Database Technologies

UE19CS344

6th Semester, Academic Year 2021-22

Week #: 4 - LOAD LARGE DATA

(A1)

Date: 15/2/2022

Name : SUMUKH RAJU BHAT	SRN : PES1UG19CS519	Section : H

(RDBMS used - postgres

Miniworld - Ticket Reservation System)

1. DDL for a schema called TICKETS

Command:

create table TICKETS(tid int, seat_no int, price int, isBooked int, date date, primary key (tid, seat_no));

Output:

```
company=# create table TICKETS(tid int, seat_no int, price int, isBooked int, date date, primary key (tid, seat_no));
CREATE TABLE
```

```
company=# \d TICKETS;
                Table "public.tickets"
  Column  |  Type   | Collation | Nullable | Default
----------+---------+-----------+----------+---------
 tid      | integer |           | not null |
 seat_no  | integer |           | not null |
 price    | integer |           |          |
 isbooked | integer |           |          |
 date     | date    |           |          |
Indexes:
    "tickets_pkey" PRIMARY KEY, btree (tid, seat_no)
```

2. Load large amount of data without insert statement

Code:

```
COPY TICKETS(tid, seat_no, price, isbooked, date)
from '/home/sumukhbhat2701/Downloads/myFile0.csv'
DELIMITER ','
CSV HEADER;
```

Output:

```
company=# COPY TICKETS(tid, seat_no, price, isbooked, date)
from '/home/sumukhbhat2701/Downloads/myFile0.csv'
DELIMITER ','
CSV HEADER;
COPY 12789
```

- COPY commands allows PostgresSQL server to move data to/from database tables to standard file system files.

- When using the command, each field in the file is inserted sequentially to the specified column. Table columns not specified in the column list get their default values.

- Files in COPY are read/written directly by the server and not by the client application. Therefore, it must be located on or accessible to the database server machine, not the client either.

- It is necessary to grant SELECT privilige on the table read by COPY to and INSERT previlige in the table where the values are inserted with COPY FROM

- It cannot work for views.

- COPY to and COPY from are the 2 variants of COPY command

- DELIMTER and HEADER options are used in the command to specify the file type.

3. **Explain-analyze**:

```
company=# explain analyze select * from tickets;
                                                  QUERY PLAN
-------------------------------------------------------------------------------------------------------------
 Seq Scan on tickets  (cost=0.00..209.89 rows=12789 width=20) (actual time=0.470..5.194 rows=12789 loops=1)
 Planning Time: 2.167 ms
 Execution Time: 6.522 ms
(3 rows)
```

```
dbt519=# explain analyze select * from tickets where price < 500 and isbooked = 0 and date between '2019-08-01' and '2019-12-29';
                                                 QUERY PLAN
-------------------------------------------------------------------------------------------------------------
 Seq Scan on tickets  (cost=0.00..337.78 rows=9 width=20) (actual time=0.489..5.243 rows=6 loops=1)
   Filter: ((price < 500) AND (date >= '2019-08-01'::date) AND (date <= '2019-12-29'::date) AND (isbooked = 0))
   Rows Removed by Filter: 12783
 Planning Time: 0.214 ms
 Execution Time: 5.286 ms
(5 rows)
```

Note that it is doing a sequential scan, as there are no secondary indexes setup.

(The 2nd screenshot is one from the next assignments. Change in OS from ubuntu to popOS is the reason for change in terminal background color in the above screenshots)

4. select count(*) statement:

Code:

select count(*) from TICKETS;

Output:

```
company=# select count(*) from TICKETS;
 count
-------
 12789
(1 row)

company=#
```