

Part A (2+3+2+3 marks)

Write SQL script to do the following operations (for question 1 and 2).

1. Create a database with name “University”.
2. Create following tables under “University” database.

“Institute “, with fields Institute_Name, Institute_ID

“Programs”, with fields Program_Name, Program_ID, Institute_Name, Number_of_seats

3. Write Sqoop command to transfer the table “Program” to HDFS by excluding ‘Institute Name’.
4. Create Hive table for “Program” and load the data into it.

Part B (15 marks)

Execute the below question

The dataset has fields like IP address, Date, Request type, URL and Status Code. Each fields are separated by tab ('\t').

Find out the count of different response code (Status Code) for individual IP addresses using GET and Post request types.

Expected output should be in the form

10.130.2.1 GET 200	35
10.130.2.1 POST 200	5
10.130.2.1 GET 302	12
10.131.2.1 POST 200	25