National University of Computer and Emerging Sciences Lahore Campus

Uname	Age	Gender	
Sara	25	F	
Zara	42	F	
Ali	15	M	
Ahmad	19	M	
Aliya	27	F	
Tania	29	F	
Hamza	34	M	

	TOPIC	
XId	Name	Subject
1	Deep Learning	Computer Science
2	Big Data	Computer Science
3	Databases	Computer Science
4	Algorithms	Computer Science
5	Human Interactions	Philosophy

		BLOG	t_	
	Bld	Bname	Bwriter	TopicId
/	10	BigData Frameworks	Ahmad	2
	20	Generation Gap	Sara	5
152	100	Map Reduce	Hamza	2
/	30	The world of CNN	All	1
1	50	Cassandra	Ali	3
1	70	Neural Nets	Tania	1
V.	60	MongoDB	Tania	3
	120	Emerging aends	Sara	2
1	80	Hbase	Ali	1

Cfq	Blogld		Cwriter
1		20	Hamze
2	6	100	Hamzo
3		20	Zara
20	9	80	Hamza
7	0	30	Zera
9	0	50	Zara
5	O	80 1	Ali
12	0	50	Ahmad
15	0	50	Tania

Write sql queries and rdd

- a) Find the name and age of the users who have never written any blog and have naver given any comment.
- b) Find the name of the users who have given comments on all the blogs written in the Computer science area.
- c) Find the Cwriter whose comments on the topic name "Database"

CLO # 1: Basic Concept of RDD

Q2: [10 marks] Suppose you have a dataset containing information about students and their grades. Each record in the dataset consists of a tuple (student_id, [(subject, grade)]), where student_id is the unique identifier of the student and [(subject, grade)] represents a list of tuples containing subject and the corresponding grade obtained by the student in that subject.

Tasks:

- Calculate the average grade for each student across all subjects.
- Find the highest grade obtained by each student.

Note: Don't used DataFrame

Spring 2024

FAST School of Computing

Page 2 of 3



National University of Compager & Lahore Crimir is

Question 3: Marks 5

CLO # 1: Basic concept of Hive

a) What are the different types of tables supported by hine b) Explain the role of metastore in Hive. Why is 1 in portent