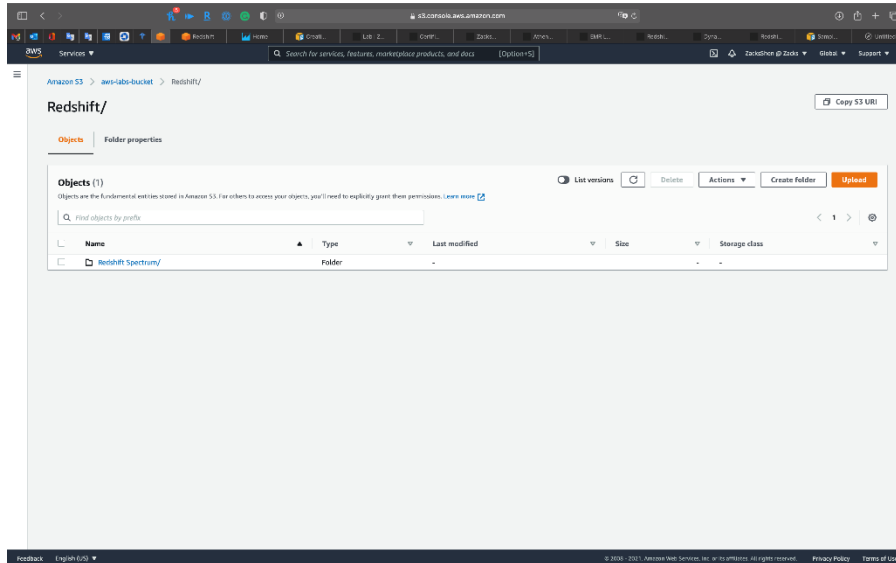


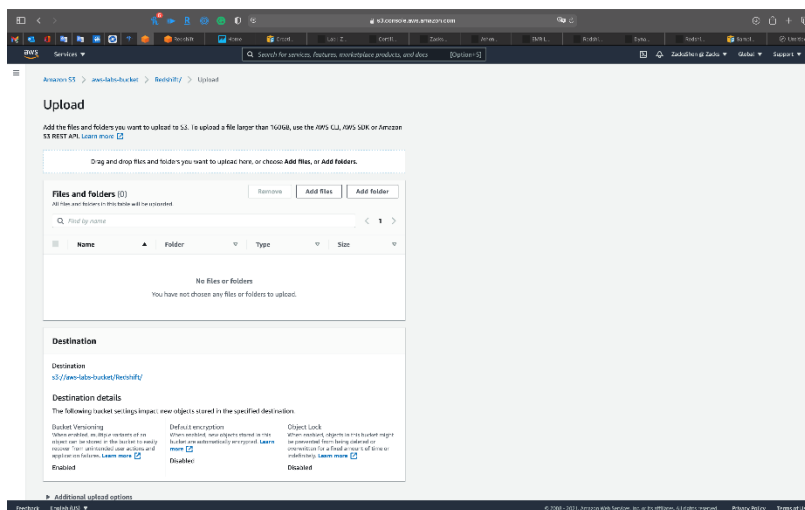
Implementación de Redshift Spectrum con QuickSight

Vamos a acceder a los datos almacenados en el depósito S3 desde QuickSight a través de Redshift Spectrum hasta el depósito S3.

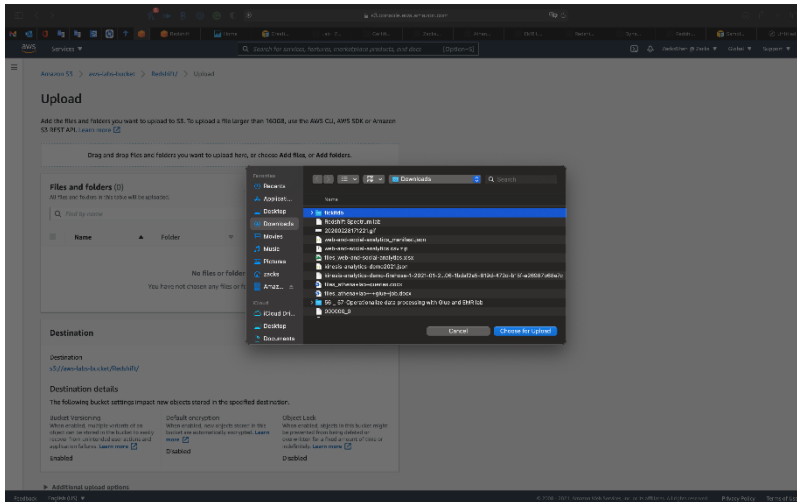
Descarga el conjunto de datos de muestra tickitdb. Carga la carpeta en su depósito S3. Haz clic en Upload



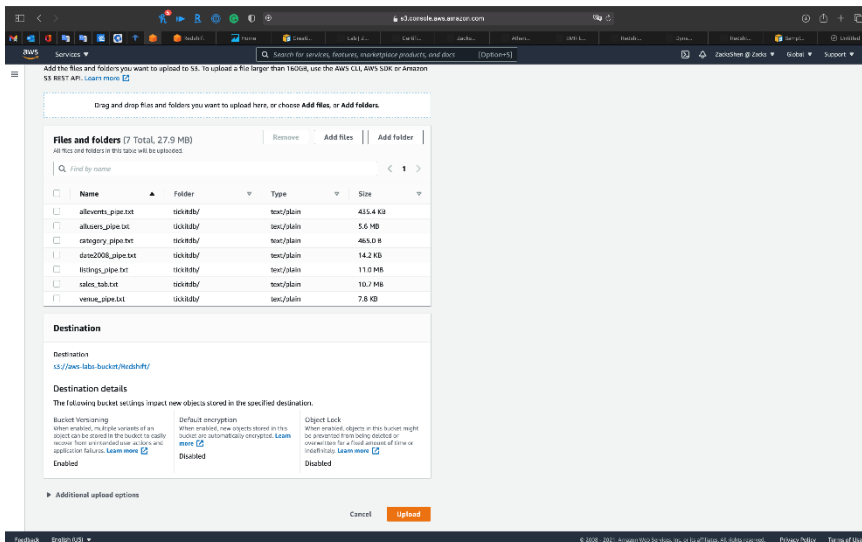
Haz clic en Add folder



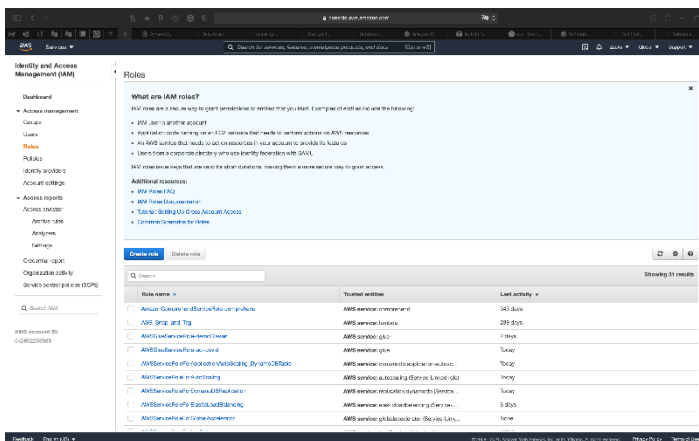
Seleccione tickitdb y luego haga clic en Choose for Upload



