.	What values does the count (*) function ignore? a. Repetitive values b. Characters c. Null values d. Integers
Ans	Option (C)
Q7.	What is the meaning of "GROUP BY" clause inMySQL? a. Group data by column and row values b. Group data by row values c. Group data by column values

	d. None of these
Ans	Option (C)
Q8.	Which of the following SQL functions does not belong to the Math functions category? a. POWER () b. LENGTH () c. ROUND () d. MOD ()
Ans	Option (B)
Q9.	Which clause is similar to "HAVING" clause inMySQL? a. SELECT b. FROM c. WHERE d. MOD ()
Ans	Option (B)
Q10.	In SQL, the equivalent of UCASE () is: a. UPPERCASE () b. CAPITALCASE () c. UPPER () d. TITLE ()
Ans	Option (A)
Q11	Only _____ functions are used with the GROUP BY clause. a. Text b. Math c. Date/ Time d. Aggregate
Ans	Option (D)
Q12	Write the output of the following SQL command: SELECT left ("Jammu Region", 5); a. Region b. Jammu c. Jammu Region d. None of the Above
Ans	Option (B)

Q13	Raj, a Database Administrator, needs to display the average pay of workers from those departments which have more than five employees. He is experiencing a problem while running the following query:
	
	SELECT DEPT, AVG(SAL) FROM EMP WHERE COUNT (*) > 5
	GROUP BY DEPT;
	
	Which of the following is a correct query to perform the given task?
	a. SELECT DEPT, AVG(SAL) FROM EMP WHERE COUNT (*) > 5 GROUP BY DEPT;
	b. SELECT DEPT, AVG(SAL) FROM EMP HAVING COUNT (*) > 5 GROUP BY DEPT;
	c. SELECT DEPT, AVG(SAL) FROM EMP GROUP BY DEPT WHERE COUNT (*) > 5;
	d. SELECT DEPT, AVG(SAL) FROM EMP GROUP BY DEPT HAVING COUNT (*) > 5;

Ans	Option (D)
Q14	Which of the following aggregation operationadds up all the values of the attribute? a. ADD () b. AVG () c. MAX () d. SUM ()
Ans	Option (D)
Q15	Predict the output of the following query: SELECT LCASE (MONTHNAME ('2023-03-05')); a. May b. March c. march d. may

Ans	Option (C)
05 Assertion and reason Based question (1 Mark)	
Q1.	**Assertion (A):** RTRIM () function takes a stringargument and returns a new string with all thetrailing space characters removed. **Reason (R):** TRIM () function enables you toremove leading and trailing white space from the string.
Ans	Option (B) Both assertion (A) and reason (R) are true, but reason (R) is not the correct explanation of assertion (A).
Q2.	**Assertion (A):** The ORDER BY clause sorts the result set in descending order by default. **Reason (R):**To sort a result set in ascending order we can use ASC keyword with ORDER BY clause.
Ans	Option (D) Assertion (A) is false, but Reason (R) is true.
Q3.	**Assertion (A):** SELECT MOD (15, 3); producethe output as O (Zero). **Reason (R):**POWER ( ) function is used to get thepower of the given values.
Ans	Option (B) Both assertion (A) and reason (R) are true, but reason (R) is not the correct explanation of assertion (A).
Q4.	**Assertion (A):**HAVING clause is often used withthe GROUP BY statement. **Reason (R):** HAVING clause is used to checkspecified condition.
Ans	Option (A) Both assertion (A) and reason (R) are true andreason (R) is the correct explanation of assertion (A).
Q5.	**Assertion (A):** UPPER () function is used to convertstring argument into upper case characters. **Reason (R):**UCASE () function is used to convert the string argument into lower case characters.
Ans	Option (C) Assertion (A) is true, but Reason (R) is false.
05 Short Knowledge/Understanding/Application Based Questions (2 Marks)	
Q1.	Considering the same string "Preoccupied" WriteSQL commands to display: (i) the position of the substring 'cup' in thestring"Preoccupied" (ii) the first 4 letters of the string.
Ans	(i) SELECT INSTR ("Preoccupied", "cup"); (ii) SELECT LEFT ("Preoccupied", 4);
Q2.	Consider the given SQL string: "12#All the Best!" Write suitable SQL queries for the following: i. Returns the position of the first occurrence of the substring "the" in the given string. I ii. To extract last five characters from the string.
Ans	i. SELECT INSTR ("12#All the Best!","the"); ii. SELECT RIGHT ("12#All the Best!",5);
Q3.	Consider the decimal number x with value8459.2654. Write commands in SQL to: (i) round it off to a whole number. (ii) round it to 2 places before the decimal.
Ans	(i) SELECT ROUND (8459.2654) ; (ii) SELECT ROUND (8459.2654, -2) ;
Q4.	What are aggregate functions in SQL? Name any two.
Ans	**Aggregate functions:** These are also called multiple row functions. These functions work on a set of records as a whole and return a single value for each column of the records on which the function is applied.

Max (), Min (), Avg(), Sum(), Count() and Count(*) are few examples of multiple row functions.

Q5. State any two differences between single rowfunctions and multiple row functions.

Ans Differences between single row functions and multiple row functions are as follows:

Sl No	Single Row Functions	Multiple Row Functions
(i)	Singlerowfunctions work onone row only.	Multiplerowfunctions work on group of rows.
(ii)	These functionsreturn one outputper row.	These functionsreturn only oneoutput foraspecified group ofrows.

05 Short Knowledge/Understanding/Application Based Questions (3 Marks)

Q1. Predict the output of the following queries:

(i) select instr('exams@cbse.nic.in','.');
(ii) select substr('exams@cbse.nic.in',7,4);
(iii) select left('exams@cbse.nic.in',5);

Ans
(i) 11
(ii) cbse
(iii) exams

Q2. Based on the SQL table CAR_SALES, write suitable queries for the following:

NUMBER	SEGMENT	FUEL	QT1	QT2
1	Compact HatchBack	Petrol	56000	70000
2	Compact HatchBack	Diesel	34000	40000
3	MUV	Petrol	33000	35000
4	MUV	Diesel	14000	15000
5	SUV	Petrol	27000	54000
6	SUV	Diesel	18000	30000
7	Sedan	Petrol	8000	10000
8	Sedan	Diesel	1000	5000

i. Display fuel wise average sales in the first quarter.
ii. Display segment wise highest sales in the second quarter.
iii. Display the records in the descending order of sales in the secondquarter.

OR

Predict the output of the following queries based on the table CAR_SALES given above:

i. SELECT LEFT(SEGMENT,2) FROM CAR_SALES WHERE FUEL= "PETROL";
ii. SELECT (QT2-QT1)/2 "AVG SALE" FROM CAR_SALES WHERE SEGMENT= "SUV";
iii. SELECT SUM(QT1) "TOT SALE" FROM CAR_SALES WHERE FUEL= "DIESEL";

Ans
i. SELECT FUEL, AVG(QT1) FROM CAR_SALES GROUP BY FUEL;
ii. SELECT SEGMENT, MAX(QT2) FROM CAR_SALES GROUP BY SEGMENT;
iii. SELECT * FROM CAR_SALES ORDER BY QT2 DESC;

OR

i.

```
+------------------------+
| LEFT(SEGMENT,2) |
+------------------------+
| Co                     |
| MU                     |
| SU                     |
| Se                     |
+------------------------+
```

ii.

```
+--------------+
| AVG SALE |
+--------------+
| 13500.0000 |
| 6000.0000 |
+--------------+
```

iii.

```
+--------------+
| TOT SALE |
+--------------+
| 67000 |
+--------------+
```

Q3. Mr. Soumyadeep is working on a MySQL table named'Hotel' having following structure:

Field	Type	Null	Key	Default	Extra
user_id	varchar(20)	YES		NULL	
name	varchar(20)	YES		NULL	
city	varchar(20)	YES		NULL	
mobile_no	varchar(20)	YES		NULL	

He needs to perform the following tasks on the table:

(i) To fetch last 2 characters from the user id column.
(ii) To display the values of name column in lower case.
(iii) To display 3 characters from 3rd place from the column city.

Suggest suitable SQL function for the same. Alsowrite the query to achieve the desired task.

Ans

(i) right ()
select right (user_id,2) from hotel;
(ii) lower ()
select lower (name) from hotel;
(iii)mid ()/substr()/substring()
select mid (city,3,3) from hotel;

Q4. While dealing with string data type in MySQL, its observed that sometimes unnecessary space character comes in between which hampers the successful execution of a string manipulation module. Name the suitable MySQL function (s) to remove leading, trailing and both type of space characters from a string.
Also give MySQL queries to depict the same.

Ans

(i) To remove leading space characters: ltrim ()
(ii) To remove trailing space characters: rtrim ()
(iii) To remove both type of space characters: trim ()

MySQL Queries:

```
select ltrim(' Hello ');
select rtrim(' Hello ');
select trim(' Hello ');
```

Output:

Hello

Q5. Consider the table GAMES given below:

Gcode	Games_Name	Type	Number	Prize_Money	Schedule_Date
101	Carrom Board	Indoor	2	5000	23.01.2004
102	Badminton	Outdoor	2	12000	12.122003
103	Table Tennis	Indoor	4	8000	14.02.2004
105	Chess	Indoor	2	9000	01.01.2004
108	Lawn Tennis	Outdoor	4	25000	19.03.1994

Write SQL commands for the following:

(i) To display details of those games which are having Prize Money more than 7000;
(ii) To display sum of Prize Money for each type of GAME.
(iii) To display the total number of games available in the above table GAMES.

Ans

(i) SELECT * FROM GAMES WHERE Prize_Money> 7000;
(ii) SELECT SUM (Prize_Money), Type FROM GAMES GROUP BY Type;
(iii) SELECT COUNT (Games_Name) FROM GAMES;

05 Short Knowledge/Understanding/Application Based Questions (4 Marks)

Q1. Satyam. a database analyst has created thefollowing table:

RegNo	SName	Stream	Optional	Marks
S1001	Akshat	Science	CS	99
S1002	Harshit	Commerce	IP	95
S1003	Devika	Humanities	IP	100
S1004	Manreen	Commerce	IP	98
S1005	Gaurave	Humanities	IP	82
S1006	Saurave	Science	CS	NULL
S1007	Bhaskar	Science	CS	95
S1007	Bhaskar	Science	CS	96

He has written following queries:

(i) select sum (Marks) from student whereOptional= 'IP' and STREAM= 'Commerce';
(ii) select max (Marks) + min (Marks) fromstudent where Optional = ‘CS';
(iii) select avg (Marks) from student where Optional = 'IP';
(iv) select length (SName) from student where Marks is NULL;

Help him in predicting the output of the abovegiven queries

Ans

i. 193 ii. 194
iii. 93.75 iv. 6

Q2. What is math function? Explain any two math functions with an example.

Ans

There are various built-in functions include inMySQL for mathematical calculations. Thesemathematical functions accept numeric values,perform some operation on it and also returnnumeric value as result.

Some mathematical functions used in MySQL areas follows.

- POWER ( ) : This function is used to get thepower of the given values.

  Syntax:

  POWER (m, n)

This function returns m raised to the power.
e.g.,

SELECT POWER (4, 3);

Output:

POWER (4, 3)
64

- ROUND ( ): This function is used to round up the number to the upwards or downwardswhichever the nearest whole number.
  Syntax:

  ROUND (number)

  If you want to get number with certain numberof decimal places, you can also pass that numberand use the following syntax:

  ROUND (number, decimal place);

  e.g,,

  SELECT ROUND(56.567);

  Output:

ROUND (56.567)
57.0

Q3. Preeti manages database in a blockchain start-up. For business purposes, she created a table named BLOCKCHAIN. Assist her by writing the following queries:

TABLE : BLOCKCHAIN

id	user	value	hash	transaction_date
1	**Steve**	**900**	**ERTYU**	**2020-09-19**
2	**Meesha**	**145**	**@345r**	**2021-03-23**
3	**Nimisha**	**567**	**#wert5**	**2020-05-06**
4	**Pihu**	**678**	**%rtyu**	**2022-07-13**
5	**Kopal**	**768**	**rrt4%**	**2021-05-15**
7	**Palakshi**	**534**	**wer@3**	**2022-11-29**

i. Write a query to display the year of oldest transaction.
ii. Write a query to display the month of most recent transaction.
iii. Write a query to display all the transactions done in the month of May.
iv. Write a query to count total number of transactions in the year 2022.

Ans

i. SELECT YEAR(MIN(TRANSACTION_DATE)) FROM BLOCKCHAIN;
ii. SELECT MONTH(MAX(TRANSACTION_DATE)) FROM BLOCKCHAIN;
iii. SELECT * FROM BLOCKCHAIN WHERE MONTHNAME (TRANSACTION_DATE)='MAY';
iv. SELECT COUNT(ID) FROM BLOCKCHAIN WHERE YEAR(TRANSACTION_DATE)=2022;

Q4. Explain the following:
(i) DAY ()
(ii) DAYNAME ()

Ans (i) **DAY ():** This function returns the day of the month of a given date. If the date

argument is zero, it returns 0. In case, the date is NULLthis function returns NULL.

Syntax:

DAY (date);

e.g.

SELECT DAY (‘2023-09-28’);

Output

DAY (‘2023-09-28’)
28

(ii) **DAYNAME ( )**: This function returns the name of week day for a specified date. It returnsstring value, means Sunday to Saturday

Syntax:

DANINAME (date);

e.g.

SELECT DAYNAME (‘2023-09-28’);

Output

DAYNAME (‘2023-09-28’)
Thursday

Q5. Carefully observe the following table named ‘stock’:

Pid	PName	Category	Qty	Price
1	Keyboard	IO	15	450
2	Mouse	IO	10	350
3	Wifi-router	NW	5	2600
4	Switch	NW	3	3000
5	Monitor	O	10	4500
6	Printer	O	4	17000

Write SQL queries for the following:

(A) To display the records in decreasing order of price.
(B) To display category and category wise total quantities of products.
(C) To display the category and its average price.
(D) To display category and category wise highest price of the products.

