#table zomato\_jassi\_source

#total number of rows

select count(\*) from zomato\_jassi\_example;

29765 rows

#total number of columns

SELECT COUNT(\*) FROM INFORMATION\_SCHEMA.COLUMNS WHERE table\_name = 'zomato\_jassi\_example';

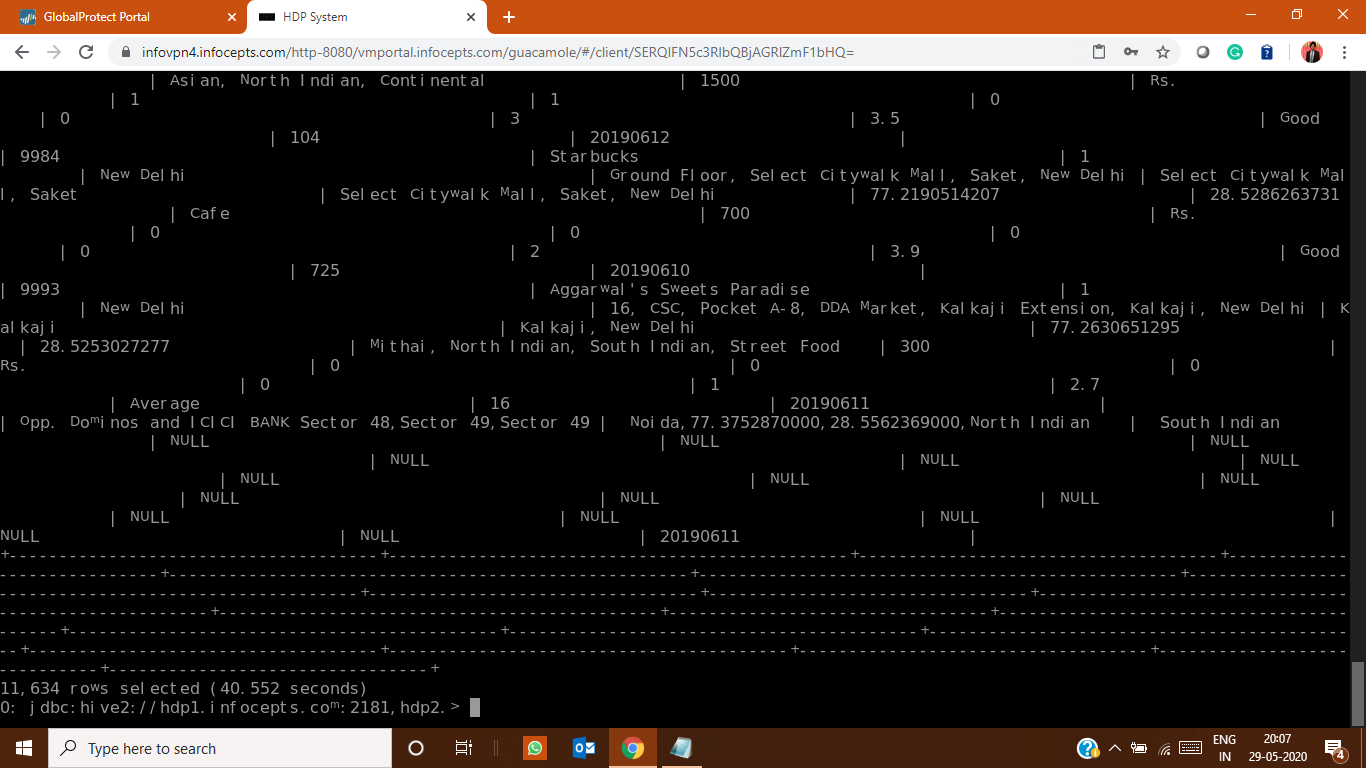
describe zomato\_jassi\_example;

21

#getting distinct rows

select distinct \* from zomato\_jassi\_source

11634 rows



#number of duplicates

29765-11634=18131

#duplicate check

select `restaurant id`, `restaurant name`, `country code`, city, address, locality,

`locality verbose`, longitude, latitude, cuisines, `average cost for two`, currency, `has table booking`,