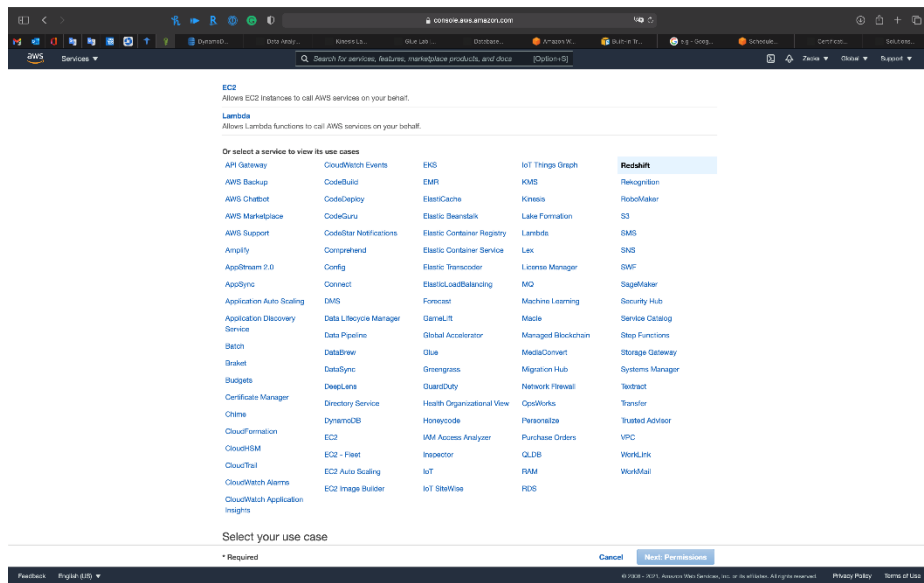
Haz clic en Upload



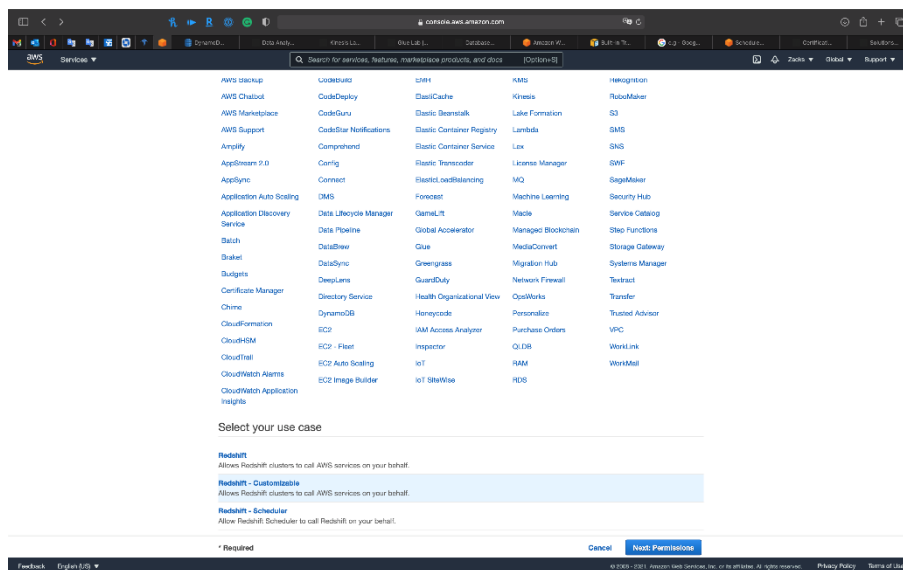
Configurar un rol de IAM para Redshift en **Servicios -> IAM -> Roles**. Haz clic en Create role



Seleccionar Redshift

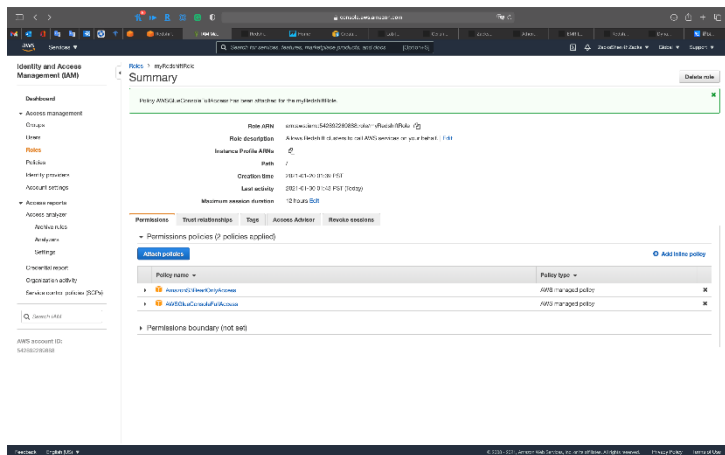


Seleccione Redshift – Customizable.