Ans

(A) select * from stock order by price desc;
(B) select category, sum(qty) from stock group bycategory;
(C) select category, avg(price) from stock group bycategory;
(D) select category, max(price) from stock group by category;

05 Case Based Questions (5 Marks)

Q1. Consider a table SALESMAN with the following data:

S.No.	SNAME	SALARY	BONUS	DATE_OF_JOIN
A01	Beena Mehta	30000	45.23	29-10-2019
A02	K L. Sahay	50000	25.34	13-03-2018
B03	Nisha Thakkar	30000	35.00	18-03-2017
B04	Leela Yadav	80000	NULL	31-12-2016
C05	Gautam Gola	20000	NULL	23-01-1989

C06	Trapti Garg	70000	12.37	15-06-1987
D07	Neena Sharma	50000	27.89	18-03-1999

Write SQL queries using SQL functions to perform the following operations:

(i) Display salesman name and bonus after rounding off to zero decimal places.
(ii) Display the position of occurrence of the string "ta" in salesman names.
(iii) Display the four characters from salesman name starting from second character.
(iv) Display the month name for the date of join of salesman.
(v) Display the name of the weekday for the date of join of salesman.

Ans

(i) select SNAME, ROUND (BONUS, 0) from SALESMAN;
(ii) select instr (SNANIE, “ta") from SALESMAN;
(iii) select substring (SNAME, 2, 4) from SALESMAN;
(iv) select monthname (DATE_OF_JOIN) from SALESMAN;
(v) select dayname (DATE_OF_JOIN) from SALESMAN;

Q2. Write suitable SQL queries for the following:

i. To calculate the exponent for 3 raised to the power of 4.
ii. To display current date and time.
iii. To round off the value -34.4567 to 2 decimal place.
iv. To remove all the probable leading and trailing spaces from the column userid of the table named user.
v. To display the length of the string ‘FIFA World Cup’.

OR

Kabir has created following table named exam:

RegNo	Name	Subject	Marks
1	Sanya	Computer Science	98
2	Sanchay	IP	100
3	Vinesh	CS	90
4	Sneha	IP	99
5	Akshita	IP	100

Help him in writing SQL queries to the perform the following task:

i. Insert a new record in the table having following values: [6,'Khushi','CS',85]
ii. To change the value “IP” to “Informatics Practices” in subject column.
iii. To remove the records of those students whose marks are less than 30.
iv. To add a new column Grade of suitable datatype.
v. To display records of “Informatics Practices” subject.

Ans

i. SELECT POWER (3,4);
ii. SELECT NOW ();
iii. SELECT ROUND (-34.4567,2);
iv. SELECT TRIM(USERID) FROM USER;
v. SELECT LENGTH ("FIFA World Cup");

OR

i. INSERT INTO EXAM VALUES(6,'Khushi','CS',85);
ii. UPDATE EXAM SET subject= "Informatics Practices" where subject = "IP";

iii. DELETE FROM EXAM WHERE marks<30;
iv. ALTER TABLE EXAM ADD COLUMN grade varchar (2);
v. Select * from exam where subject="Informatics Practices";

Q3. Write the SQL functions which will perform thefollowing operations:

(i) Display the string "chest" from the string"Manchester United".
(ii) Trim the blanks from both ends from the string"hello " .
(iii) To display the name of the month e.g.- Januaryor February from the current date.
(iv) To convert "world" into uppercase letters.
(v) SELECT INSTR ("HELLO WORLD", "HE")

Ans

(i) SELECT SUBSTRING ("Manchester United", 4, 5);
(ii) SELECT TRIM (" hello ");
(iii) SELECT MONTHNAME(NOW ());
(iv) SELECT UCASE ("world");
(v) 1

Q4. Consider the following table 'Furniture'. Write SQL commands for the statements (i)to (iii) and write outputfor SQL queries (iv) and (v).

Table: Furniture

FCODE	NAME	PRICE	MANUFDATE	WCODE
10023	Coffee table	4000	19-DEC-2016	W03
10001	Dining table	20500	12-JAN-2017	W01
10012	Sofa	35000	06-JUN-2016	W02
10024	Chair	2500	07-APR-2017	W03
10090	Cabinet	18000	31-MAR-2015	W02

(i) To display FCODE and NAME of each Furniture Item in descending order of FCODE.
(ii) To display the average PRICE of all the Furniture Items, which are made of Wood with WCODE as W02.
(iii) To display WCODE wise, WCODE and the highest price of Furniture Items.
(iv) SELECT SUM (PRICE) FROM Furniture WHERE WCODE = 'W03';
(v) SELECT COUNT (DISTINCT PRICE) FROM Furniture;

Ans

(i) SELECT FCODE, NAME FROM FurnitureORDER BY FCODE DESC;
(ii) SELECT AVG (PRICE) FROM FurnitureWHERE WCODE = 'W02';
(iii) SELECT WCODE, MAX(PRICE) FROM FurnitureGROUP BY WCODE;
(iv)

SUM (PRICE)
6,500

(v)

COUNT (DISTINCT PRICE)
5

Q5. Consider the following table Activity. Write SQL Commands for the statements (i) to (ii) and output for SQLqueries (iii) to (v).

Table: Activity

PID	PARTICIPANT	GRADE	EVENT	POINTS	EVENTDATE	HOUSE
101	Amit Dubey	A	Hopping	200	2018-12-19	Gandhi

			bag			
102	Shivraj Singh	B	Skipping	300	2019-01-12	Bose
103	Raj Arora	B	Bean bag	200	2018-12-19	Gandhi
104	Kapil Raj	A	Obstacle	250	2018-12-19	Bhagat
105	Deepshlkha Sen	A	Egg & Spoon	350	2018-03-31	Bose
106	Saloni Raj		Hopping bag	200	2018-12-20	Bose

(i) To display names of Participants and pointsin descending order of points.
(ii) To display House wise total points scored along with House name. (i.e. display the HOUSE and total pointsscored by each HOUSE.)
(iii) SELECT AVERAGE (POINTS) FROM Activity WHERE HOUSE = 'Gandhi' orHOUSE = 'Bose':
(iv) SELECT COUNT (DISTINCT POINTS)FROM ACTIVITY;
(v) SELECT SUM(POINTS) FROM ACTIVITY;

Ans

(i) SELECT PARTICIPANTPOINTS FROM Activity ORDER BY POINTS DESC;
(ii) SELECT HOUSE, SUM(POINTS) FROM Activity GROUP BY HOUSE;
(iii) 250
(iv) 4
(v) 1500

Name of Chapter: DBMS (MySQL)

My SQL Aggregate Functions, Group By and Joins

Topics Covered:

- Aggregate Functions →**max( ), min( ), sum( ), avg( ), count( )/count(*)**
- Querying and manipulating data using **Group by, Having, Order by**

Key Points

AGGREGATE FUNCTIONS:

An aggregate function in SQL performs a calculation on multiple values and returns a single value. SQL provides many aggregate functions that include avg, count, sum, min, max, etc. An aggregate function ignores NULL values when it performs the calculation, except for the count function.

MAX( )

Max() function returns the maximum value from a given column or expressions

MIN( )

Max() function returns the minimum value from a given column or expressions

SUM()

Sum() function returns the sum of values in the given parameter(input) column or expression .

AVG( )

Avg() function returns the average value of the given parameter(input) value.

COUNT( ) and COUNT(*)

Count function returns the total no. of values/records under the specified column or expression.

Illustrating AGGREGATE FUNCTIONS using PRODUCT Table

Pno	Pname	DOP	Company	Price	Qty
101	Router	2003-12-09	TitBit	3599.990	100
102	Switch	2005-10-06	Cosmos	3890.890	50
103	RAM	2004-01-01	Universal	2899.990	60
104	WebCam	2004-08-24	Starlite	1950.490	20
105	Memory Card	2004-07-03	PCWorks	295.000	200
106	Head Phone	2003-02-16	NULL	NULL	NULL
107	Bluetooth Headset	2005-05-19	Cosmos	1190.000	10
108	Speaker	2004-07-10	StarMark	1659.890	25

```
mysql> select MAX(Qty),MIN(Qty),SUM(Qty),AVG(Qty),COUNT(Qty),COUNT(*) from product;
+----------+----------+----------+----------+------------+----------+
| MAX(Qty) | MIN(Qty) | SUM(Qty) | AVG(Qty) | COUNT(Qty) | COUNT(*) |
+----------+----------+----------+----------+------------+----------+
|      200 |       10 |      465 |  66.4286 |          7 |        8 |
+----------+----------+----------+----------+------------+----------+
1 row in set (0.00 sec)
```

NOTE:

The * is the only argument that includes NULLs when it is used only with COUNT( ), functions other than COUNT disregard NULLs in any case.

ORDER BY

Usually, the output displayed after the execution of SELECT query is predecided. To sort the result output of a query in a specific order, we use ORDER BY clause. Sorting can be done in either **ascending (asc)** or**descending (desc) order.**
If nothing is mentioned, by default it is done in the **, (ascending) order**

```
mysql> select * from student;
+-------+----------+------------+----------+-------+--------+
| admno | sname    | DOB        | Stream   | Marks | deptno |
+-------+----------+------------+----------+-------+--------+
|   101 | Ranjita  | 2003-12-09 | Science  |    96 |     10 |
|   102 | Amita    | 2005-10-06 | Commerce |    90 |     10 |
|   103 | Ranjan   | 2004-01-01 | Commerce |    85 |     10 |
|   104 | Ashok    | 2004-08-24 | Science  |   100 |     10 |
|   105 | Dharna   | 2004-07-03 | Science  |    96 |     10 |
|   106 | Aman     | 2003-02-16 | Arts     |  NULL |   NULL |
|   107 | Priyanka | 2005-05-19 | Arts     |    90 |     10 |
|   108 | Vihaan   | 2004-07-10 | Science  |    65 |     10 |
+-------+----------+------------+----------+-------+--------+
```

```
mysql> select * from student
    -> order by marks;
+-------+----------+------------+----------+-------+--------+
| admno | sname    | DOB        | Stream   | Marks | deptno |
+-------+----------+------------+----------+-------+--------+
|   106 | Aman     | 2003-02-16 | Arts     |  NULL |   NULL |
|   108 | Vihaan   | 2004-07-10 | Science  |    65 |     10 |
|   103 | Ranjan   | 2004-01-01 | Commerce |    85 |     10 |
|   102 | Amita    | 2005-10-06 | Commerce |    90 |     10 |
|   107 | Priyanka | 2005-05-19 | Arts     |    90 |     10 |
|   101 | Ranjita  | 2003-12-09 | Science  |    96 |     10 |
|   105 | Dharna   | 2004-07-03 | Science  |    96 |     10 |
|   104 | Ashok    | 2004-08-24 | Science  |   100 |     10 |
+-------+----------+------------+----------+-------+--------+
8 rows in set (0.00 sec)
```

```
mysql> select * from student
    -> order by marks desc;
+-------+----------+------------+----------+-------+--------+
| admno | sname    | DOB        | Stream   | Marks | deptno |
+-------+----------+------------+----------+-------+--------+
|   104 | Ashok    | 2004-08-24 | Science  |   100 |     10 |
|   101 | Ranjita  | 2003-12-09 | Science  |    96 |     10 |
|   105 | Dharna   | 2004-07-03 | Science  |    96 |     10 |
|   102 | Amita    | 2005-10-06 | Commerce |    90 |     10 |
|   107 | Priyanka | 2005-05-19 | Arts     |    90 |     10 |
|   103 | Ranjan   | 2004-01-01 | Commerce |    85 |     10 |
|   108 | Vihaan   | 2004-07-10 | Science  |    65 |     10 |
|   106 | Aman     | 2003-02-16 | Arts     |  NULL |   NULL |
+-------+----------+------------+----------+-------+--------+
8 rows in set (0.00 sec)
```

```
mysql> SELECT dob
    -> FROM student
    -> order by DOB;
+------------+
| dob        |
+------------+
| 2003-02-16 |
| 2003-12-09 |
| 2004-01-01 |
| 2004-07-03 |
| 2004-07-10 |
| 2004-08-24 |
| 2005-05-19 |
| 2005-10-06 |
+------------+
8 rows in set (0.00 sec)
```

GROUP BY CLAUSE

- The Group By clause **combines all those records that have identical values in a particular field or a group of fields**. This grouping results into **one summary record per group** if group functions are used with it.

- Grouping can be done **by a column name, or with aggregate functions** in which case the aggregate produces a value for each group.

- All rows with a **NULL** in the column are treated as if **NULL** was another **value**. If a **grouping** column **contains null values**, all **null values** are considered equal, and they are put into a single **group**.

SYNTAX

SELECT <column name> [,<*column name*>,....]
FROM <table name>
GROUP BY <coumn name>;

Illustration of Group By clause using a table STUDENT

```
mysql> SELECT * FROM STUDENT;
+-------+----------+------------+----------+-------+--------+
| admno | sname    | DOB        | Stream   | Marks | deptno |
+-------+----------+------------+----------+-------+--------+
|   101 | Ranjita  | 2003-12-09 | Science  |    96 |     10 |
|   102 | Amita    | 2005-10-06 | Commerce |    90 |     20 |
|   103 | Ranjan   | 2004-01-01 | Commerce |    85 |     20 |
|   104 | Ashok    | 2004-08-24 | Science  |   100 |     10 |
|   105 | Dharna   | 2004-07-03 | Science  |    96 |     10 |
|   106 | Aman     | 2003-02-16 | NULL     |  NULL |     10 |
|   107 | Priyanka | 2005-05-19 | Arts     |    90 |     30 |
|   108 | Vihaan   | 2004-07-10 | Science  |    65 |     10 |
+-------+----------+------------+----------+-------+--------+
8 rows in set (0.00 sec)
```

```
mysql> select stream,count(*)
    -> From student
    -> GROUP BY stream;
+----------+----------+
| stream   | count(*) |
+----------+----------+
| Science  |        4 |
| Commerce |        2 |
| NULL     |        1 |
| Arts     |        1 |
+----------+----------+
4 rows in set (0.00 sec)
```

HAVING condition used along with GROUP BY clause

The **Having clause** places conditions on groups and it can also include aggregate functions.