`has online delivery`, `is delivering now`, `switch to order menu`, `price range`,

`aggregate rating`, `rating text`, votes, count(\*)

from zomato\_jassi\_example

group by `restaurant id`, `restaurant name`, `country code`, city, address, locality,

`locality verbose`, longitude, latitude, cuisines, `average cost for two`, currency, `has table booking`,

`has online delivery`, `is delivering now`, `switch to order menu`, `price range`,

`aggregate rating`, `rating text`, votes

having count(\*) > 1;

2455

=========================================================================================

#table dim\_country\_jassi

#total number of rows

select count(\*) from dim\_country\_jassi;

17

#total number of columns

describe dim\_country\_jassi

2

#getting distinct rows

select distinct \* from dim\_country\_jassi;

17

#numbers of null

select count(\*) from dim\_country\_jassi where `country code` is null OR `country` is null;

2

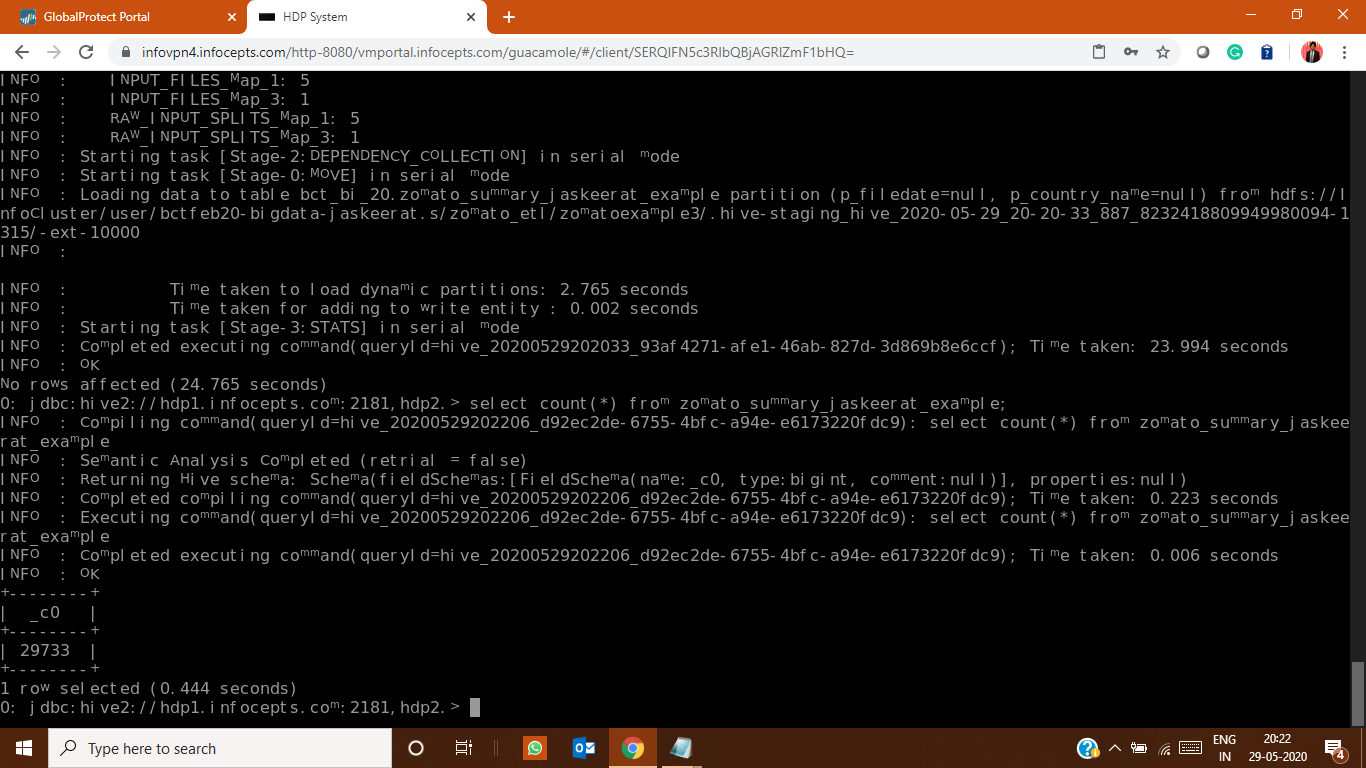
========================================================================================

