Aim:- Perform Analytics on the Books Database.

\*\*\*Before performing analytics create index on all the tables

\*\*\*Go through every fields in Book Recommendation Project dataset and appropriately create the tables

Create users,books and ratings table in hive terminal

Create users table in hive terminal through below command

**CREATE TABLE IF NOT EXISTS users ( user\_idint, location String, age int, updated\_at timestamp) COMMENT 'users details' ROW FORMAT DELIMITED FIELDS TERMINATED BY '|' LINES TERMINATED BY '\n' STORED AS TEXTFILE;**

* **Similarly create tables for other dataset like Books and ratings**
* **Next execute the below commands for all tables that are created and observe the**

**difference**

**Hive Tables:(available three tables users,books and ratings in hive)**

**show tables;**

Analyze the ratings table through below command

**ANALYZE TABLE ratings COMPUTE STATISTICS**

Describe ratings table

**DESCRIBE extended ratings;**

Describe the ratings table through below command.

**DESCRIBE FORMATTED ratings;**

* **Next create index on the above created tables and observe the difference**

Create index user\_index on table user through below command

Show index on users table.

**Show index on users;**

Perform the below analytics on Books database

1. How many users are given rating for any book;
2. Top 10 rated book;
3. How many books have not rated;
4. Top 10 users who have given maximum rating for maximum book; //which user has given 10 ratings for maximum no. of books.//
5. how many  books rated by user id 104,200, and 700;

**create table users(**

**uid int , location string , age int , updated\_at timestamp)**

**row format delimited fields terminated by ‘,’ lines terminated by ‘\n’ stored as textfile;**

load data local inpath '/home/ak/Downloads/ratings.csv' overwrite into table rating;

**create table books(**

**isbn string , title string , author string, yob string, publisher string, urls string , urlm string , urll string)**

**row format delimited fields terminated by ‘,’ lines terminated by ‘\n’ stored as textfile;**

**create table ratings(**

**uid int , isbn string , rating int)**

**row format delimited fields terminated by ‘,’ lines terminated by ‘\n’ stored as textfile;**