Illustration of GROUP BY-HAVING clause using a table STUDENT

```
mysql> select stream,count(*)
    -> From student
    -> GROUP BY stream
    -> HAVING count(*)<=2;
+----------+----------+
| stream   | count(*) |
+----------+----------+
| Commerce |        2 |
| NULL     |        1 |
| Arts     |        1 |
+----------+----------+
3 rows in set (0.00 sec)
```

JOINS in MySQL

A **Join** is a query that combines rows from two or more tables. In a Join-query, more than one table are written in the FROM clause. So, we can say that Joining is the function of combining data from multiple relations. **Join Condition e**xpression with WHERE clause is the **Join Predicate.**

SYNTAX of JOIN QUERY:

SELECT <field list>

FROM <table1>,<table2> [,<table3>…]

[WHERE <Join condition for the tables>];

Example:

Select * From Subject,Teacher Where Subject.sub_no=Teacher.sub_no;

EQUI-JOIN

The join in which columns are compared for **equality (=)**, is called Equi Join. This join combines tables based on matching values in specified columns in the join condition.

Example:

Select * From Subject,Teacher Where Subject.sub_no=Teacher.sub_no;

15 Objective Question (1 Mark)

Q1.	Identify the aggregate function out of the following: i) round() ii) distinct() iii) sum() iv) length()
Ans	iii) sum()
Q2.	Identify the odd one, out of the following functions: i) max() ii) trim() iii) sum() iv) count()
Ans	ii) trim()
Q3.	Which clause is appropriate to display the marks secured by all the students of a class starting from the highest marks to the lowest marks. i) where ii) max iii) group by iv) order by
Ans	iv) order by
Q4.	Which clause is used to specify a condition with a GROUP BY clause? i) Having ii) Where iii) From iv) Table
Ans	i) Having
Q5.	By default, Order By clause lists the records in _____ order. i) Descending ii) Ascending iii) Alternate iv) None of the above
Ans	ii) Ascending
Q6.	In MySQL, _____ function computes the cardinality of the table.

	i) Sum() ii) All() iii) Count(*) iv) Count()
Ans	iii) Count(*)
Q7.	In MySQL, identify the correct query used in ordering the values of a field namely marks in the table STUDENT in ascending order: i) Select * From Student Order by marks asc; ii) Select * From Student Order by marks desc; iii) Select * From Student Order by marks; iv) Both (i) and (iii)
Ans	iv) Both (i) and (iii)
Q8.	A/An_____ is a query that retrieves rows from more than one table: i) Update ii) Concatenate iii) Combine iv) Join
Ans	iv) Join
Q9.	Which clause is used in sorting field values in Select queries? (i) Group by (ii) Order By (iii) Order As (iv) Where
Ans	(ii) Order By
Q10.	If column "salary" contains the data set (45000, 5000, 55000, 45000, 55000), what will be the output after the execution of the given query? SELECT AVG (DISTINCT salary) FROM employee; (i) 38500 (ii) 40000 (iii) 41000 (iv) 35000
Ans	(iv) 35000
Q11	The correct SQL from below to display the city and temperature in increasing order of the temperature. (i) SELECT city FROM weather order by temperature ;

	(ii) SELECT city, temperature FROM weather ; (iii) SELECT city, temperature FROM weather ORDER BY temperature ; (iv) SELECT city, temperature FROM weather ORDER BY city ;
Ans	(iii) SELECT city, temperature FROM weather ORDER BY temperature ;
Q12	Where and Having clauses can be used interchangeably in SELECT queries? (i) True (ii) False (iii) Only in views (iv) With order by
Ans	(ii) False
Q13	Which one of the following functions is used to find the smallest value from the given data in MySQL? (i) MINIMUM( ) (ii) MIN( ) (iii) SMALL( ) (iv) TINY()
Ans	(ii) MIN( )
Q14	The following SQL is which type of join? Select * From Subject,Teacher Where Subject.sub_no=Teacher.sub_no; (i) Equi-join (ii) Natural join (iii) Outer join (iv) Cartesian join
Ans	(i) Equi-Join

Q15	Mr. Dev, a Database Administrator, needs to display the average salary of employees from those departments which have less than five employees. He is experiencing a problem while running the following query:
	SELECT DEPT, AVG(SAL) FROM EMP WHERE COUNT(*) < 5 GROUP BY DEPT;
	Which of the following is a correct query to perform the given task?
	i) SELECT DEPT, AVG(SAL) FROM EMP WHERE COUNT(*) > 5 GROUP BY DEPT;
	ii) SELECT DEPT, AVG(SAL) FROM EMP GROUP BY DEPT HAVING COUNT(*) > 5;
	iii) SELECT DEPT, AVG(SAL) FROM EMP GROUP BY DEPT WHERE COUNT(*) > 5;
	iv) SELECT DEPT, AVG(SAL) FROM EMP GROUP BY DEPT HAVING COUNT(*) < 5;

Ans	iv) SELECT DEPT, AVG(SAL) FROM EMP GROUP BY DEPT HAVING COUNT(*) < 5;

05 Assertion and reason Based question (1 Mark)

Q1.	**Assertion (A):-** Order by clause sorts fields in a table in ascending or descending order. **Reasoning (R): -** The WHERE clause is placed before the ORDER BY clause. i. Both A and R are true and R is the correct explanation for A ii. Both A and R are true and R is not the correct explanation for A

	iii. A is True but R is False iv. A is false but R is True
Ans	(ii) Both A and R are true and R is not the correct explanation for A
Q2.	**Assertion (A):-**All rows with a NULL in the column are treated as if NULL was another value in Grouping records as per a field. **Reasoning (R): -**If a grouping column contains null values, all null values are considered equal, and they are put into a single group. i. Both A and R are true and R is the correct explanation for A ii. Both A and R are true and R is not the correct explanation for A iii. A is True but R is False iv. A is false but R is True
Ans	(i) Both A and R are true and R is the correct explanation for A
Q3.	**Assertion (A):-** Aggregate functions perform a calculation on multiple values. **Reasoning (R): -** Aggregate functions return multiple values as output for a given column. i. Both A and R are true and R is the correct explanation for A ii. Both A and R are true and R is not the correct explanation for A iii. A is True but R is False iv. A is false but R is True
Ans	(iii) A is True but R is False
Q4.	**Assertion (A):-** In a Date of Birth (DOB) column, max( ) function returns the DOB of the eldest student as per age. **Reasoning (R): -** Max() aggregate function returns the highest number or latest date. i. Both A and R are true and R is the correct explanation for A ii. Both A and R are true and R is not the correct explanation for A iii. A is True but R is False iv. A is false but R is True
Ans	(iv) A is false but R is True

Q5.	**Assertion (A):-** A JOIN clause is used to combine rows from two or more tables, based on a related column between them. **Reasoning (R): -** A Table in a database is also known as a relation. i. Both A and R are true and R is the correct explanation for A ii. Both A and R are true and R is not the correct explanation for A iii. A is True but R is False iv. A is false but R is True
Ans	(ii) Both A and R are true and R is not the correct explanation for A

05 Short Knowledge/Understanding/Application Based Questions (2 Marks)

Q1. Differentiate between Order By and Group By in terms of MySQL. Also give one example query for each.

Ans

ORDER BY	**GROUP BY**
An ORDER BY allows us to organize result sets alphabetically or numerically and in ascending or descending order.	A GROUP BY statement sorts data by grouping it based on column(s) we specify in the query and is used with aggregate functions.
The Order By clause is always placed after the Group by clause in a Select query.	The Group By clause is always placed before the Order by clause in a Select query.
Example Query: Select * From Student Order by marks desc;	**Example Query:** Select Stream,count(*) From STUDENT Group by Stream;

Q2. Aryan is confused while working on a table in MySQL. Help him in distinguishing between Where clause and Having clause along with one query example for each.

Ans

Where clause	**Having clause**
WHERE clause allows us to filter data from specific rows (individual rows) from a table based on certain conditions.	HAVING clause allows you to filter data from a group of rows in a query based on conditions involving aggregate values.
Aggregate functions cannot be used in a where clause directly. It can be used in a	Aggregate functions can be used in a Having clause.

	sub-query.	
	Example: Select * From Employee where salary>50000;	**Example:** Select stream,count(*) From student Group by stream Having count(*)>=2;

Q3. Class teacher of XII Commerce, Ms. Payal is working on a table namely MONTHLYTEST(Admno, Sname, Marks). She is confused between the following two queries which are giving two different outputs. Explain the outputs with justification.

QUERY1:	**QUERY2:**
Select count (*) From MONTHLYTEST;	Select count (Marks) From MONTHLYTEST;
OUTPUT: 30	**OUTPUT:** 27

Ans COUNT(*) returns the count of all rows in the table even if there is any NULL value in any of the columns, whereas COUNT() is used with Column_Name passed as an argument and counts the number of non-NULL values in a column that is given as an argument.
The Marks column must be having 3 NULL values. Hence, count(Marks) is giving the output as 27 and count(*) is giving the total number of records present in the table.

Q4. What is JOIN in MySQL? Illustrate with one example command.

Ans A JOIN clause is used to combine rows from two or more tables, based on a related column between them. *JOINS* are used to retrieve data from multiple tables in a single query

Example:

Select * From Subject,Teacher
Where Subject.sub_no=Teacher.sub_no;

Q5. Consider the relation VEHICLE given below and write the output of the following questions:

VNO	TYPE	COMPANY	PRICE	QTY
WB101	Sedan	Maruti	1200000	2
JH134	Sumo	Tata	1000000	8
MH782	SUV	Maruti	1500000	5
BH451	Van	Datsun	900000	NULL

a) SELECT COMPANY , SUM(PRICE) FROM VEHICLE GROUP BY COMPANY;
b) SELECT MIN(QTY) FROM VEHICLE;

Ans

a)

COMPANY	SUM(PRICE)
Maruti	2700000
Tata	1000000
Datsun	900000

b)

MIN(QTY)

2

05 Short Knowledge/Understanding/Application Based Questions (3 Marks)

Q1. What is the purpose of GROUP BY clause in SQL? Explain with the help of suitable example.

Ans

- The Group By clause combines all those records that have identical values in a particular field or a group of fields. This grouping results into one summary record per group if group functions are used with it.
- Grouping can be done by a column name, or with aggregate functions in which case the aggregate produces a value for each group.
- All rows with a NULL in the column are treated as if NULL was another value.

Example: Table namely STUDENT along with the fields.

STUDENT(Admno,sname,marks,DOB,stream)

Query based on the above table STUDENT:

Select Stream, count(*) From STUDENT Group by Stream;

Q2. What is the purpose of Order By clause in SQL? Explain with the help of suitable example.

Ans

- In SQL ORDER BY command is used to sort the results of the query.
- Sorting can be done on fields either in ascending order or descending order.
- Order By clause is used in the query and the keywords for ascending is (asc) and descending is (desc).
- In the absence of the keywords asc or desc in the query, sorting is done in ascending order by Default.

Example:

SELECT Salary FROM Employee ORDER BY Salary DESC;

The above statement will display the Salary of Employees in descending order from table

	Employee.
Q3.	What are Aggregate functions in SQL? Name any two.
Ans	**Aggregate functions:** These are also called multiple row functions. These functions work on a set of records as a whole, and return a single value for each column of the records on which the function is applied. An aggregate function ignores NULL values when it performs the calculation, except for the count function. Max(), Min(), Avg(), Sum(), Count() / Count(*) are the examples of multiple row functions.
Q4.	Explain the following aggregate functions with example queries: **min( ), sum( ), count( )**
Ans	Min() function returns the minimum value from a given column or expressions SELECT MIN(SALARY) FROM EMPLOYEE; Sum() function returns the sum of values in the given parameter(input) column or expression. SELECT SUM(SALARY) FROM EMPLOYEE; Count( ) function returns the total no. of values/records under the specified column or expression. SELECT COUNT(SALARY),COUNT(*) FROM EMPLOYEE;
Q5.	Consider the relation VEHICLE given below and write the output of the following questions:

VNO	TYPE	COMPANY	PRICE	QTY
WB101	Sedan	Maruti	1200000	2
JH134	Sumo	Tata	1000000	8
MH782	SUV	Maruti	1500000	5
BH451	Van	Tata	900000	NULL
MH562	Jeep	Mahindra	2500000	12
WB726	Sedan	Toyota	1500000	4

a) SELECT COMPANY , COUNT(*), FROM VEHICLE GROUP BY COMPANY;

b) SELECT AVG(QTY) FROM VEHICLE;

c) SELECT SUM(PRICE) From Vehicle where Qty>5;

Ans

a)

COMPANY	COUNT(*)
Maruti	2
Tata	2
Mahindra	1

Toyota	1

b)

AVG(QTY)

6.2

c)

SUM(PRICE)

3500000

05 Short Knowledge/Understanding/Application Based Questions (4 Marks)

Q1. Consider the relation VEHICLE given below and write the queries for the following questions:

VNO	TYPE	COMPANY	PRICE	QTY
WB101	Sedan	Maruti	1200000	2
JH134	Sumo	Tata	1000000	8
MH782	SUV	Maruti	1500000	5
BH451	Van	Tata	900000	NULL
MH562	Jeep	Mahindra	2500000	12
WB726	Sedan	Toyota	1500000	4

a) To display the cardinality of the table vehicle

b) To find the average quantity from the table vehicle

c) To find the aggregate price from the table vehicle whose quantity is greater than 5.

d) To find the minimum quantity from the table vehicle.