#table zomato\_summary\_jaskeerat\_managed and zomato\_summary\_jaskeerat\_example

#total number of rows

select count(\*) from zomato\_summary\_jaskeerat\_example;

29733

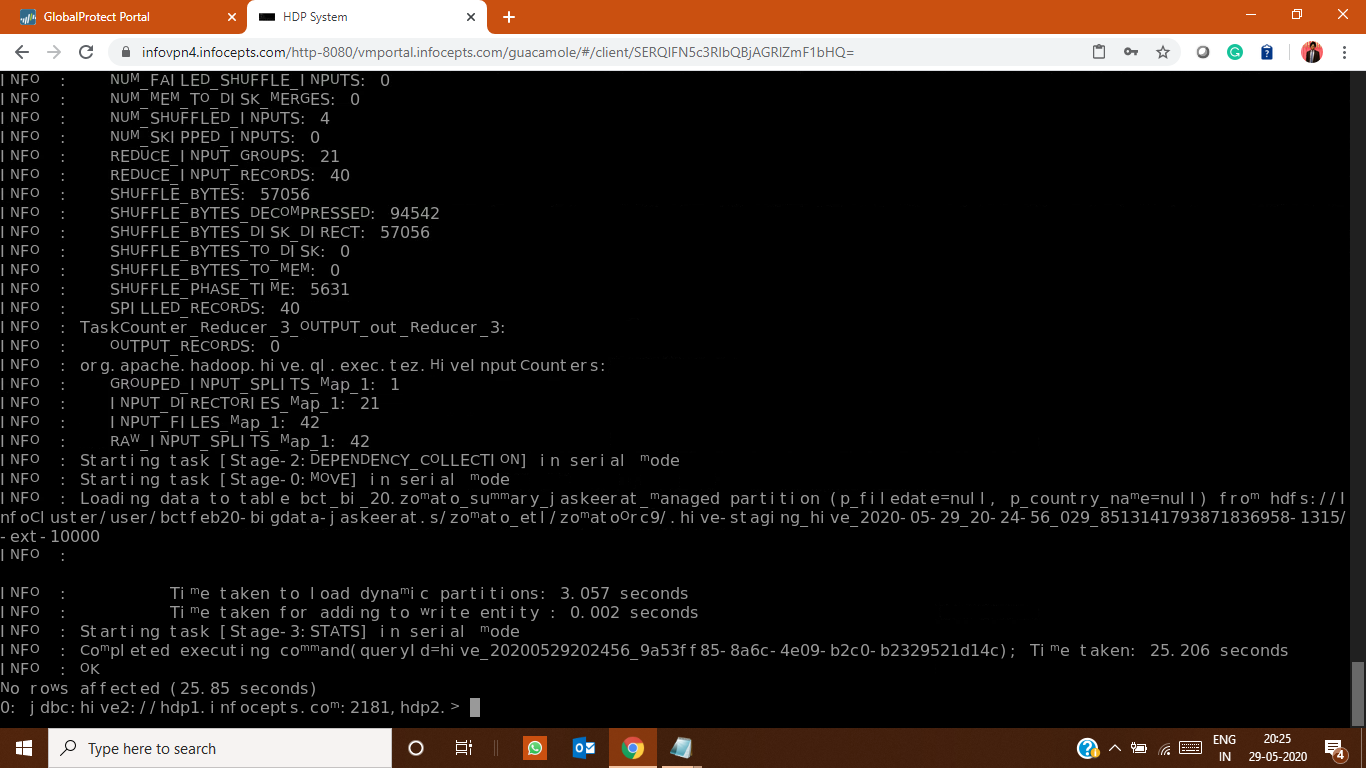


#total number of columns

describe zomato\_summary\_jaskeerat\_example;