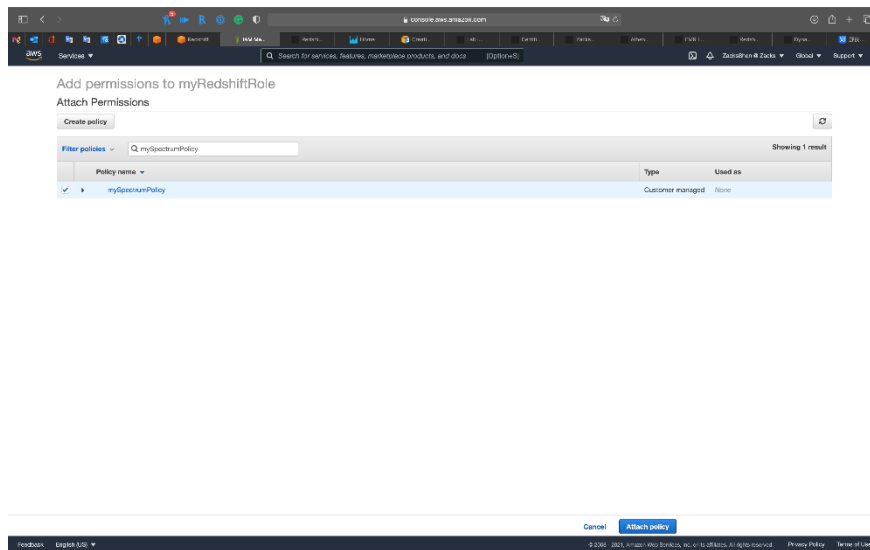


Buscar S3 y luego seleccionar AmazonS3ReadOnlyAccess. Buscar Glue y luego seleccionar AWSGlueConsoleFullAccess. Dale un nombre al rol y crealo

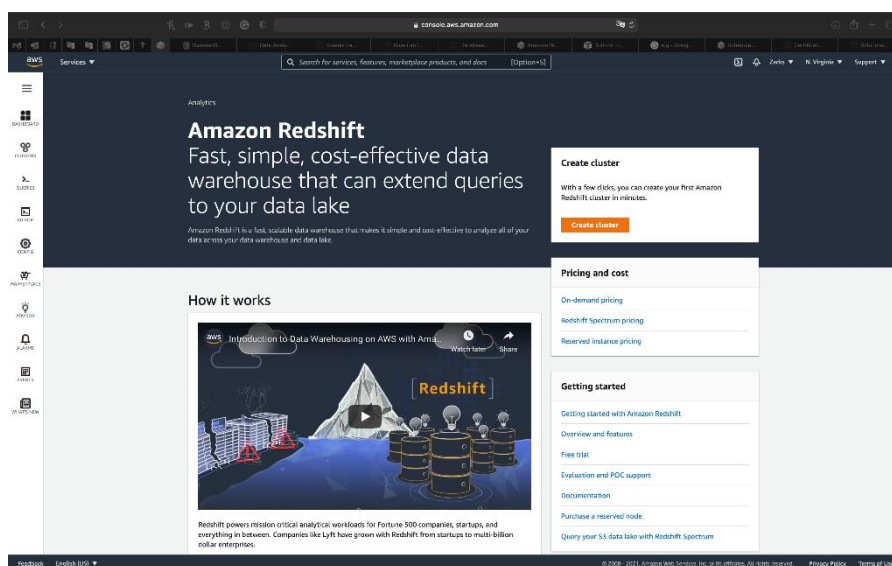
Ahora vamos a crear la política que nos falta mediante un documento JSON. Haz clic en Políticas en el panel de navegación. Haz clic en Create policy