Ans

a) SELECT COUNT(*) FROM VEHICLE;

b) SELECT AVG(QTY) FROM VEHICLE;

c) SELECT SUM(PRICE) From Vehicle where Qty>5;

d) SELECT MIN(QTY) From Vehicle;

Q2. Consider the relation VEHICLE given below and write the queries for the following questions:

VNO	TYPE	COMPANY	PRICE	QTY
WB101	Sedan	Maruti	1200000	2
JH134	Sumo	Tata	1000000	8
MH782	SUV	Maruti	1500000	5
BH451	Van	Tata	900000	NULL
MH562	Jeep	Mahindra	2500000	12

WB726	Sedan	Toyota	1500000	4

(a) To display the company name and their average quantity company wise for those companies whose quantity is less than or equal to 5. (2marks)

b) To display number of records, type wise from the table Vehicle. (2marks)

Ans

a) Select Company, avg(Qty) From Vehicle Group By Company Having Qty<=5;

b) Select count(*), Type From Vehicle Group by Type;

Q3. Consider the following table STUDENT, and give the output for the following questions:

ADMNO	SNAME	STREAM	MARKS	DOB
101	DISHA	SCIENCE	99	12/09/2005
234	MADHU	ARTS	88	23/08/2004
145	BASANT	COMMERCE	74	31/12/2004
137	JUNAID	SCIENCE	NULL	15/10/2006
452	MARY	ARTS	74	07/02/2006

a) SELECT MIN(DOB) FROM STUDENT;

b) SELECT COUNT(MARKS) FROM STUDENT;

c) SELECT COUNT(DISTINCT(MARKS)),SUM(MARKS) FROM STUDENT;

d)SELECT AVG(MARKS) FROM STUDENT WHERE STREAM='ARTS';

Ans

a)

MIN(DOB)

23/08/2004

b)

COUNT(MARKS)

4

c)

COUNT(DISTINCT(MARKS))	SUM(MARKS)
3	335

d)

AVG(MARKS)

81

Q4. Consider the following table STUDENT, and write the commands for the following:

ADMNO	SNAME	STREAM	MARKS	DOB
101	DISHA	SCIENCE	99	12/09/2005
234	MADHU	ARTS	88	23/08/2004

145	BASANT	COMMERCE	74	31/12/2004
137	JUNAID	SCIENCE	NULL	15/10/2006
452	MARY	ARTS	74	07/02/2006

a) To display streamwise maximum marks for each stream. (2)

b) To display DOB yearwise number of students whose year of birth are same. (2)

Ans

a) SELECT STREAM,MAX(MARKS) FROM STUDENT GROUP BY STREAM;

b) SELECT YEAR(DOB),COUNT(*) FROM STUDENT GROUPBY YEAR(DOB);

Q5. Consider the relation VEHICLE given below and write the output of the following questions:

VNO	**TYPE**	**COMPANY**	**PRICE**	**QTY**
WB101	Sedan	Maruti	1200000	2
JH134	Sumo	Tata	1000000	8
MH782	SUV	Maruti	1500000	5
BH451	Van	Tata	900000	NULL
MH562	Jeep	Mahindra	2500000	12
WB726	Sedan	Toyota	1500000	4

a) SELECT COUNT(DISTINCT(PRICE)) FROM VEHICLE;

b) SELECT COMPANY, MAX(PRICE) FROM VEHICLE GROUP BY COMPANY;

c) SELECT SUM(PRICE) From Vehicle where Qty>5;

d) SELECT COUNT(DISTINCT(TYPE)) FROM VEHICLE;

Ans

a)

COUNT(DISTINCT(PRICE))

5

b)

COMPANY	MAX(PRICE)
Maruti	1500000
Tata	1000000
Mahindra	2500000
Toyota	1500000

c)

SUM(PRICE)

3500000

d)

COUNT(DISTINCT(TYPE))

	5

04 Case Based Questions (5 Marks)

Q1. Dona is the Database administrator of a famous beauty salon in Kolkata. She has prepared a rate chart for different services given by the experts for a particular day. Help her in providing the output for the following queries based on the table SALON:

Sno	Brand	Service_name	Expert_name	Rate	No_of_Cust
H1	Habib	Hair Cut	Suparna	500	15
H2	Habib	Hair Colour	Trima	1150	8
H3	Shehnaz	Hair Spa	Zarin	800	7
F1	Lotus	Face clean	Piu	1200	12
F2	Shehnaz	Facial	Tiyasha	1500	20
M	Shehnaz	Manicure	Sarita	450	5
P	Lotus	Pedicure	Priya	600	10

a) SELECT MIN(RATE) FROM SALON WHERE NO_OF_CUST>10; (1)

b) SELECT BRAND,SUM(NO_OF_CUST) FROM SALON GROUP BY BRAND; (1)

c) SELECT SERVICE_NAME,RATE FROM SALON ORDER BY RATE; (1)

d) SELECT SERVICE_NAME, RATE*NO_OF_CUST "TOTAL CHARGE" FROM SALON ORDER BY SERVICE_NAME; (2)

Ans a)

MIN(RATE)

500

b)

BRAND	SUM(NO_OF_CUST)
Habib	23
Shehnaz	32
Lotus	22

c)

SERVICE_NAME	RATE
Manicure	450

Hair Cut	500
Pedicure	600
Hair Spa	800
Hair Colour	1150
Face Clean	1200
Facial	1500

d)

Service_name	Total Charge
Face clean	14400
Facial	30000
Hair Colour	9200
Hair Cut	7500
Hair Spa	5600
Manicure	2250
Pedicure	6000

Q2. Dona is the Database administrator of a famous beauty salon in Kolkata. She has prepared a rate chart for different services given by the experts for a particular day. Help her in writing commands for the following questions based on the table SALON:

Sno	Brand	Service_name	Expert_name	Rate	No_of_Cust
H1	Habib	Hair Cut	Suparna	500	15
H2	Habib	Hair Colour	Trima	1150	8
H3	Shehnaz	Spa Hair	Zarin	800	7
F1	Lotus	Face clean	Piu	1200	12
F2	Shehnaz	Facial	Tiyasha	1500	20
M	Shehnaz	Manicure	Sarita	450	5
P	Lotus	Pedicure	Priya	600	10

a) Display brandwise number of customers. (2)

b) Display the details of the salon in descending order of rate of all the Hair related services. (2)

c) Display total clients visited on that day. (1)

Ans

a) SELECT BRAND,SUM(NO_OF_CUST) FROM SALON GROUP BY BRAND;

b) SELECT * FROM SALON WHERE SERVICE_NAME LIKE “%HAIR%” ORDER BY

RATE DESC;

c) SELECT SUM(NO_OF_CUST) FROM SALON;

Q3. Consider the relation VEHICLE of a distributor given below and write the output of the following questions:

VNO	TYPE	COMPANY	PRICE	QTY
WB101	Sedan	Maruti	1200000	2
JH134	Sumo	Tata	1000000	8
MH782	SUV	Maruti	1500000	5
BH451	Van	Tata	900000	NULL
MH562	Jeep	Mahindra	2500000	12
WB726	Sedan	Toyota	1500000	4

a) SELECT * FROM VEHICLE ORDER BY QTY; (1)

b) SELECT COMPANY, PRICE FROM VEHICLE ORDER BY PRICE DESC; (1)

c) SELECT VNO FROM VEHICLE ORDER BY TYPE; (1)

d) SELECT VNO,COMPANY,TYPE FROM VEHICLE WHERE PRICE<1500000 ORDER BY PRICE; (2)

Ans a)

VNO	TYPE	COMPANY	PRICE	QTY
BH451	Van	Tata	900000	NULL
WB101	Sedan	Maruti	1200000	2
WB726	Sedan	Toyota	1500000	4
MH782	SUV	Maruti	1500000	5
JH134	Sumo	Tata	1000000	8
MH562	Jeep	Mahindra	2500000	12

b)

COMPANY	PRICE
Mahindra	2500000
Maruti	1500000
Toyota	1500000
Maruti	1200000
Tata	1000000

Tata	900000

c)

VNO
MH562
WB101
WB726
JH134
MH782
BH451

d)

VNO	COMPANY	TYPE
BH451	Tata	Van
JH134	Tata	Sumo
WB101	Maruti	Sedan

Q 4. Agniswar works in a Bank as their IT Administrator on the following two tables CUSTOMER and ENQUIRY. Help him to write the MySQL commands for the following questions:

CUSTOMER

ACTNO	CNAME	CITY	DOA
3098	RAGHAV	KOLKATA	21/01/1995
4065	MAYANK	BHOPAL	03/01/1995
2451	MANDIRA	PATNA	12/03/1997
3472	SONAM	KOLKATA	19/08/1996
5508	RONIT	BHUBANESHWAR	05/07/1995

ENQUIRY

ACCTNO	BALANCE
3098	45000
4065	30550
2451	66000
3472	72000

5508	23000

(i) To display customer name in order of their date of Account (DOA) creation from the oldest customer to the newest customer of the bank. (2)

(ii) To display Customer's Account number, name, Date of Account creation with their balance of only those customers who live in Kolkata. (2)

(iii) To display the average balance maintained by the customers of the bank. (1)

Ans.

(i) SELECT CNAME,DOA FROM CUSTOMER ORDER BY DOA;

(ii) SELECT CUSTOMER**.**ACCTNO,CNAME,DOA,BALANCE
FROM CUSTOMER,ENQUIRY
WHERE CUSTOMER.ACCTNO=ENQUIRY.ACCTNO AND CITY="KOLKATA";

(iii) SELECT AVG(BALANCE) FROM ENQUIRY;

Name of Chapter: Introduction to computer Networks

Topics Covered : Introduction to networks, Types of network: PAN, LAN, MAN, WAN. Network Devices: modem, hub, switch, repeater, router, gateway Network Topologies: Star, Bus, Tree, Mesh. Introduction to Internet, URL, WWW, and its applications- Web, email, Chat, VoIP.

Website: Introduction, difference between a website and webpage, static vs dynamic web page, web server and hosting of a website.

Key Points

1) Introduction to Networks:

- A network is a crucial part of modern computing, allowing devices to communicate and share information. It can be a local network within a home or office or a global network like the internet.
- Networking media, also known as transmission media, provide the physical means for data to travel from one device to another. These media come in various forms, each with its own advantages and limitations.
- Ethernet cables, often categorized as twisted-pair cables, are commonly used in local area networks (LANs). They offer reliability and high-speed data transmission over short distances.
- Optical fiber is an advanced networking medium that uses light signals to transmit data. It provides high bandwidth and is immune to electromagnetic interference, making it suitable for long-distance communication.
- Microwave communication relies on radio waves in the microwave frequency range. It is frequently used for point-to-point links, such as connecting two buildings in a city.
- Radiowave communication utilizes radio frequency waves for wireless data transmission. This is common in Wi-Fi networks and mobile communication.
- Satellite communication involves the use of geostationary or low-earth-orbit satellites to relay signals over vast distances. It's essential for global communication and broadcasting.
- Infrared communication uses infrared light to transmit data wirelessly, typically over short distances. It's commonly found in remote controls and some short-range data transfer applications.

2) Types of Networks:

- Personal Area Network (PAN) is the smallest type of network, often covering a range of a few meters. Bluetooth devices, like wireless headphones or keyboards, are examples of PANs.
- Local Area Network (LAN) is a network that encompasses a relatively small geographic area, such as a home, office, or campus. LANs enable devices to share resources like printers or files.
- Metropolitan Area Network (MAN) extends over a city or a large campus. It connects multiple LANs within a geographic area, providing higher connectivity.
- Wide Area Network (WAN) spans a large geographical area and can cover a city, country, or even the entire globe. WANs use various technologies, including the internet, to connect devices over long distances.

3) Network Devices and Topologies:

- A modem is a device that modulates and demodulates digital data, allowing it to travel over analog communication lines like telephone or cable.
- Hubs are basic devices that connect multiple network devices in a LAN. They operate at the physical layer of the OSI model and simply broadcast data to all connected devices.
- Switches are more intelligent than hubs. They operate at the data link layer and forward data only to the device that needs it, reducing network traffic.
- Repeaters are used to boost signal strength and extend the range of a network. They are commonly employed in wireless networks and in situations where long cable runs are necessary.
- Routers connect different networks and determine the best path for data to travel between them. They operate at the network layer of the OSI model.
- Gateways are devices that connect networks using different communication protocols or technologies, allowing them to communicate with each other.
- **Star Topology:**
  Description: In a star topology, all devices are connected directly to a central hub or switch.

  Advantages:

  Easy to install and manage.

  Fault isolation - if one connection fails, it doesn't affect the entire network.

  High reliability for individual connections.

  Disadvantages:

  Dependent on the central hub; if it fails, the entire network can go down.

  Requires more cabling compared to some other topologies.

- **Bus Topology**:

  Description: In a bus topology, all devices share a single communication line.

  Advantages:

  Simplicity in design and installation.

  Cost-effective for small networks.

  Easy to add or remove devices.

  Disadvantages:

  Susceptible to signal reflections and collisions, which can degrade performance.

  Difficult to troubleshoot if a fault occurs.

- **Ring Topology**:

  Description: In a ring topology, devices are connected in a closed loop, where data

circulates in one direction.

Advantages:

Equal access to the network for all devices.

No collisions, as data travels in a predefined path.

Disadvantages:

If one device or connection fails, it can disrupt the entire ring.

Adding or removing devices can be complex.

- **Tree Topology**:

  Description: A tree topology combines characteristics of a star and a bus. It consists of multiple star-configured networks connected to a central bus.

  Advantages:

  Scalability and flexibility.

  Redundancy can be built in for improved reliability.

  Disadvantages:

  Complex to set up and maintain.

  Failure of the central bus can affect multiple branches.

- **Mesh Topology**:

  Description: In a mesh topology, every device is connected to every other device, creating redundant paths.

  Advantages:

  High redundancy and fault tolerance.

  No single point of failure.

  Disadvantages:

  High cabling and configuration complexity.

  Cost-prohibitive for large networks due to the number of connections.

4) Introduction to Internet and Its Applications:

- The internet is a global network of interconnected computers and servers. It's a massive network of networks, enabling communication, information sharing, and access to various services.
- URL, which stands for Uniform Resource Locator, is a standardized address system used to locate resources on the internet. It includes the protocol (e.g., "http://" or "https://"), domain name (e.g., www.example.com), and specific path to the resource.

- WWW, or the World Wide Web, is a system of interconnected webpages and multimedia content accessible via the internet. It's a significant part of what most people experience when they use the internet.
- Internet applications include web browsing, which allows users to access websites and navigate through webpages. Email is used for electronic communication, while chat applications enable real-time text-based conversations. VoIP (Voice over Internet Protocol) allows voice communication over the internet, often at lower costs than traditional phone services.

5) Website:

- A website is a collection of webpages stored on a web server and accessible via the internet. Websites can serve various purposes, such as providing information, selling products or services, or hosting online communities.
- A webpage is a single document within a website, typically composed of HTML (Hypertext Markup Language) code. Webpages can contain text, images, videos, forms, and links to other pages.
- Static web pages have fixed content that doesn't change unless manually updated by a web developer. They are suitable for presenting unchanging information.
- Dynamic web pages are generated on-the-fly based on user interactions or data from databases. They can display personalized content and are commonly used for e-commerce websites or social media platforms.
- A web server is a specialized computer that stores and delivers web content to users' web browsers. It responds to requests for webpages by sending the appropriate data.
- Hosting of a website refers to the service of making a website accessible on the internet. Web hosting providers store the website's files, ensuring it's available 24/7.