26

after running- insert overwrite table zomato\_summary\_jaskeerat\_managed select distinct \* from zomato\_summary\_jaskeerat\_managed;



#getting distinct rows

select distinct \* from zomato\_summary\_jaskeerat\_example;

11602



#getting nulls

select votes from zomato\_summary\_jaskeerat\_managed where votes is NULL;

6

UPDATE zomato\_summary\_jaskeerat\_managed SET `votes` = "NA" WHERE `votes` IS NULL;

select `locality` from zomato\_summary\_jaskeerat\_managed where `locality` is NULL;

6

UPDATE zomato\_summary\_jaskeerat\_managed SET `locality` = "NA" WHERE `locality` IS NULL;

select `locality verbose` from zomato\_summary\_jaskeerat\_managed where `locality verbose` is NULL;

6

UPDATE zomato\_summary\_jaskeerat\_managed SET `locality verbose` = "NA" WHERE `locality verbose` IS NULL;

select `address` from zomato\_summary\_jaskeerat\_managed where `address` is NULL;

6

UPDATE zomato\_summary\_jaskeerat\_managed SET `address` = "NA" WHERE `address` IS NULL;

select `longitude` from zomato\_summary\_jaskeerat\_managed where `longitude` is NULL;

6

UPDATE zomato\_summary\_jaskeerat\_managed SET `longitude` = "NA" WHERE `longitude` IS NULL;

select `latitude` from zomato\_summary\_jaskeerat\_managed where `latitude` is NULL;

6

UPDATE zomato\_summary\_jaskeerat\_managed SET `latitude` = "NA" WHERE `latitude` IS NULL;

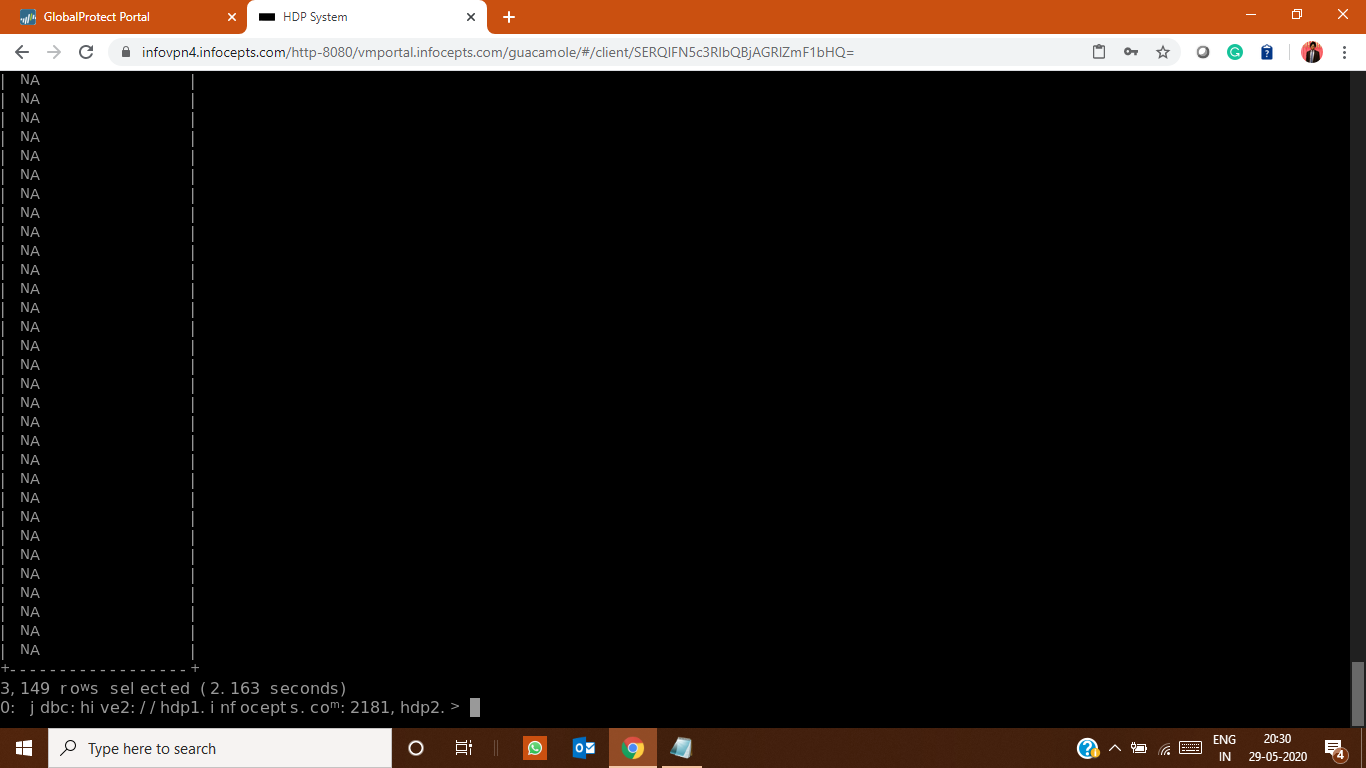
select `rating text` from zomato\_summary\_jaskeerat\_managed where `rating text` is NULL;

