

```
create schema redshift_demo;
```

```
create table redshift_demo.users(  
  userid integer not null distkey sortkey,  
  username char(8),  
  firstname varchar(30),  
  lastname varchar(30),  
  city varchar(30),  
  state char(2),  
  email varchar(100),  
  phone char(14),  
  likesports boolean,  
  liketheatre boolean,  
  likeconcerts boolean,  
  likejazz boolean,  
  likeclassical boolean,  
  likeopera boolean,  
  likerock boolean,  
  likevegas boolean,  
  likebroadway boolean,  
  likemusicals boolean);
```

```
create table redshift_demo.venue(  
  venueid smallint not null distkey sortkey,  
  venuename varchar(100),  
  venuecity varchar(30),  
  venuestate char(2),  
  venuesseats integer);
```

```
create table redshift_demo.category(  
  catid smallint not null distkey sortkey,  
  catgroup varchar(10),  
  catname varchar(10),  
  catdesc varchar(50));
```

```
create table redshift_demo.date(  
  dateid smallint not null distkey sortkey,  
  caldate date not null,  
  day character(3) not null,  
  week smallint not null,  
  month character(5) not null,  
  qtr character(5) not null,  
  year smallint not null,  
  holiday boolean default('N'));
```

```
create table redshift_demo.event(  
  eventid integer not null distkey,  
  venueid smallint not null,  
  catid smallint not null,
```

```
dateid smallint not null sortkey,  
eventname varchar(200),  
starttime timestamp);
```

```
create table redshift_demo.listing(  
listid integer not null distkey,  
sellerid integer not null,  
eventid integer not null,  
dateid smallint not null sortkey,  
numtickets smallint not null,  
priceperticket decimal(8,2),  
totalprice decimal(8,2),  
listtime timestamp);
```

```
create table redshift_demo.sales(  
salesid integer not null,  
listid integer not null distkey,  
sellerid integer not null,  
buyerid integer not null,  
eventid integer not null,  
dateid smallint not null sortkey,  
qtysold smallint not null,  
pricepaid decimal(8,2),  
commission decimal(8,2),  
saletime timestamp);
```

```
copy redshift_demo.users from 's3://mybucketashu1/allusers_pipe.txt'  
iam_role 'arn:aws:iam::751448659042:role/myRedshiftRole' delimiter '|' region 'us-east-1';
```

```
copy redshift_demo.venue from 's3://mybucketashu1/venue_pipe.txt' iam_role  
'arn:aws:iam::751448659042:role/myRedshiftRole' delimiter '|' region 'us-east-1';
```

```
copy redshift_demo.category from 's3://mybucketashu1/category_pipe.txt' iam_role  
'arn:aws:iam::751448659042:role/myRedshiftRole' delimiter '|' region 'us-east-1';
```

```
copy redshift_demo.date from 's3://mybucketashu1/date2008_pipe.txt' iam_role  
'arn:aws:iam::751448659042:role/myRedshiftRole' delimiter '|' region 'us-east-1';
```

```
copy redshift_demo.event from 's3://mybucketashu1/allevnts_pipe.txt' iam_role  
'arn:aws:iam::751448659042:role/myRedshiftRole' delimiter '|' region 'us-east-1';
```

```
copy redshift_demo.listing from 's3://mybucketashu1/listings_pipe.txt' iam_role  
'arn:aws:iam::751448659042:role/myRedshiftRole' delimiter '|' region 'us-east-1'
```

```
copy redshift_demo.sales from 's3://mybucketashu1/sales_tab.txt' iam_role  
'arn:aws:iam::751448659042:role/myRedshiftRole' delimiter '\t' region 'us-east-1' timeformat  
'MM/DD/YYYY HH:MI:SS';
```

```
select * from redshift_demo.sales;
```

```

// eda on the data
        // Find the total sales on a given calendar date.
        select sum(qtysold)
from redshift_demo.sales, redshift_demo.date
where sales.dateid = date.dateid
and caldate='2008-01-05';

// Find the top 10 buyers by the quantity that they purchased.

        select firstname, lastname, total_quantity
        from (select buyerid, sum(qtysold) total_quantity
        from redshift_demo.sales
        group by buyerid
        order by total_quantity desc limit 10) Q, redshift_demo.users
        where Q.buyerid = userid
        order by Q.total_quantity desc;

        // Find events in the 99.9 percentile in terms of all-time gross sales
        select eventname, total_price
from (select eventid, total_price, ntile(1000) over (order by total_price desc) as percentile
        from (select eventid, sum(pricepaid) total_price
        from redshift_demo.sales
        group by eventid )) Q, redshift_demo.event E
Where Q.eventid = E.eventid and percentile = 1
order by total_price desc;

```