6) Web Browsers:

- Web browsers are software applications designed for accessing and viewing web content. They interpret HTML and other web technologies to display webpages.
- Commonly used web browsers include Google Chrome, Mozilla Firefox, Apple Safari, Microsoft Edge, and others, each with its features and user interface.
- Browser settings allow users to customize their browsing experience. Users can adjust preferences for security, privacy, appearance, and functionality.
- Add-ons and plug-ins are additional software components that enhance browser functionality. Examples include ad blockers, password managers, and productivity tools.
- Cookies are small text files stored on a user's device by websites they visit. Cookies are used for various purposes, including tracking user preferences, authentication, and maintaining session information.

15 Objective Question (1 Mark)

Q1.	Which of the following is an example of a LAN device? a) Modem b) Router

	c) Hub d) Gateway
Ans	c) Hub
Q2.	Which network topology connects all devices to a central hub or switch? a) Star b) Bus c) Mesh d) Tree
Ans	a) Star
Q3.	What is the function of a router in a network? a) Amplify network signals b) Connect devices within a LAN c) Determine the best path for data packets d) Broadcast data to all devices on the network
Ans	c) Determine the best path for data packets
Q4.	Which term is commonly used for the World Wide Web? a) URL b) LAN c) WWW d) PAN
Ans	c) WWW
Q5.	What does URL stand for in the context of the Internet? a) Universal Resource Locator b) Uniform Resource Link c) Unique Resource Locator d) Universal Routing Link
Ans	a) Universal Resource Locator
Q6.	Which of the following is an application of the World Wide Web? a) Email b) Chat

	c) VoIP d) All of the above
Ans	d) All of the above
Q7.	What is the primary difference between a website and a webpage? a) A website is a collection of webpages. b) A webpage is a type of website. c) They are synonyms and mean the same thing. d) A website has animations, while a webpage does not.
Ans	a) A website is a collection of webpages.
Q8.	Which type of web page can display real-time information and interact with users? a) Static web page b) Dynamic web page c) Text-only web page d) Offline web page
Ans	b) Dynamic web page
Q9.	What is the purpose of a web server in hosting a website? a) To store the website's images and videos b) To display advertisements on the website c) To process and respond to user requests for webpages d) To secure the website's domain name
Ans	c) To process and respond to user requests for webpages
Q10.	Which of the following is not a commonly used web browser? a) Google Chrome b) Mozilla Firefox c) Microsoft Word d) Apple Safari
Ans	c) Microsoft Word
Q11	What are add-ons and plug-ins in web browsers? a) Small edible snacks b) Extensions that add functionality to the browser

	c) A type of bookmark d) Browser themes
Ans	b) Extensions that add functionality to the browser
Q12	What is the primary purpose of cookies in web browsing? a) To track user activity and preferences b) To block unwanted websites c) To display advertisements d) To save webpages for offline viewing
Ans	a) To track user activity and preferences
Q13	Which network type has the smallest geographic coverage? a) PAN b) LAN c) MAN d) WAN
Ans	a) PAN
Q14	What is the main function of a gateway in a network? a) Connects devices within a LAN b) Filters incoming network traffic c) Translates between different network protocols d) Manages IP addresses
Ans	c) Translates between different network protocols
Q15	Which of the following is not an example of a network topology? a) Star b) Cloud c) Tree d) Mesh
Ans	b) Cloud
05 Assertion and reason Based question (1 Mark)	
Q1.	Assertion (A): PAN stands for Personal Area Network.

	Reason (R): PANs are typically larger in scale than LANs. a. Both A and R are true and R is the correct explanation for A. b. Both A and R are true and R is not the correct explanation for A. c. A is true, but R is false. d. A is false, but R is true.
Ans	c) A is true but R is false.
Q2.	Assertion (A): In a star topology, if center device fails, it can affect the entire network. Reason (R): Star topologies provide redundancy to prevent network failures. a. Both A and R are true and R is the correct explanation for A. b. Both A and R are true and R is not the correct explanation for A. c. A is true, but R is false. d. A is false, but R is true.
Ans	c) A is true but R is false.
Q3.	Assertion (A): Web servers are responsible for hosting websites. Reason (R): Websites and webpages are the same thing. a. Both A and R are true and R is the correct explanation for A. b. Both A and R are true and R is not the correct explanation for A. c. A is true, but R is false. d. A is false, but R is true.
Ans	c) A is true but R is false.
Q4.	Assertion (A): Cookies are a type of malicious software used to hack websites. Reason (R): Cookies are small pieces of data stored on a user's computer by websites for various purposes. a. Both A and R are true and R is the correct explanation for A. b. Both A and R are true and R is not the correct explanation for A. c. A is true, but R is false. d. A is false, but R is true.
Ans	b) Both A and R are true and R is not the correct explanation for A.
Q5.	Assertion (A): A modem is a device that connects a computer to the Internet. Reason (R): A modem is responsible for modulating and demodulating digital data for

	transmission over analog lines. a. Both A and R are true and R is the correct explanation for A. b. Both A and R are true and R is not the correct explanation for A. c. A is true, but R is false. d. A is false, but R is true.
Ans	a) Both A and R are true and R is the correct explanation for A.
10 Short Knowledge/Understanding/Application Based Questions (2 Marks)	
Q1.	You are setting up a small office network. What type of network would be most suitable, and why?
Ans	A LAN (Local Area Network) would be most suitable for a small office. LANs provide high-speed connections for devices within a limited geographical area, making them efficient for office use.
Q2.	Imagine a scenario where a company has multiple offices in different cities, and they need to connect all offices to share resources efficiently. Which type of network topology would you recommend, and why?
Ans	A MAN (Metropolitan Area Network) topology would be ideal in this scenario. MANs cover a larger geographical area than LANs and are suitable for connecting offices in different cities within a metropolitan area.
Q3.	Alice is setting up a new network for her small office. She wants a cost-effective solution that allows for easy scalability. Explain to Alice the advantages and disadvantages of using a LAN (Local Area Network) compared to a MAN (Metropolitan Area Network) in her situation.
Ans	A Local Area Network (LAN) is a network that covers a small geographical area, typically within a single building or campus. LANs are cost-effective and suitable for small offices like Alice's, offering high-speed data transfer and ease of management. However, they are limited in terms of coverage. On the other hand, a Metropolitan Area Network (MAN) covers a larger area, such as a city, and allows for scalability. The disadvantage of a MAN is that it can be more expensive to set up and maintain. Alice should consider her current needs and potential future growth when making her decision.
Q4.	Explain the differences between the Internet, the World Wide Web (WWW), and a URL (Uniform Resource Locator). How do they interrelate?
Ans	Knowledge and Understanding: The Internet is a global network of networks, while the WWW is a system of interlinked documents accessed via the Internet. A URL is a specific web address. URLs are used to access websites on the WWW through the Internet.
Q5.	Explain the role of a web server in hosting a website. How does web hosting differ from using a personal computer as a web server?
Ans	Understanding and Application: A web server stores and serves web content to users. Web hosting involves renting server space from a hosting provider, ensuring 24/7 availability and

	security. Using a personal computer as a server lacks the reliability and scalability of professional hosting.
Q6.	You are concerned about online privacy and security. Explain what cookies are, why they are used, and how you can manage them in your web browser.
Ans	Cookies are small text files that websites store on a user's device to track user preferences and activities. They are used for personalization and tracking. To manage cookies, you can typically adjust settings in your browser to block or delete them, thereby enhancing privacy and security.
Q7.	Mary is an aspiring web developer. Describe the differences between a static web page and a dynamic web page. Provide examples of when each type is typically used.
Ans	A static web page is fixed and displays the same content to all users. It is typically created using HTML and CSS and does not change based on user input or external data. Examples include informational websites and personal blogs. A dynamic web page is generated on the fly and can display different content to different users or based on user interactions. It often relies on server-side scripting languages like PHP, Python, or JavaScript. Examples include e-commerce sites, social media platforms, and web applications that require user accounts and real-time updates.
Q8.	Imagine a scenario where a company has offices in multiple cities within a country, and they need to establish a network that connects all these offices. Which type of network would you recommend, and why?
Ans	WAN, because wide area network spans beyond a single building or large campus to include multiple locations spread across a specific geographic area, or even the world.
Q9.	Describe the role of a modem in a network setup. How does it differ from a router?
Ans	A modem (modulator-demodulator) converts digital data from a computer into the analog signal required for transmission over a telephone or cable line. A router, on the other hand, manages data traffic between devices within a network and connects to the internet.
Q10.	What is DNS
Ans	DNS, or the Domain Name System, translates human readable domain names (for example, www.abc.com) to machine readable IP addresses (for example, 192.12.42.102.
05 Short Knowledge/Understanding/Application Based Questions (3 Marks)	
Q1.	John is confused about the differences between a hub, a switch, and a router in a network. Provide a detailed explanation of the functions and purposes of each of these network devices.
Ans	A hub is a basic network device that operates at the physical layer of the OSI model. It simply broadcasts data to all connected devices without intelligence. Hubs are not efficient for modern networks as they cause unnecessary network traffic. A switch operates at the data link layer and is more intelligent than a hub. It forwards data only to the specific device it's intended for, reducing network congestion and improving

	performance. A router operates at the network layer and is responsible for routing data between different networks. It connects networks and determines the best path for data transmission based on IP addresses. Routers are essential for connecting a local network to the internet.
Q2.	You are setting up a new network for a small business. Explain why choosing a star topology for this network might be a better choice than a bus topology. Include considerations such as scalability, fault tolerance, and ease of management.
Ans	In a star topology, all devices are connected directly to a central hub or switch. This topology offers several advantages for a small business network: Scalability: It's easy to add or remove devices without disrupting the entire network. Fault Tolerance: If one device fails, it doesn't affect the rest of the network. Ease of Management: Troubleshooting and managing the network is simplified as each connection is separate. In contrast, a bus topology can be more challenging to scale and maintain, as adding or removing devices can disrupt the entire network. Fault tolerance is lower, and identifying issues can be more complex.
Q3.	Elaborate on the concept of cookies in web browsing. Describe how they work, their benefits, and potential privacy concerns associated with cookies.
Ans	Cookies are small pieces of data stored on a user's device by websites they visit. They serve various purposes, including session management, personalization, and tracking. Cookies work by sending and receiving data between a user's browser and a web server. Benefits include improved user experience and website functionality. However, privacy concerns arise when cookies are used for tracking without user consent, potentially leading to invasive tracking and data collection practices.
Q4.	Sarah is interested in understanding the difference between a website and a webpage. Provide a detailed explanation, including how webpages contribute to the overall structure of a website.
Ans	A website is a collection of related webpages hosted on a web server. Each webpage is an individual document or resource that is part of the website. Webpages contribute to the overall structure of a website by containing content, such as text, images, and multimedia, and they are linked together through navigation menus and hyperlinks. A website may consist of multiple webpages, each serving a specific purpose or providing different information.
Q5.	You have been tasked with explaining the concept of VoIP (Voice over Internet Protocol) to a group of non-technical individuals. Describe what VoIP is, how it works, and some of its advantages and disadvantages.
Ans	VoIP, or Voice over Internet Protocol, is a technology that enables voice communication (phone calls) over the internet instead of traditional telephone networks. It works by converting voice

	signals into digital data packets that are transmitted over the internet and reconverted into voice at the receiving end. Advantages of VoIP include cost savings, flexibility, and integration with other digital services. However, disadvantages can include call quality issues, reliance on internet connectivity, and potential security concerns.
05 Short Knowledge/Understanding/Application Based Questions (4 Marks)	
Q1.	Briefly describe the function of following devices in a network. Modem, Hub, Switch, Repeater
Ans	● A modem is a device that modulates and demodulates digital data to allow it to be transmitted over analog communication channels, such as telephone lines, for internet access. ● A hub is a basic network device that connects multiple devices in a LAN, forwarding data to all connected devices without any intelligence to filter or manage traffic. ● A switch is a network device that operates at the data link layer (Layer 2) and intelligently forwards data only to the specific device it is intended for, improving network efficiency. ● A repeater is used to extend the range of a network by regenerating and amplifying signals, allowing data to travel over longer distances without significant degradation.
Q2.	Describe the following topologies in short. Star, Bus, Ring, Tree
Ans	● In a star network topology, all devices are connected to a central hub or switch, forming a star-like structure. All communication passes through the central hub. ● In a bus network topology, all devices are connected to a single central cable or "bus," and data is transmitted along the cable. Devices tap into the bus to send or receive data. ● Ring topology is a type of network configuration where devices are connected in a circular manner, forming a closed loop. In this setup, each device is connected to exactly two other devices, creating a continuous pathway for data transmission ● The tree network topology combines elements of both the star and bus topologies. It consists of multiple star-configured networks connected to a central bus.
Q3.	Write two examples each of the following: Static Webpage, Dynamic webpage, Browser, online text massaging app
Ans	● Static Webpages: Personal Blog: A website with unchanging content, like a personal blog. Online Brochure: Business websites with fixed information. ● Dynamic Webpages:

	Social Media Feed: Platforms like Facebook with real-time updates. E-commerce Listings: Sites like Amazon with changing product info. ● Browsers: Google Chrome: Fast, feature-rich web browser. Mozilla Firefox: Open-source, privacy-focused browser. ● Online Text Messaging Apps: WhatsApp: Popular messaging app for texts, calls, and more. Telegram: Secure messaging app with large file support.
Q4.	Answer the following questions: I. What are cookies in the context of web browsing? II. How do cookies work on websites? III. What is the primary purpose of cookies? IV. Are cookies a security risk for users?
Ans	I. Cookies are small text files that websites store on a user's device, containing data such as user preferences, login credentials, and tracking information. II. Cookies work by being sent from a web server to a user's browser when they visit a website. III. The primary purpose of cookies is to enhance the user's browsing experience. They can remember login credentials, save user preferences, and track user behavior for various purposes, including advertising and analytics. IV. While cookies themselves are not inherently a security risk, they can pose privacy concerns if misused. Some third-party cookies can track users across different websites, potentially compromising privacy.
Q5.	Write the Full form of the following: VoIP, Http, DNS, TCP/IP
Ans	VoIP: Voice over Internet Protocol HTTP: Hypertext Transfer Protocol DNS: Domain Name System TCP/IP: Transmission Control Protocol/Internet Protocol
05 Case Based Questions (5 Marks)	
Q1.	A company called TechLink Solutions has 3 branches named Branch A, Branch B, and Branch

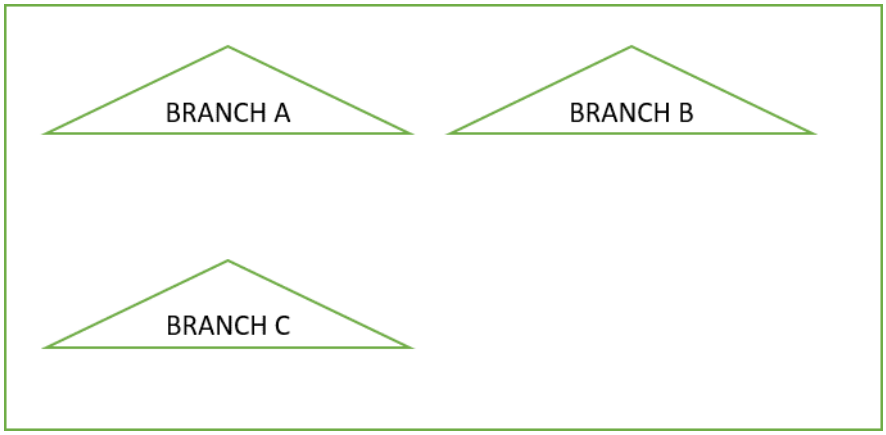
C.