6

UPDATE zomato\_summary\_jaskeerat\_managed SET `rating text` = "NA" WHERE `rating text` IS NULL;

#anomalies in m\_rating\_colour coloumn

select m\_rating\_colour from zomato\_summary\_jaskeerat\_example where m\_rating\_colour="NA";

3149

**Data column validation of data Input from source table to summary table:**

|  |  |  |  |
| --- | --- | --- | --- |
| zomato\_jassi\_source | zomato\_summary\_jaskeerat\_example | True/False | True/False |
| `Restaurant ID` | `Restaurant ID` | TRUE | TRUE |
| `Restaurant Name` | `Restaurant Name` | TRUE | TRUE |
| `Country Code` | `Country Code` | TRUE | TRUE |
| `City` | `City` | TRUE | TRUE |
| `Address` | `Address` | TRUE | TRUE |
| `Locality` | `Locality` | TRUE | TRUE |
| `Locality Verbose` | `Locality Verbose` | TRUE | TRUE |
| `Longitude` | `Longitude` | TRUE | TRUE |
| `Latitude` | `Latitude` | TRUE | TRUE |
| `Cuisines` | `Cuisines` | TRUE | TRUE |
| `Average Cost for two` | `Average Cost for two` | TRUE | TRUE |
| `Currency` | `Currency` | TRUE | TRUE |
| `Has Table booking` | `Has Table booking` | TRUE | TRUE |
| `Has Online delivery` | `Has Online delivery` | TRUE | TRUE |
| `Is delivering now` | `Is delivering now` | TRUE | TRUE |
| `Switch to order menu` | `Switch to order menu` | TRUE | TRUE |
| `Price range` | `Price range` | TRUE | TRUE |
| `Aggregate rating` | `Aggregate rating` | TRUE | TRUE |
| `Rating text` | `Rating text` | TRUE | TRUE |
| `Votes` | `Votes | TRUE | TRUE |

**Data validation for comparing the country code on both table:**

|  |  |  |  |
| --- | --- | --- | --- |
| dim\_country\_jassi | zomato\_summary\_jaskeerat\_example | True/False | True/Faslse |
| 1 : India | 1 : India | True | True |
| 14 : Australia | 14 : Australia | True | True |
| 30 : Brazil | 30 : Brazil | True | True |
| 37 : Canada | 37 : Canada | True | True |
| 94 : Indonesia | 94 : Indonesia | True | True |
| 148 : New Zealand | 148 : New Zealand | True | True |
| 162 : Phillipines | 162 : Phillipines | True | True |
| 166 : Qatar | 166 : Qatar | True | True |
| 184 : Singapore | 184 : Singapore | True | True |
| 189 : South Africa | 189 : South Africa | True | True |
| 191 : Sri Lanka | 191 : Sri Lanka | True | True |
| 208 : Turkey | 208 : Turkey | True | True |
| 214 : UAE | 214 : UAE | True | True |
| 215 : United Kingdom | 215 : United Kingdom | True | True |
| 216 : United States | 216 : United States | True | True |