Busca la política. Seleccionarla y luego hacer click en Attach policy

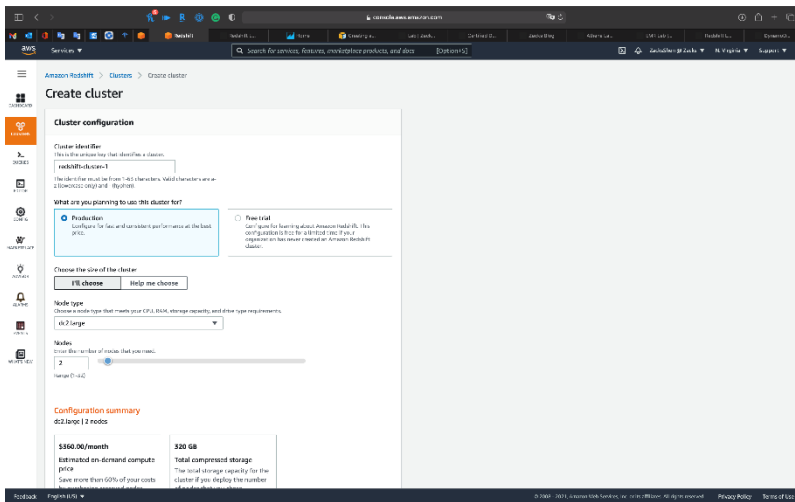


Ahora vamos a crear nuestro cluster de Redshift. Haz clic en Create cluster

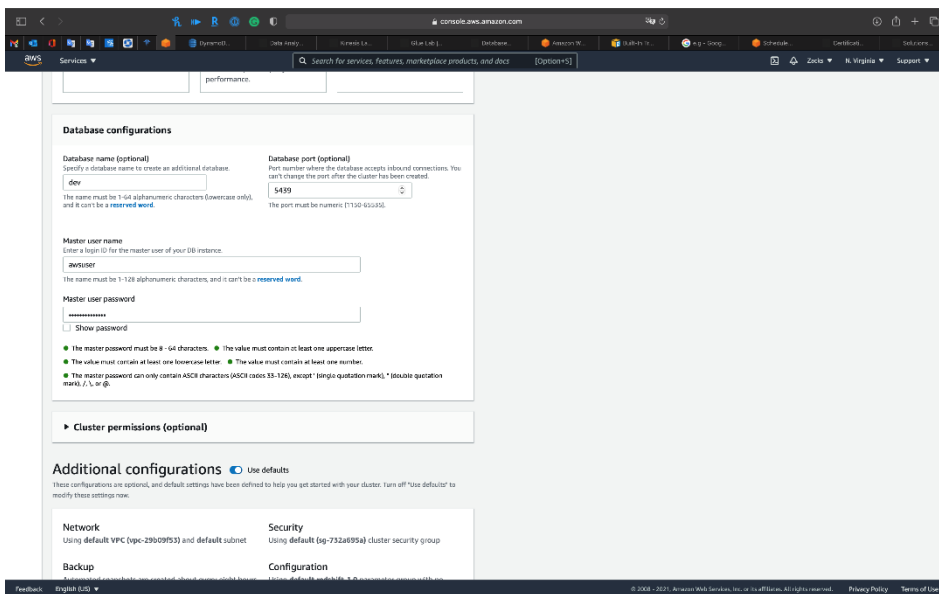


Dale un nombre al cluster

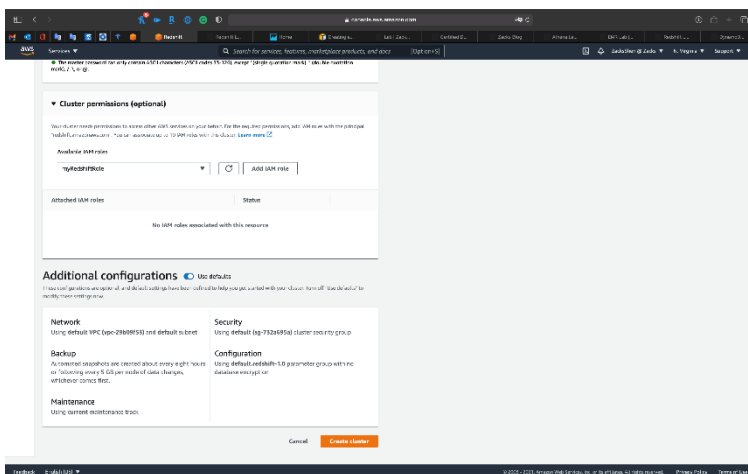
En **Elige el tamaño del clúster**, haz click en **I'll choose**
Elige el tipo de nodo: **dc2.large**. Marca que vamos a utilizar 2 nodos(1 para el nodo líder, 1 para el nodo de cálculo)



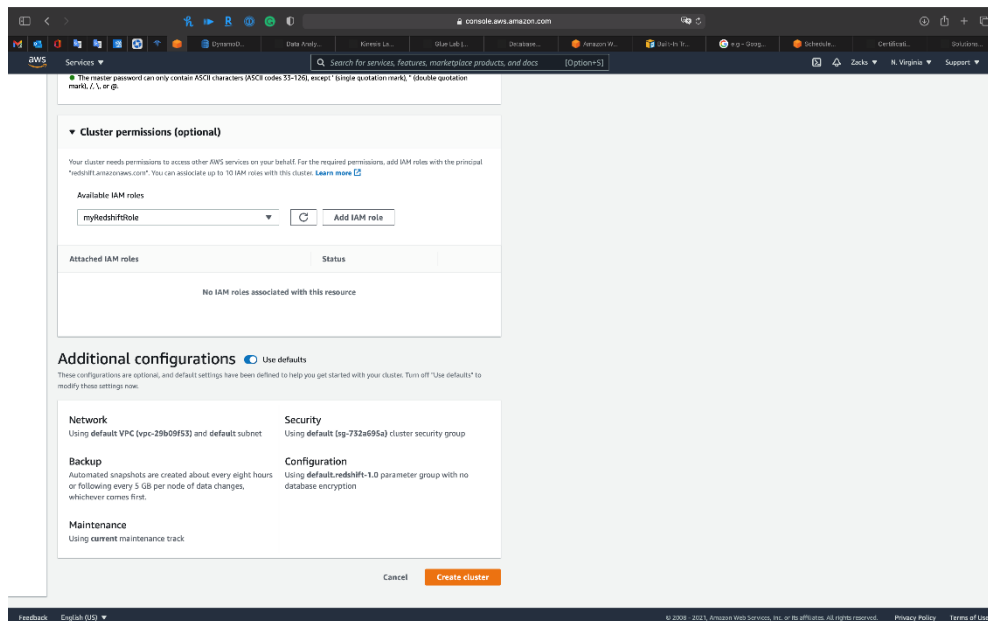
Contraseña maestra de usuario: ingresa su contraseña



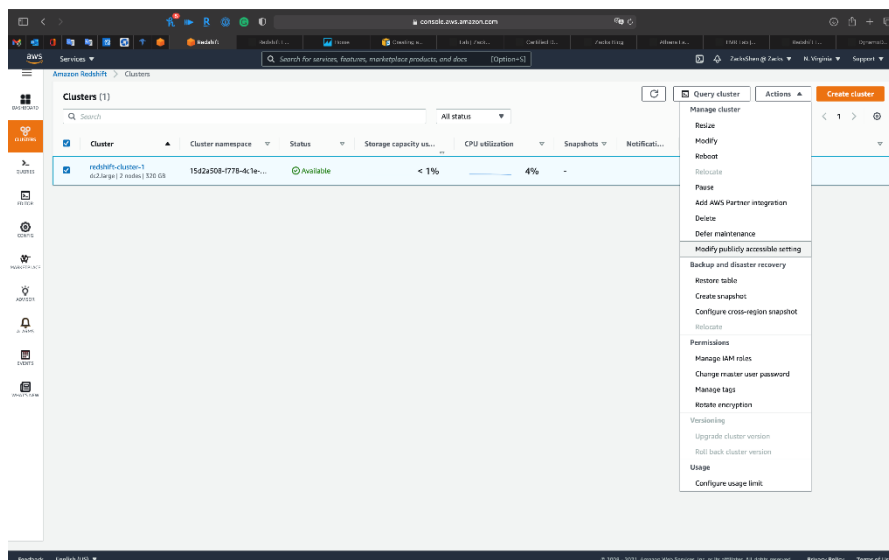
Agrega el rol de IAM en cluster permissions