Distances between Branches:

Branch A to Branch B - 50 m

Branch B to Branch C - 75 m

Branch C to Branch A - 60 m



Number of computers at each branch:

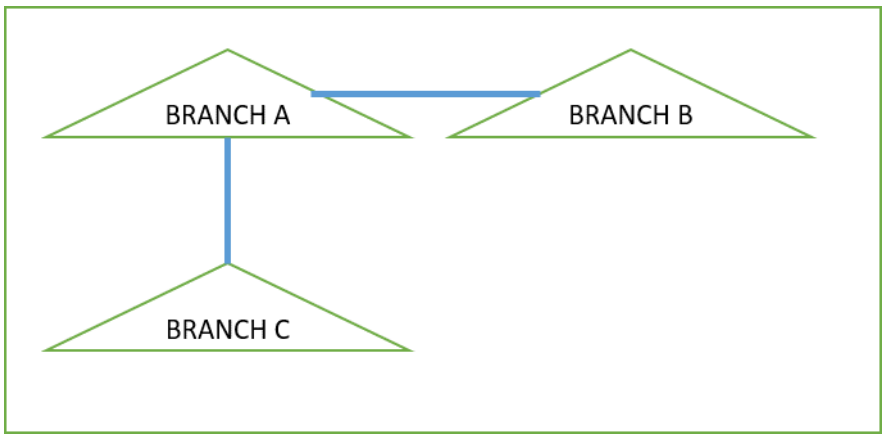
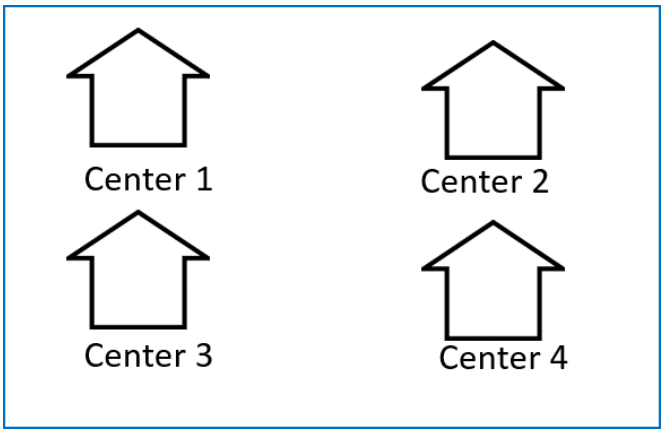
Branch A - 50

Branch B - 30

Branch C - 40

All branches are currently isolated. The company has decided to connect all the branches together.

a. Suggest a suitable cable layout for connecting the branches.
b. Suggest the most appropriate topology for the connection between the branches.
c. TechLink Solutions wants internet accessibility in all branches. Suggest a suitable technology out of.
    i. DSL
    ii. Fiber Optic
    iii. Satellite
    iv. 4G LTE
d. Suggest the placement of the following devices with justification if the company wants minimized network traffic:
    i. Repeater
    ii. Hub/Switch
e. TechLink Solutions is planning to link its head office situated in New York with the branches in remote areas. Suggest an economical way to connect them.
    i. MPLS VPN
    ii. Leased Line
    iii. Virtual Private Network (VPN)

	iv. Point-to-Point Protocol (PPP)
Ans	Answers: a)  b) Star Topology c) Fiber Optic d) Repeater – No repeater required as the distance between none of the building is around 100m or more. Hub/Switch - At each branch to connect computers within the branch. e) Virtual Private Network (VPN)
Q2.	A company named Global Networks operates 4 data centers, named Center 1, Center 2, Center 3, and Center 4 in Mumbai.  **Distances between Data Centers:** Center 1 to Center 2 - 80 meters Center 2 to Center 3 - 70 meters Center 3 to Center 4 - 60 meters

Center 4 to Center 1 - 120 meters

Center 1 to Center 3 - 150 meters

Center 2 to Center 4 - 90 meters

Number of computers in each data center:

Center 1 - 200

Center 2 - 250

Center 3 - 100

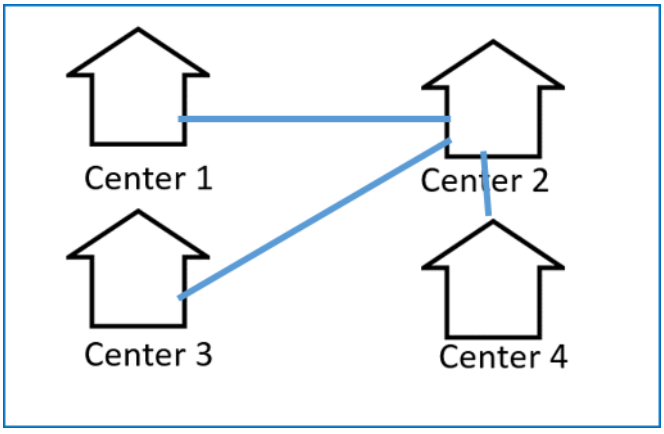
Center 4 - 120

All data centers are isolated. The company has now decided to connect all the data centers together.

a. Suggest a suitable cable layout for connecting the data centers.
b. Suggest the most appropriate topology for the connection between the data centers.
c. Suggest the most suitable place to house the server of this company.
d. Suggest the placement of the following devices with justification if the company wants minimized network traffic:
   i. Repeater
   ii. Hub/Switch
e. Global Networks is planning to establish a disaster recovery site in a city in hilly region. Suggest a way to connect it economically with reasonably high speed.

Ans

a)

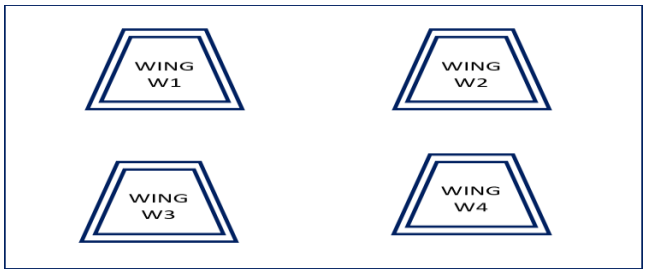


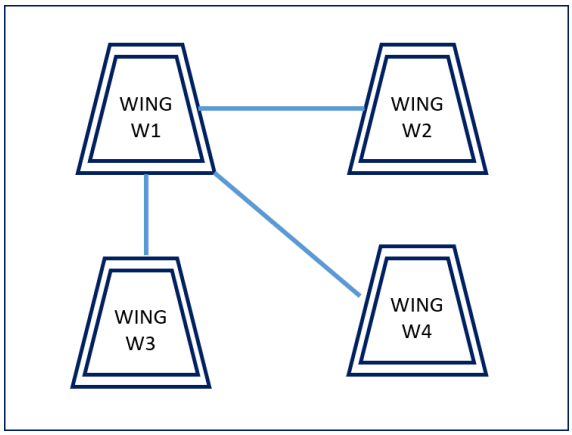
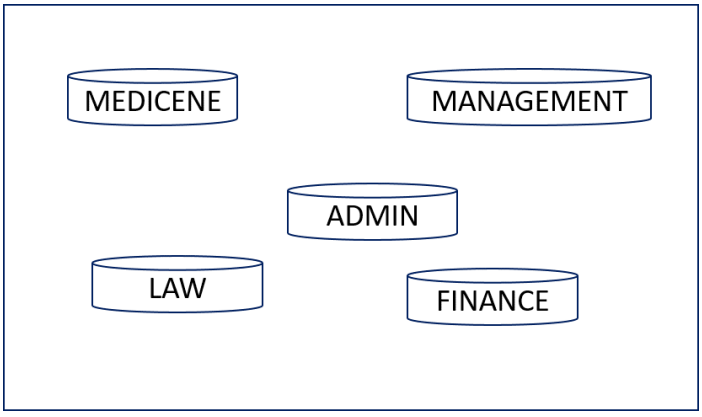
b) Star Topology

c) Centre 2 is the most suitable place to house the server of the company, as this center has maximum number of computers, thus reducing the cabling cost and increase the efficiency of the network.

d)

Repeater - Between every center to boost the signal above the 80-meter distance.
Hub/Switch - At each data center to connect servers within the data center.

	e) Radio waves
Q3.	A company in Mega Enterprises has 4 wings of buildings as shown in the diagram :  **Center to center distances between various Buildings:** W3 to W1 - 50m W1 to W2 - 60m W2 to W4 - 25m W4 to W3 - 170m W3 to W2 - 125m W1 to w4 - 90m **Number of computers in each of the wing:** W1 - 150 W2 - 15 W3 - 15 W4 - 25 Computers in each wing are networked but wings are not networked .The company has now decided to connect the wings also. i. Suggest a most suitable cable layout for the above connections. ii. Suggest the most appropriate topology of the connection between the wings. iii.Suggest the most suitable place to house the server of this company.. iv. Suggest the placement of the following devices with justification if the company wants minimized network traffic a)Repeater b)Hub / switch v. The company is planning to link its head office situated in New Delhi with the offices in hilly areas. Suggest a way to connect it economically.

Ans	i.  ii. Star topology iii. W1 is the most suitable place to house the server of the company, as this center has maximum number of computers, thus reducing the cabling cost and increase the efficiency of the network iv. a) Repeater - Between W1 to W4 to boost the signal above the 90-meter distance. b) Hub/Switch - At each data center to connect servers within the data center. v. Radio Waves
Q4.	A company named MediConnect has 5 wings of Centers, namely Medicine, Management, Law, Admin, Finance.  **Center-to-center distances between various Buildings:** Medicine to Management - 60m Medicine to law - 40m Medicine to Admin - 90m Medicine to Finance - 70m

Management to Admin - 120m

Management Finance – 100m

Management to law – 130m

Law to Admin – 70m

Law to Finance – 60m

Number of computers in each of the wing:

Medicine - 200

Management - 50

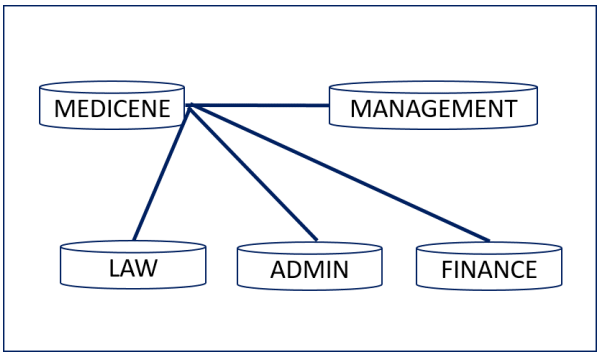
Law - 100

Admin - 75

Finance - 30

Computers in each wing are networked, but wings are not networked. The company has now decided to connect the wings also.

a. Suggest a most suitable cable layout for the above connections along with topology.
b. Suggest the most suitable location to install the main server of this institution to get efficient connectivity.
c. The company wants internet accessibility in all the wings. Suggest a suitable technology.
    i. Fiber Optic Broadband
    ii. Cable Internet
    iii. Wi-Fi
d. Suggest the placement of the following devices with justification if the company wants minimized network traffic.
    i. Repeater
    ii. Hub / Switch
e. The company is planning to link its head office situated in Chicago with the offices in remote rural areas. Suggest a way to connect it economically out of the following.
    i. DSL (Digital Subscriber Line)
    ii. Satellite Internet
    iii. Leased Line

Ans	a)  Star topology is best option here to connect all the local offices. b) Medicine office is the best place to install the server as the office contains maximum number of computers and this will efficiently manage the network traffic. c) Fibre Optic Broadband d) i. Repeater - Between Center Medicine and Admin to boost signal due to more distance of about 90m. ii. Hub / Switch - At the center where multiple computers need to communicate within the same wing. Hence Hub/Switch is required in every building. e) Satellite Internet - A cost-effective solution for connecting remote rural areas.

Q5. Mandalam Education service has 4 offices namely Engineering, Business, Admin, Media.

