

**BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI (RAJ)**  
**I SEMESTER 2021-2022**  
**ASSIGNMENT-1**

**Course No.:** CS/SS G527  
**Deadline:** 28<sup>th</sup> September

**Course Title:** Cloud Computing  
**Maximum Marks:** 40 (10%)

---

**Note:**

- **Group of maximum five students**
- **Deliverables: (a) Well commented and neatly formatted source code (b) Design document containing system design, and the rationale behind design choices.**
- **Submit deliverables on Canvas**

---

In this assignment, you are required to develop a REST-based service for the following dataset. This service will accept a query in the form of one of the patterns mentioned below and translate into MapReduce jobs, Spark job and Storm job. Service should run these three jobs separately and return the following in a JSON object:

- Time taken for Hadoop MapReduce execution
- Time taken for Spark execution
- Time taken for Storm execution
- Result of the query separately for each job.

Expected SQL query clauses and their combinations to be supported are given below.

- (1) SELECT (\* or any set of columns), FROM, WHERE (=, >, <, >=, <=, <>, *LIKE*, *IN*)
- (2) SELECT, FROM, WHERE, JOIN (NATURAL JOIN, INNER JOIN, LEFT OUTER JOIN)
- (3) SELECT, FROM, WHERE, GROUP BY, HAVING (count, max, min, avg, sum)

Description of the dataset is given below:

Users file stores the demographic information about the users. The zipcodes file contains the city and state information for the zipcodes. Zipcode is referenced by users file. Movies file stores the information about movies. The last 19 fields in movies table are the genres, a 1 indicates the movie is of that genre, a 0 indicates it is not; movies can be in several genres. Rating file contains the ratings given by users for movies. Userid and movieid refer to the users and movies files respectively.

**Users:**

userid | age | gender | occupation | zipcode

**Zipcodes:**

zipcode | zipcodetype | city | state

**Movies table:**

movieid | title | releasedate | unknown | Action | Adventure | Animation |  
Children | Comedy | Crime | Documentary | Drama | Fantasy |  
Film\_Noir | Horror | Musical | Mystery | Romance | Sci\_Fi |  
Thriller | War | Western |

**Rating:**

userid | movieid | rating | timestamp

These files can be downloaded from the links given below.

[https://drive.google.com/file/d/1LDngsYyPyilgkVdVMuenHzRv\\_PR5TPjF/view?usp=sharing](https://drive.google.com/file/d/1LDngsYyPyilgkVdVMuenHzRv_PR5TPjF/view?usp=sharing)

--&--