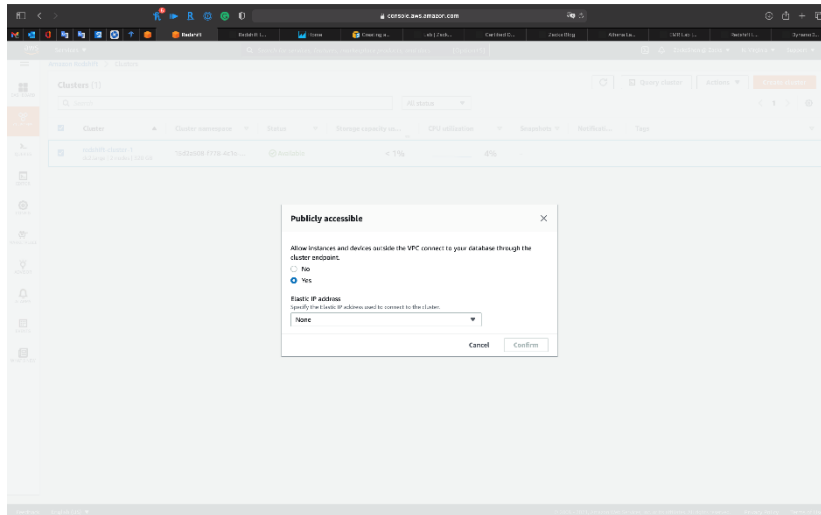
Haz clic en Create cluster



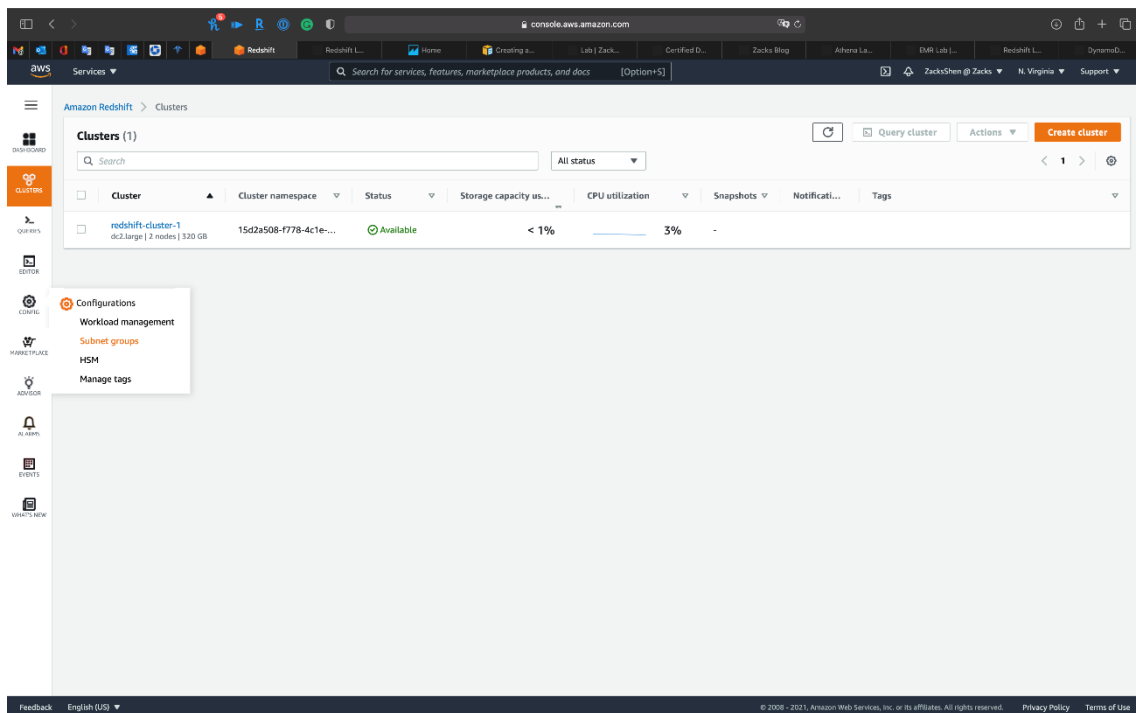
Vamos a agregar el acceso publico. Selecciona el cluster, haz clic en Modify publicly accessible setting



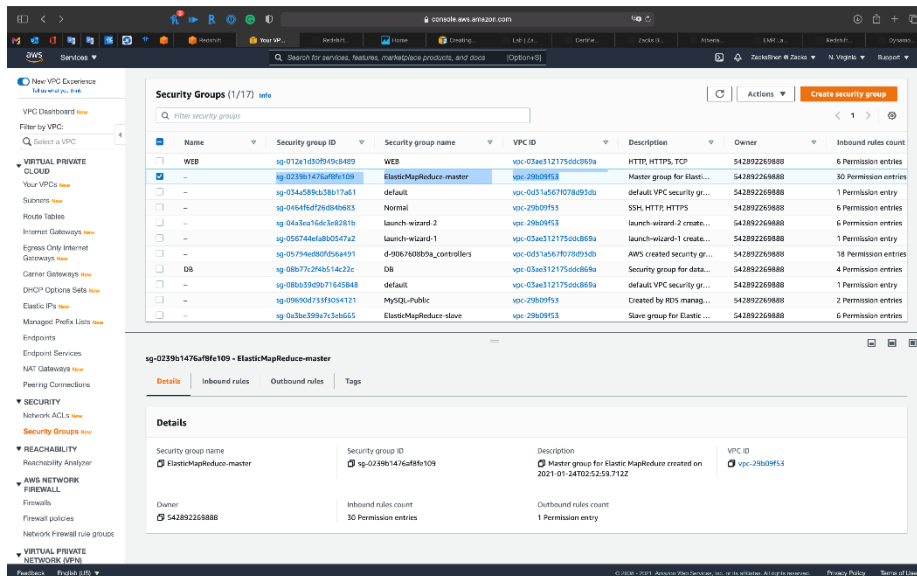
Seleccione Yes y haga clic en Confirm



Y ahora vamos a modificar el grpo de seguridad. Haz clic en CONFIG -> Subnet groups en el panel de navegación



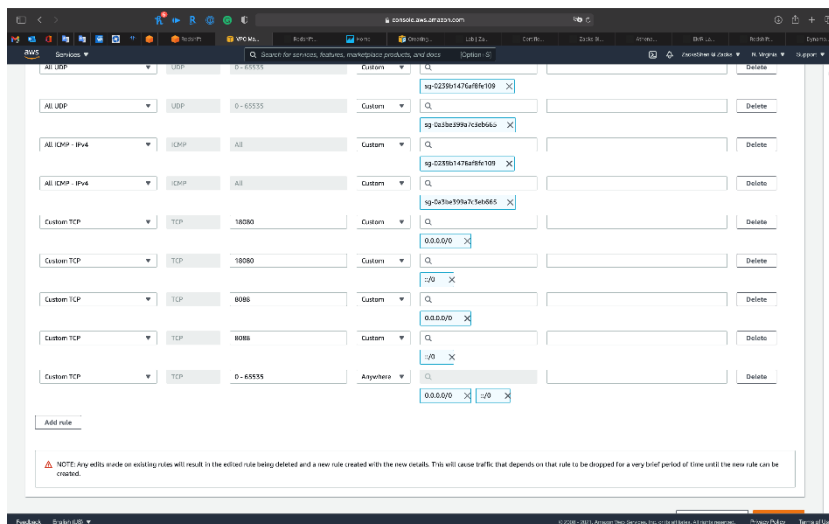
Haz clic en el ID del grupo de seguridad de la vpc de tu clúster