Office to Office distances are:

Engineering to Business - 60m

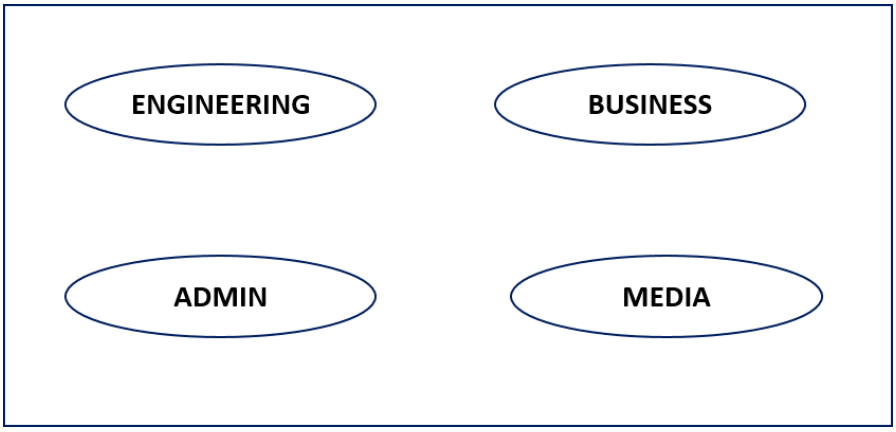
Engineering to Admin - 75m

Engineering to Media - 40m

Business to Admin - 20m

Business to Media - 90m

Admin to Media - 30m



Number of computers in each of the wing:

Engineering - 100

Business - 20

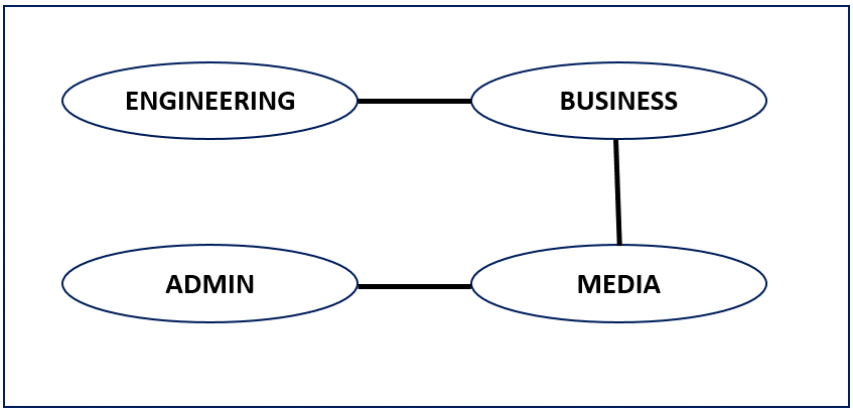
Admin - 20

Media – 30

Computers in each office are networked but offices are not networked. The company has now decided to connect the offices also.

a. Suggest a most suitable cable layout and topology for the above connections.
b. Suggest the most appropriate office for installing the server of the institution to get best connectivity. Justify your answer.
c. Which hardware device will you suggest to be procured by the institution to be installed to protect and control the internet users within the campus.

d. Suggest the device of the following with justification for connecting all the computers in each office

	i. Repeater ii. Hub / Switch iii. Router iv. Bridge e. Which of the following will you suggest to establish online face to face communication between the people in admin office of Kolkata withDelhi Head Office i. Cable TV ii. Email iii. Video Conferencing iv.Text chat
Ans	a)  b) Medicine office is the best place to install the server as the office contains maximum number of computers and this will efficiently manage the network traffic. c) Firewall d) Hub/Switch e) Video Conferencing

Name of Chapter: Societal Impacts

Topics covered : Digital footprint, net and communication etiquettes, data protection, intellectual property rights (IPR), plagiarism, licensing, and copyright, free and open-source software (FOSS) cybercrime and cyber laws, hacking, phishing, cyber bullying, overview of Indian IT Act. E-waste hazards and management. Awareness about health concerns related to the usage of technology.

Key Points

Digital footprint : A digital footprint is like a trail of footsteps we leave behind when we use the internet and electronic devices. It's all the information and actions we do online, such as playing games, chatting, or posting pictures. Just like footprints in the sand show where we've been, our digital footprint shows what we've done on the internet.

Net and communication etiquettes:

Respect Privacy: It's essential to respect others' privacy when communicating online. Avoid sharing personal information, such as phone numbers or addresses, without permission. Be cautious when posting pictures or information about others, and always ask for consent before sharing someone else's content.

Use Proper Language and Tone: Maintain a polite and respectful tone when communicating online, whether through emails, social media, or chat platforms. Avoid using offensive or hurtful language, and think before you post or send a message to ensure it's appropriate and considerate.

Avoid Cyberbullying: Cyberbullying is harmful and unacceptable. Class XII students should never engage in bullying behaviour online, such as sending hurtful messages, spreading rumors, or sharing embarrassing photos or videos. If they witness cyberbullying, they should report it to the appropriate authorities.

Cite Sources and Give Credit: When using information or content from the internet in projects, research, or social media posts, students should always give credit to the original source. Properly citing sources not only shows respect for others' work but also avoids plagiarism.

Practice Netiquette in Emails: When sending emails, follow email etiquette or "netiquette." This includes using a clear subject line, addressing recipients appropriately, keeping emails concise and to the point, and proofreading for errors before sending. Always reply promptly to emails from teachers, peers, or professional contacts.

Data protection : Data protection can be achieved using
Encryption: Encryption is the process of converting data into a code to prevent unauthorized access. It involves using algorithms to scramble the data so that only those with the decryption key can read it. For example, data on a smartphone or laptop can be encrypted to protect it from being accessed if the device is lost or stolen.

Regular Backups: Regularly backing up data to secure locations, such as external hard drives or cloud storage services, is essential. This ensures that even if data is lost due to a hardware failure or other issues, a copy of it is safely stored and can be restored.

Access Control: Implementing access controls ensures that only authorized users can access sensitive

data. This involves setting up user accounts with strong passwords , limit access to specific roles or responsibilities, and employing multi-factor authentication (MFA) for added security.

Firewalls :Firewalls block unauthorized access to a network

Antivirus and Anti-Malware Software:
Installing and regularly updating antivirus and anti-malware software helps protect against viruses, malware, and other malicious software that can compromise data security. These programs scan files and emails for potential threats and remove or quarantine them.

Intellectual property rights (IPR) :

IPR refer to a set of legal rights and protections granted to individuals or entities for their intellectual creations. These creations can include inventions, artistic works, literary works, and symbols used in commerce. IPR grants the creators or owners exclusive rights to use, control, and profit from their intellectual creations.

There are several forms of IPR, including patents, copyrights, trademarks, and trade secrets, each serving a specific purpose:

Patents: Patents protect new and inventive inventions, granting inventors exclusive rights to use, make, and sell their inventions for a specified period. This protection encourages innovation and allows inventors to benefit from their ideas.

Copyrights: Copyrights protect original literary, artistic, and musical works. This includes books, paintings, music, films, and computer software. Copyright owners have the exclusive right to reproduce, distribute, and display their works.

Trademarks: Trademarks protect symbols, names, and slogans used to identify products or services in commerce. They help consumers identify and distinguish between brands and prevent others from using similar marks.

Plagiarism : Plagiarism is the act of using someone else's words, ideas, or work without giving them proper credit and presenting it as one's own. It involves copying, paraphrasing, or reusing text, research, or creative content (such as essays, articles, artwork, or software) without obtaining permission or acknowledging the original source.

Copyright: A copyright is a collection of rights vested to someone who has created an original work. The copyright owner has the authority to transfer the rights to use / distribute to more people. When someone uses a copyrighted material without permission , it is called copyright infringement

free and open source software (FOSS) : Free and Open Source Software (FOSS) refers to software programs that are distributed with certain freedoms and characteristics. FOSS can be used, modified, and distributed by anyone without the need for restrictive licenses or fees. It promotes collaboration and sharing within the software community.
Free Software : Free software means the software is freely accessible and can be freely used, changed, improved and distributed by all wish to do so
Open Source Software : Open source software can be freely used but it does not have to be free of cost
FSF : Free Software foundation

OSI : Open Source Initiative

GPL : General Public License

CC licence : Creative Commons licences

Hacking: Hacking refers to the act of gaining unauthorized access to computer systems, networks, or data. It involves exploring, manipulating, or exploiting vulnerabilities or weaknesses in security measures to access information or perform actions that were not intended by the system's owner or administrator.

Phishing: Phishing is when someone tries to trick you online by pretending to be someone or something you trust, like a friend, a company, or a website. They send you messages, emails, or links that look real, but they're actually trying to get your personal information, like passwords or credit card numbers, so they can do bad things with it. So, it's important to be really careful and not share your important stuff with anyone you don't know or trust online.

Cyber bullying : Cyberbullying is like being mean or unkind to someone using the internet or digital devices, like sending mean messages, spreading rumours, or making fun of them online It involves individuals using technology, such as social media, text messages, or emails, to engage in hostile, hurtful, or threatening behaviour towards others. This can take the form of insults, spreading false information, sharing embarrassing photos or videos without consent, or any other actions meant to intimidate, humiliate, or harm someone emotionally..

Overview of Indian IT Act. : The Information Technology Act, 2000 is a significant piece of legislation in India that was enacted to provide legal recognition to electronic transactions and to govern various aspects of electronic commerce and digital communication. It aims to create a secure and conducive environment for electronic transactions, data protection, and the prevention of cybercrimes

India's IT Act and IT (Amendment) Act , 2008 incluses following changes :
Legal Recognition of Electronic Records . Digital Signatures: The Act establishes the framework for the use of digital signatures as a means of verifying the authenticity and integrity of electronic documents and messages
Data Protection and Privacy: The IT Act includes provisions for the protection of sensitive personal data and information.
Cybercrimes and Penalties: The Act addresses various cybercrimes, such as hacking, identity theft, and the distribution of malicious software. It prescribes penalties for cyber offenses

Electronic Governance: The IT Act promotes electronic governance by allowing government agencies to use electronic records and digital signatures for official purposes.

E-waste: hazards and management : E-waste refers to discarded electronic and electrical equipment, including computers, smartphones, televisions, refrigerators, and other electronic devices that are no longer in use or have reached the end of their operational life. E-waste poses significant environmental

and health hazards due to the presence of toxic substances

Hazards of E-Waste:

- Environmental Pollution
- Resource Depletion
- Health Risks
-

Management of E-Waste:

Collection and Segregation, Dismantling: Establish collection centres and systems for e-waste, ensuring proper segregation of different types of electronic equipment.

Recycling: Encourage the recycling of e-waste through authorized recycling facilities. Recycling can recover valuable materials and reduce the environmental impact.

Safe Disposal: Ensure the safe disposal of non-recyclable e-waste in controlled incineration facilities or hazardous waste disposal sites.

Awareness and Education: Educate consumers, manufacturers, and the public about the importance of responsible e-waste disposal and the hazards associated with improper handling.

15 Objective Question (1 Mark)	
Q1.	OSI stands for (a) Open Source Index (b) Open Source Image (c)Open Source Initiative (d) Open Syntax Index
Ans	(C) Open Source Initiative
Q2.	Which of the followings comes under the category of Cyber Crime ? (a) Identity Theft (b)Phishing (c) Cyber Bullying (d)All of above
Ans	(d) All of above
Q3.	The Information Technology act was notified in which year? (A)1999 (B) 2001 (C) 2000 (D) 2002
Ans	(C) 2000
Q4.	Our digital footprint can be created by _____ (A) visiting any website (B) Sending email (C). posting online (D). All of the above
Ans	(D) All of the above
Q5.	Digital footprints can be used to _____ (A) To know user's location (B). Tracing users activity online (C) know users like or dislikes on a website . (D) All of the above
Ans	(D). All of the above
Q6.	Which of the following are not part of Net Etiquette? (A). Be Respectful (C) . Be Responsible (C) Posting your phone number (D). All of the above
Ans	(C) Posting your phone number
Q7.	IPR stands for _____ (A) Indian Property Right (B) Intellectual Property Right (C)Intelligent Property Resource (D)Internet Property Resource
Ans	(B) Intellectual Property Right

Q8.	Copying someone's work or idea is referred as (A) IPR (B) Plagiarism (C) Patent (D) Trademarks
Ans	(B) Plagiarism
Q9.	Source code of software will be protected by _____ a. copyright b. patent c. registered trademark d. None of the above
Ans	(A) copyright
Q10.	The name and logo of the software will be protected by _____ a. copyright b. patent c. registered trademark d. None of the above
Ans	(C) Registered trademark
Q11	The _____ include right to copy (reproduce) a work, right to distribute copies of the work to the public, and right to publicly display or perform the work. a. Copyright b. Patent c. Plagarism d. None of the above
Ans	(A) Copyright
Q12	A mail or message sent to a large number of people indiscriminately without their consent is called_____
Ans	SPAM
Q13	FOSS stands for _____ a. For open source software b. Free and open set software c. Free and open source software d. Free and On Server Software
Ans	(C)Free and open source software
Q14	Which of the following is cybercrime? a. Hacking b. Phishing c. Spamming d. All of the above
Ans	(D) All of the above
Q15	_____ is an activity where fake websites or emails that look original or authentic are presented to the user. a. Phishing b. Hacking c. Spamming d. Identity theft
Ans	(A) Phishing
05 Assertion and reason Based question (1 Mark)	
	For question 1 to 5 answer A, B , C or D as per following details (A) Both Assertion and reason are true and reason correctly explain Assertion (B) Both Assertion and reason are true but reason is not correct explanation of Assertion (C)Assertion is true but reason is false (D) Assertion is false but reason is true
Q1.	**Assertion**: Assertion: Digital footprints are the traces of our online activities. **reason**: Digital footprints include information like websites visited, social media posts, and online purchases.
Ans	A) Both Assertion and reason are true and reason correctly explain Assertion
Q2.	Assertion: It is impossible to control or reduce your digital footprint. Reasoning: Being mindful of what you share online and using privacy settings can help minimize your digital footprint.
Ans	(D) Assertion is false but reason is true

