HIVE ASSIGNMENT 1

1. Table Creation

: jdbc:hive2://localhost:10000> CREATE TABLE IF NOT EXISTS RATING (userID int, movieID int, ratings float, timestamps String)

. . . . . . . . . . . . . . . .> COMMENT 'Movie Rating details'

. . . . . . . . . . . . . . . .> ROW FORMAT DELIMITED

. . . . . . . . . . . . . . . .> FIELDS TERMINATED BY ','

. . . . . . . . . . . . . . . .> LINES TERMINATED BY '\n'

. . . . . . . . . . . . . . . .> STORED AS TEXTFILE;

1. Loading of Data

0: jdbc:hive2://localhost:10000> LOAD DATA LOCAL INPATH '/home/aditi/ml-latest-small

. . . . . . . . . . . . . . . .> /ratings.csv' OVERWRITE INTO TABLE RATING;

1. Display ratings data

0: jdbc:hive2://localhost:10000> SELECT \* FROM RATING LIMIT 10;

INFO : Concurrency mode is disabled, not creating a lock manager

+----------------+-----------------+-----------------+--------------------+

| rating.userid | rating.movieid | rating.ratings | rating.timestamps |

+----------------+-----------------+-----------------+--------------------+

| 1 | 1 | 4.0 | 964982703 |

| 1 | 3 | 4.0 | 964981247 |

| 1 | 6 | 4.0 | 964982224 |

| 1 | 47 | 5.0 | 964983815 |

| 1 | 50 | 5.0 | 964982931 |

| 1 | 70 | 3.0 | 964982400 |

| 1 | 101 | 5.0 | 964980868 |

| 1 | 110 | 4.0 | 964982176 |

| 1 | 151 | 5.0 | 964984041 |

| 1 | 157 | 5.0 | 964984100 |

+----------------+-----------------+-----------------+--------------------+

1. Display rating wise count

0: jdbc:hive2://localhost:10000> SELECT rating.ratings, COUNT(\*) FROM RATING GROUP BY rating.ratings;

+-----------------+--------+

| rating.ratings | \_c1 |

+-----------------+--------+

| 0.5 | 1370 |

| 1.0 | 2811 |

| 1.5 | 1791 |

| 2.0 | 7551 |

| 2.5 | 5550 |

| 3.0 | 20047 |

| 3.5 | 13136 |

| 4.0 | 26818 |

| 4.5 | 8551 |

| 5.0 | 13211 |

+-----------------+--------+