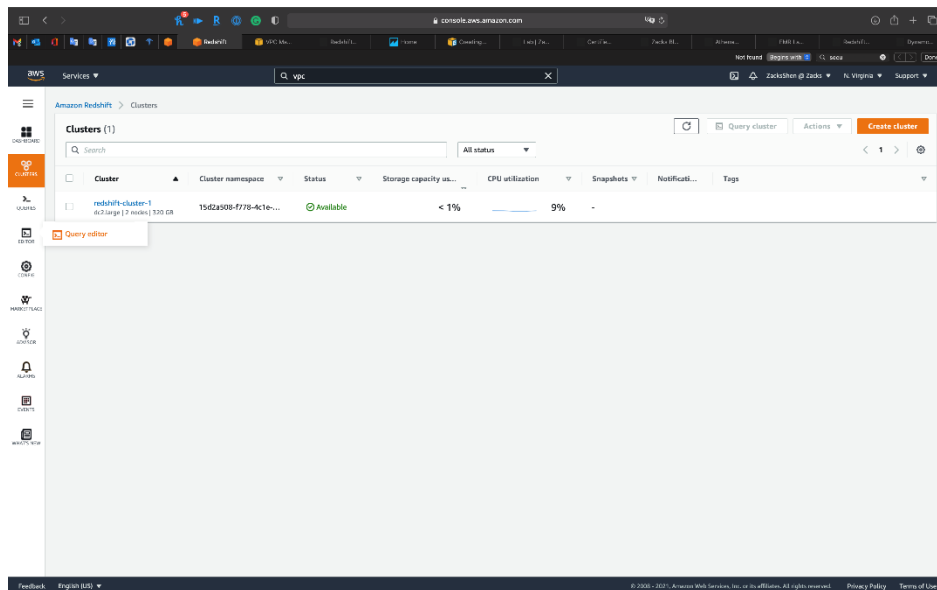
Haz clic en Edit inbound rules

- Tipo:TCP
- Rango de puertos:0 - 65535
- Fuente:Anywhere

Hacer clic Save rules

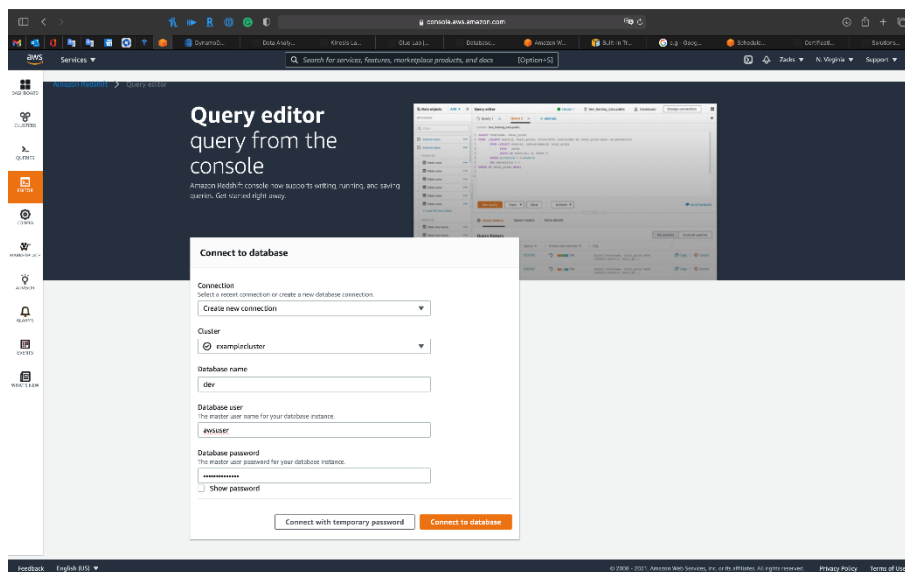


Haz clic en EDITOR -> Query Editor en el panel de navegación



- Conexión: Create new connection
- Grupo: nombre
- Nombre de la base de datos: dev
- Usuario de la base de datos: awsuser
- Contraseña de la base de datos: ingrese su contraseña

Haz clic en Connect to database



Vamos ahora a crear nuestras tablas. Copie y pegue la siguiente consulta sql en su editor.

```
create table 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 venue(  
  venueid smallint not null distkey sortkey,  
  venuename varchar(100),  
  venuecity varchar(30),  
  venuestate char(2),  
  venueseats integer);  
  
create table category(  
  catid smallint not null distkey sortkey,  
  catgroup varchar(10),  
  catname varchar(10),  
  catdesc varchar(50));  
  
create table 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 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 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);
```

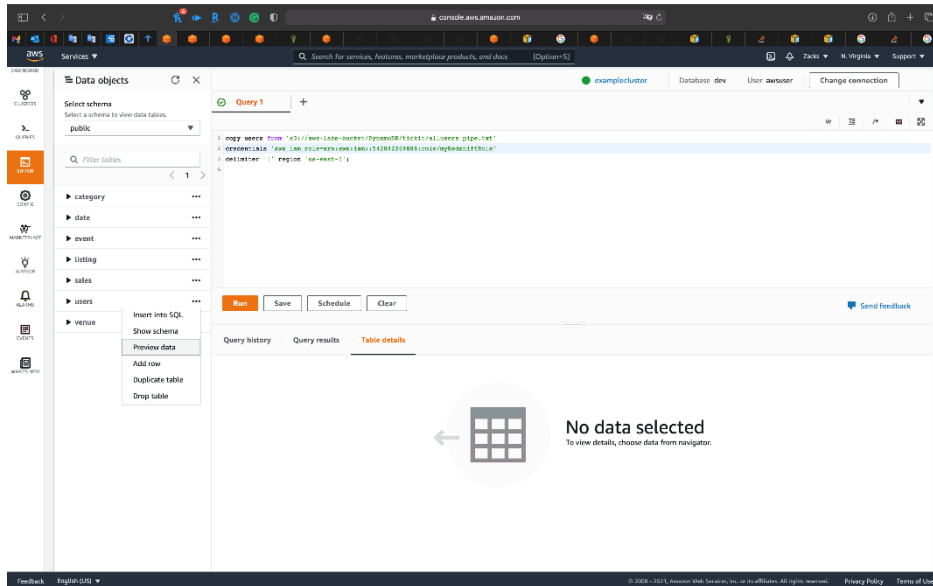
```
create table sales(  
  salesid integer not null,  
  listid integer not null distkey,  
  sellerid integer not null,  
  buyerid integer not null,  
  eventid integer not null,  
  dateid smallint not null sortkey,  
  qtysold smallint not null,  
  pricepaid decimal(8,2),  
  commission decimal(8,2),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

Haz clic en Run

Vamos a cargar **datos en tablas desde S3**. Copia y pega la siguiente consulta sql en su editor. Reemplaza el campo <> con su información.