Q3.	Assertion: Intellectual Property Rights protect tangible physical assets. Reasoning: IPR primarily safeguards creations of the mind, such as inventions and artistic works.
Ans	(D) Assertion is false but reason is true
Q4.	Assertion: Paraphrasing a source without proper citation is considered plagiarism. Reasoning: Paraphrasing without proper attribution is a form of plagiarism, as it involves using someone else's ideas or language without giving them credit.
Ans	A) Both Assertion and reason are true and reason correctly explain Assertion
Q5.	Assertion: E-waste is a significant environmental concern in today's digital age. Reason: Electronic devices contain hazardous materials that can harm the environment if not properly disposed of.
Ans	Ans : A) Both Assertion and reason are true and reason correctly explain Assertion
05 Short Knowledge/Understanding/Application Based Questions (2 Marks)	
Q1.	What do you mean by Phishing? Explain with the help of an example
Ans	Phishing is a attack often used to steal user data, including login credentials and credit card numbers Email / web form designed to get victims to purchase gift cards, or to give up personal email or phone numbers
Q2.	What are Importance of IPR in Software Development
Ans	Encouraging Innovation: IPR protection encourages software developers to invest time and resources in creating new and innovative software solutions, knowing that their creations will be protected from unauthorized use or duplication. Preventing Plagiarism and Piracy: IPR safeguards software from plagiarism and piracy
Q3.	State True / False for following two statements: i.Shareware software allows you to try the software before you buy it. ii. Copyright is not the right of the creator of creative/artistic work.
Ans	Ans. (i) True (ii) False
Q4.	Define Digital Property
Ans	Ans : any information created by a person that exists in digital form , either online or on some other electronic storage
Q5.	Mention any two damages caused by Spamming
Ans	• Spam reduces productivity • Spam eats up users time
05 Short Knowledge/Understanding/Application Based Questions (3 Marks)	
Q1.	Write two examples of intellectual property protected by IPR
Ans	• Patent: A patent is a type of IPR that protects new and inventive inventions • Copyright: Copyright is another form of IPR that safeguards original literary, artistic, and musical works. For instance, novels, paintings, songs, and computer software code are subject to copyright protection
Q2.	What measures you take to keep data secure
Ans	(i) Making regular backups of files (ii) Using updated antivirus (iii) Choosing strong password to restrict access and updating it on regular basis. (iv) No using public / open Wifi
Q3.	Mention any three amendments in Information Technology Amendment Act , 2008
Ans	• Digital Signature gets legal recognition • Electronic Documents get legal reconition • Imposition of penalty for damage to computer system

Q4.	What do you mean by Identity theft? Explain with the help of an example
Ans	Identity theft is the crime of obtaining the personal or financial information of another person for the sole purpose of assuming that person's name or identity to make transactions or use it to post inappropriate remarks , comments etc. Example: Sending text messages/ e-mails pretending some one else's identity
Q5.	Give example of Cyber bullying
Ans	Giving threats online, posting the victim's personal information, or comments aimed to publicly ridicule a victim.
05 Short Knowledge/Understanding/Application Based Questions (4 Marks)	
Q1.	Explain the concept of netiquette and provide three examples of good netiquette practices when communicating online Ans :
Ans	Netiquette, a term derived from "Internet" and "etiquette," refers to the set of rules and guidelines for polite and respectful behaviour when interacting with others on the internet. It is essential for maintaining a positive and productive online environment three examples of good netiquette practices: • Post correct content in respectful language • Proper Citation and Avoiding Plagiarism • Do not repeatedly post same content . i.e do not spam
Q2.	Explain the concept of "Digital Eye Strain" , "Repetitive Strain Injury " as health concerns of technology usage
Ans	Digital Eye Strain :Digital Eye Strain, also known as Computer Vision Syndrome (CVS), is a health concern related to prolonged and improper usage of technology devices such as computers, smartphones, and tablets. It results from the extended periods of screen time and can lead to various eye discomforts Repetitive Strain Injury : Due to prolonged sitting in same postures , repetitive movements of thumbs and shoulders , joints and muscles a injury disorder may occur known as RSI injuring mentioned body parts along with nerve , tendon and ligaments
Q3.	Explain the hazards associated with improper disposal of electronic waste (e-waste) and outline two key strategies for its effective management
Ans	Toxic Substances: E-waste often contains hazardous materials such as lead, mercury, cadmium, and brominated flame retardants. When these substances leach into the soil and water, they can contaminate the environment and harm ecosystems. Health Risks: Exposure to e-waste can lead to serious health issues for both workers involved in handling e-waste and communities living near disposal sites. Health problems may include respiratory issues, skin ailments, and the risk of cancer due to the release of toxic chemicals during recycling or disposal. To effectively manage e-waste, two key strategies are: • Recycling and Proper Disposal : Encourage the responsible recycling and disposal of e-waste through designated e-waste collection centers or authorized recycling facilities • Refurbishment and reuse
Q4.	Define hacking, phishing, and cyberbullying, providing a brief explanation of each and highlighting their distinct characteristics.
Ans	Hacking: Hacking refers to the unauthorized access, manipulation, or intrusion into computer systems or networks. It involves exploiting vulnerabilities in security to gain access to

	confidential information or to disrupt the functioning of systems. Hacking can be carried out for various purposes, including data theft, espionage, or simply to test security measures. It is primarily a technical activity focused on breaching security measures. Phishing: Phishing is a deceptive practice where cybercriminals impersonate trusted entities or individuals through emails, websites, or messages to trick users into revealing sensitive information such as login credentials, credit card details, or personal information. Phishing typically relies on social engineering tactics and psychological manipulation rather than technical prowess. It aims to exploit human trust and gullibility. Cyberbullying: Cyberbullying involves the use of digital communication tools like social media, messaging apps, or emails to harass, threaten, or intimidate others. It can take the form of hurtful messages, spreading rumours, sharing embarrassing photos, or any online behaviour intended to harm or harass individuals. Cyberbullying is a form of online harassment and abuse and is characterized by its social and psychological impact on victims
Q5.	Explain the concept of a "Digital Footprint" and provide two examples of how individuals can manage and maintain a positive digital footprint
Ans	A "Digital Footprint" refers to the trail of digital information or data that individuals leave behind as a result of their online activities and interactions. It encompasses everything from social media posts and online purchases to email communications and website visits. • Thoughtful Social Media Posting: Encourage students to be mindful of what they post on social media platforms. They should avoid sharing overly personal information, offensive content, or anything that could be considered harmful or inappropriate. Instead, they should focus on sharing positive and constructive content that reflects their interests and values. Regularly reviewing and adjusting privacy settings to limit the visibility of their posts can also help in maintaining control over their digital footprint. • Online Reputation Management: Teach students the importance of monitoring and managing their online reputation. They should regularly Google their own name to see what information about them is available online. If they come across negative or inaccurate information, they can take steps to address it, such as contacting website administrators to request content removal or posting positive and accurate information to overshadow any negative content. Building a professional online presence through platforms like LinkedIn can also help shape a positive digital footprint, especially for future career opportunities
05 Case Based Questions (5 Marks)	
Q1.	The GreenTech Foundation, an environmental nonprofit organization, is running an e-waste recycling program in a metropolitan city. They have partnered with local businesses and residents to collect and recycle electronic waste responsibly. To streamline their efforts, they want to develop a web-based platform that allows users to schedule e-waste pickups, learn about the recycling process, and track their environmental impact by participating in the program. 1.What is the primary goal of the GreenTech Foundation's e-waste recycling program, as described in the case study? A. To encourage the production of more electronic devices. B. To educate the public about e-waste dangers. C. To sell e-waste to international markets. D. To promote the disposal of e-waste in landfills.

	2.What functionalities should be included in the web-based platform, as mentioned in the case study? A. Online shopping for new electronic gadgets. B. Scheduling e-waste pickups, learning about recycling, and tracking environmental impact. C. Offering discounts on electronic products. D. Hosting virtual gaming tournaments. 3.Who are the primary participants in the GreenTech Foundation's e-waste recycling program, as mentioned in the case study? A. Electronic manufacturers and retailers. B. Local businesses and residents. C. International organizations and government agencies. D. Waste disposal companies. 4. How will the web-based platform contribute to the success of the e-waste recycling program, according to the case study? A. By promoting the sale of new electronic gadgets. B. By educating participants about responsible e-waste disposal and facilitating e-waste pickup. C. By offering entertainment and gaming options. D. By encouraging illegal e-waste dumping. 5.What is the potential impact of implementing the web-based platform, as mentioned in the case study? A. Increased production of disposable electronic devices. B. Improved e-waste recycling rates and reduced environmental impact. C. Lower costs for electronic device manufacturing. D. Decreased public interest in recycling.
Ans.	1)B (To educate the public about e-waste dangers.) 2)B (Scheduling e-waste pickups, learning about recycling, and tracking environmental impact.) 3)B (Local businesses and residents.) 4) B (By educating participants about responsible e-waste disposal and facilitating e-waste pickup.) 5) B (Improved e-waste recycling rates and reduced environmental impact.)
Q2.	Rahul, a friends of yours, recently encountered a distressing situation. He received an email from an unknown sender claiming to be his classmate. The email contained hurtful and offensive content, making Rahul uncomfortable and worried. Additionally, He was asked to click on a link to view an alleged embarrassing information about him. Rahul is concerned about his online safety and wants to take appropriate actions to address this issue. 1.What type of cyber threat does Rahul situation primarily represent, as described in the case study? A. Identity theft. B. Cyberbullying. C. Data breach. D. Phishing attack. 2.What is the main concern for Rahul, as mentioned in the case study? A. Increasing his online presence.

	B. Protecting his personal information. C. Identifying the unknown sender. D. Encouraging open communication. 3. What actions can Rahul take to address the situation and protect himself, as suggested in the case study? A. Ignore the email and continue using his email account as usual. B. Report the incident to a trusted adult, block the sender, and not click on any suspicious links. C. Reply to the email and ask the sender for more information. D. Share the offensive email with his friends. 4. Why is it important for individuals to be cautious when encountering suspicious emails or messages, as highlighted in the case study? A. To increase their online presence. B. To protect their personal information and safety. C. To encourage cyberbullying. D. To share offensive content with others 5.What is the potential consequence of clicking on a suspicious link in an email like the one Rahul received, according to the case study? A. Nothing significant will happen. B. The unknown sender will be revealed. C. It may lead to malware installation or data theft. D. Rahul will gain popularity online.
Ans	1.B (Cyberbullying.) 2.B (Protecting his personal information.) 3.B (Report the incident to a trusted adult, block the sender, and not click on any suspicious links.) 4.B (To protect their personal information and safety.) 5.C (It may lead to malware installation or data theft.)
Q3.	Rajat, a Class XII student, is working diligently on his final project, a research paper on renewable energy solutions. During his research, he discovers an excellent article on a reputable website that closely matches his topic. Without proper attribution, Rajat copies substantial portions of this article into his own project, thinking it would improve the quality of his work. 1.What ethical issue is presented in Rajat's situation? A. Difficulty in conducting research. B. Plagiarism and failure to give proper credit to the original author. C. The desire to excel academically. D. The need for better research skills. 2. Why is Rajat's action considered plagiarism in this case study? A. He found an excellent article. B. He used substantial portions of the article in his project without proper citation. C. He aimed to improve the quality of his work. D. His project received a high grade. 3. What are the potential consequences of Rajat's actions, as per the case study?

	A. He may receive praise for his project. B. He could face academic penalties, damage to his academic reputation, and legal issues. C. His project will become highly recognized. D. He will be encouraged to continue such practices. 4. What can Rajat do to rectify the situation and avoid further complications, according to the case study? A. Continue using the copied content and ignore the issue. B. Seek guidance from a teacher or mentor, properly cite the source, and create original content. C. Share the copied article with his classmates. D. Withdraw from the project. 5. Why is proper attribution and respect for copyright essential in academic and creative works, as highlighted in the case study? A. It's unnecessary in academic projects. B. It fosters a culture of honesty, recognizes intellectual property, and ensures fair evaluation. C. It complicates the research process. D. It's only important for winning competitions.
Ans	1.D (The need for better research skills.) 2.B (He used substantial portions of the article in his project without proper citation.) 3.B (He could face academic penalties, damage to his academic reputation, and legal issues.) 4.B (Seek guidance from a teacher or mentor, properly cite the source, and create original content.) 5.B (It fosters a culture of honesty, recognizes intellectual property, and ensures fair evaluation.)
Q4.	ABC Public School, a reputed institution in your city, has decided to explore the use of Free and Open Source Software (FOSS) to enhance their educational environment. As an Informatics Practices teacher, you have been assigned the task of guiding the school administration in this endeavor. 1. Identify one key advantage of implementing Free and Open Source Software (FOSS) in an educational institution like ABC Public School. A) Reduced cost of software licenses B) Proprietary software exclusivity C) Increased vendor lock-in D) Limited software customization 2. Which of the following FOSS operating systems is commonly used in educational institutions due to its stability, security, and a wide range of educational applications available? A) macOS B) Windows C) Linux D) Android 3: In the context of FOSS, what does the term "Open Source" mean?

	A) Software that is available for free download B) Software with no source code access C) Software with restricted use D) Software with freely accessible source code 4.ABC Public School has decided to use a FOSS office suite for word processing, spreadsheets, and presentations. Which FOSS office suite should they consider? A) Microsoft Office B) LibreOffice C) Google Workspace D) Apple iWork
Ans	1) A) Reduced cost of software licenses 2) C) Linux 3) D) Software with freely accessible source code 4) B) LibreOffice
Q5.	Rahul, a Class XII student, is known for his exceptional coding skills. Recently, he discovered a vulnerability in his school's online examination portal that allowed him to access the upcoming question papers before the exams. Instead of reporting the vulnerability, Rahul decided to use this knowledge to get an unfair advantage in the upcoming board examinations. He shared this information with a few friends who also decided to exploit the situation. 1.What type of cybercrime is Rahul involved in by exploiting the vulnerability in the school's online examination portal? A) Hacking B) Cyberbullying C) Identity theft D) Phishing 2. Which Indian legislation deals with cybercrimes and provides legal measures against various cyber offenses, including hacking? A) Indian Penal Code (IPC) B) Information Technology Act, 2000 C) Copyright Act, 1957 D) Consumer Protection Act, 2019 3.Rahul's actions, along with his friends, not only violated the school's policies but also the law. Which section of the Information Technology Act, 2000, deals with unauthorized access to computer material? A) Section 43 B) Section 65 C) Section 66 D) Section 69A 4. What can be the potential consequences for Rahul and his friends under the Information

	Technology Act, 2000, for their unauthorized access and use of the examination portal? A) A warning letter from the school B) Community service C) Imprisonment and/or fine D) Suspension from school activities 5. To avoid facing legal consequences, what should Rahul and his friends have done when they discovered the vulnerability in the school's online examination portal? A) Exploited the vulnerability discreetly B) Reported the vulnerability to the school authorities or website administrator C) Sold the information to other students D) Ignored the vulnerability
Ans	1. A) Hacking 2. B) Information Technology Act, 2000 3. C) Section 66 4. C) Imprisonment and/or fine 5. B) Reported the vulnerability to the school authorities or website administrator