**Queries-**

Copying Data from the S3 bucket to Redshift-

COPY dev.public.ticketviolation FROM 's3://datadenbucket/combined\_csv1.csv' IAM\_ROLE 'arn:aws:iam::724145206969:role/DataDenRole' FORMAT AS CSV DELIMITER ',' QUOTE '"' DATEFORMAT 'MM/DD/YYYY' REGION AS 'us-west-1'

Creating table in Redshift –

Create Table ticketviolation (

summons\_number varchar,

plate\_id varchar,

registration\_state varchar,

plate\_type varchar,

issue\_date date,

violation\_code varchar,

vehicle\_body\_type varchar,

vehicle\_make varchar,

issuing\_agency varchar,

street\_name varchar

);

Performing Queries-

**1. Number of ticket violations per year**

Select extract(year from issue\_date) as year, count(summons\_number) TicketCount from ticketviolation group by year having year in (2013,2014,2015,2016,2017) order by year;

**2. Count of violation codes in NY**

Select distinct registration\_state, violation\_code, count(violation\_code) as code\_count from ticketviolation where registration\_state='NY' group by registration\_state,violation\_code order by code\_count desc;

**3. Max number of violation codes in each regsitration state or violations most occurred in country**

select registration\_state, max(vc) as max\_code\_count from (select count(violation\_code) as vc, registration\_state from ticketviolation group by registration\_state) group by registration\_state order by max\_code\_count desc;

**4. violations based on plate id and vc**

select plate\_id, count(summons\_number) as no\_of\_violations from ticketviolation group by plate\_id;