```
copy users from 's3://<your Bucket  
name>/ticketdb/allusers_pipe.txt'  
credentials 'aws_iam_role=<your iam-role-arn>'  
delimiter '|' region '<your aws-region>';
```

Haz clic en Preview data a la derecha de users



Repita el paso anterior y reemplace los siguientes parámetros por los suyos.

copy venue from 's3://<your Bucket name>/tickitdb/venue_pipe.txt'

credentials 'aws_iam_role=<your iam-role-arn>'

delimiter '|' region '<your aws-region>';

copy category from 's3://<your Bucket name>/tickitdb/category_pipe.txt'

credentials 'aws_iam_role=<your iam-role-arn>'

delimiter '|' region '<your aws-region>';

copy date from 's3://<your Bucket name>/tickitdb/date2008_pipe.txt'

credentials 'aws_iam_role=<your iam-role-arn>'

delimiter '|' region '<your aws-region>';

copy event from 's3://<your Bucket name>/tickitdb/allevvents_pipe.txt'

credentials 'aws_iam_role=<your iam-role-arn>'

delimiter '|' timeformat 'YYYY-MM-DD HH:MI:SS' region '<your aws-region>';

copy listing from 's3://<your Bucket name>/tickitdb/listings_pipe.txt'

credentials 'aws_iam_role=<your iam-role-arn>'

delimiter '|' region '<your aws-region>';

copy sales from 's3://<your Bucket name>/ticketdb/sales_tab.txt'

credentials 'aws_iam_role=<your iam-role-arn>'

delimiter '\t' timeformat 'MM/DD/YYYY HH:MI:SS' region '<your aws-region>';

En este paso, creamos una base de datos y una tabla de Redshift Spectrum SIN datos de S3, lo que significa que consultamos los datos directamente de S3 a través del esquema de Redshift Spectrum.

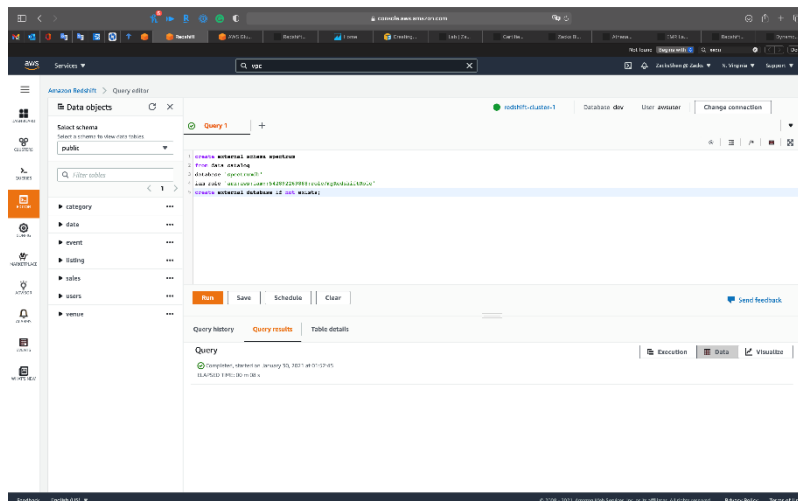
create external schema spectrum

from data catalog

database 'spectrumdb'

iam_role 'arn:aws:iam::542892269888:role/myRedshiftRole'

create external database if not exists;



create external table spectrum.sales(

salesid integer,

listid integer,

sellerid integer,

buyerid integer,

eventid integer,

dateid smallint,

qtysold smallint,

pricepaid decimal(8,2),

commission decimal(8,2),

saletime timestamp)

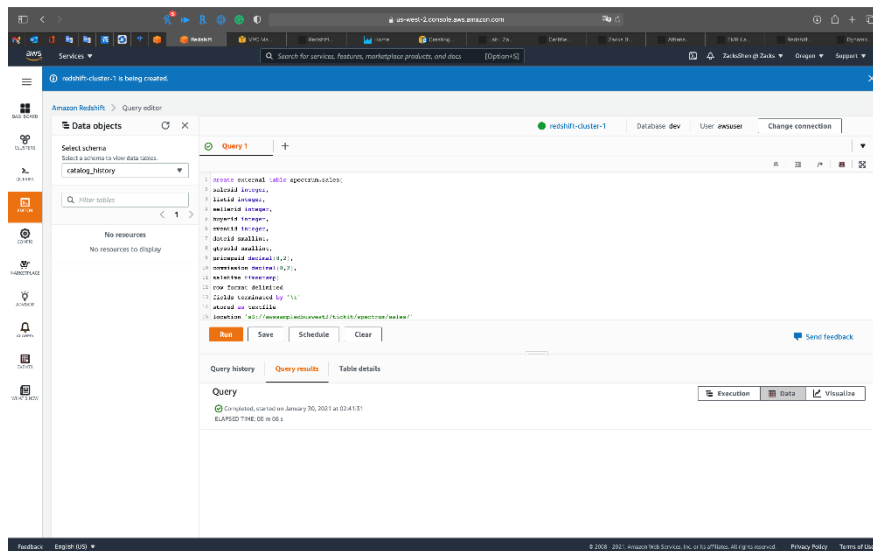
row format delimited

fields terminated by '\t'

stored as textfile

location 's3://awssampleduswest2/tickit/spectrum/sales/'

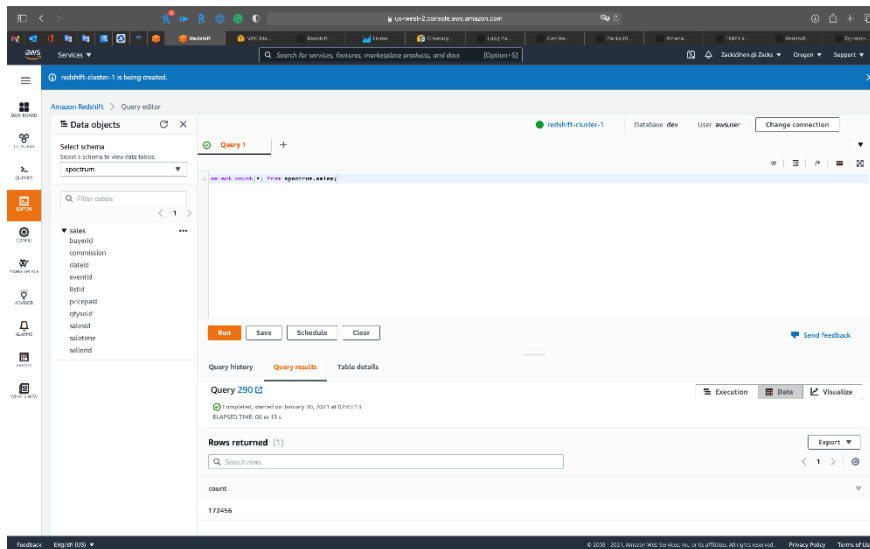
table properties ('numRows'='172000');



Consulta tus datos en Amazon S3

Obtenga el número de filas en la tabla SPECTRUM.SALES.

`select count(*) from spectrum.sales;`

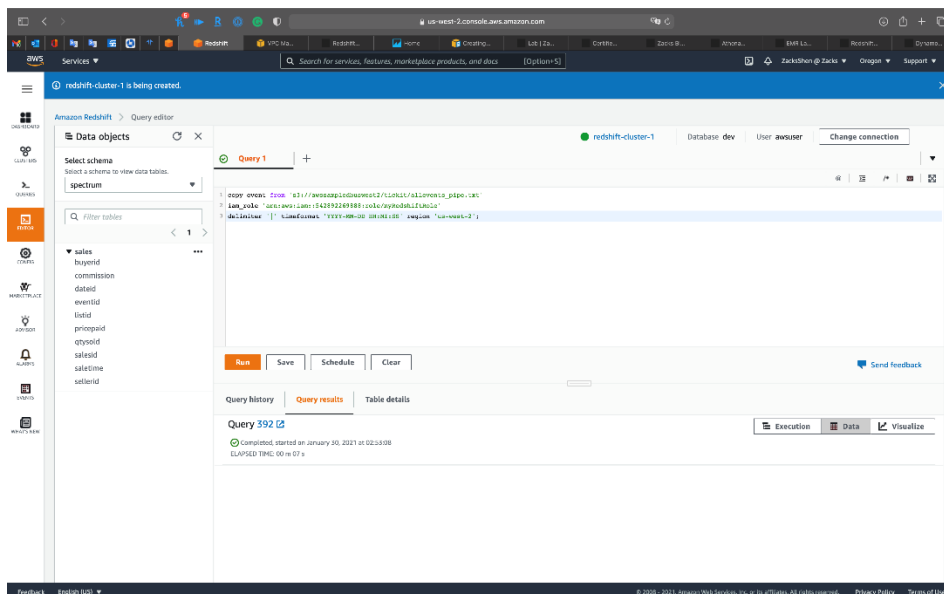


Cargue la tabla EVENT reemplazando el ARN del rol de IAM en el siguiente comando COPY con el ARN del rol que creó.

copy event from 's3://awssampleduswest2/ticket/allevnts_pipe.txt'

iam_role 'arn:aws:iam::542892269888:role/myRedshiftRole'

delimiter '|' timeformat 'YYYY-MM-DD HH:MI:SS' region 'us-west-2';



El siguiente ejemplo une la tabla externa SPECTRUM.SALES con la tabla local EVENT para encontrar las ventas totales de los 10 eventos principales.

select top 10 spectrum.sales.eventid, sum(spectrum.sales.pricepaid) from spectrum.sales, event

where spectrum.sales.eventid = event.eventid

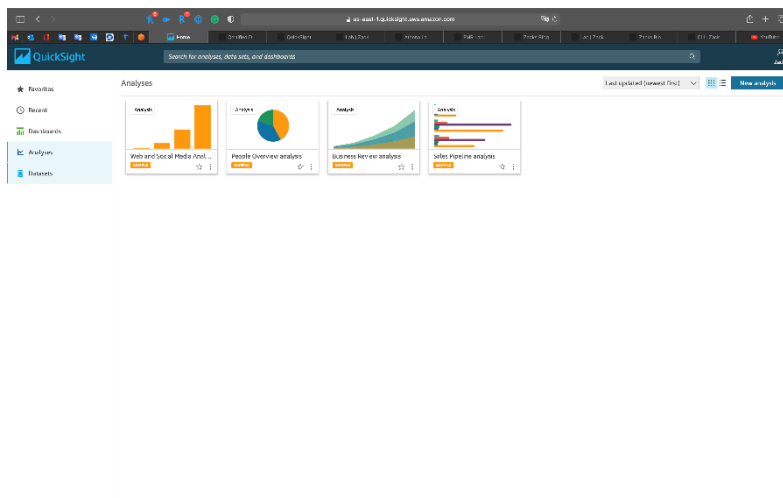
and spectrum.sales.pricepaid > 30

group by spectrum.sales.eventid

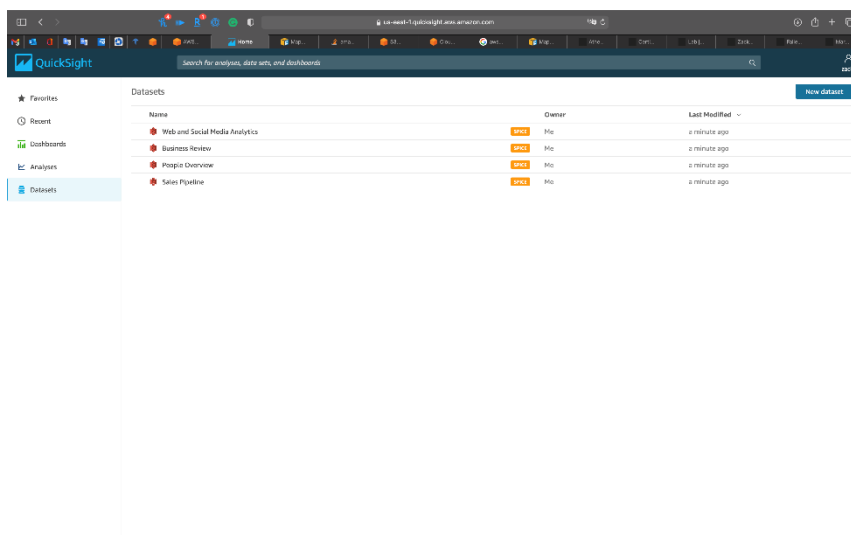
order by 2 desc;

Vamos ahora a configurar nuestro Quicksight

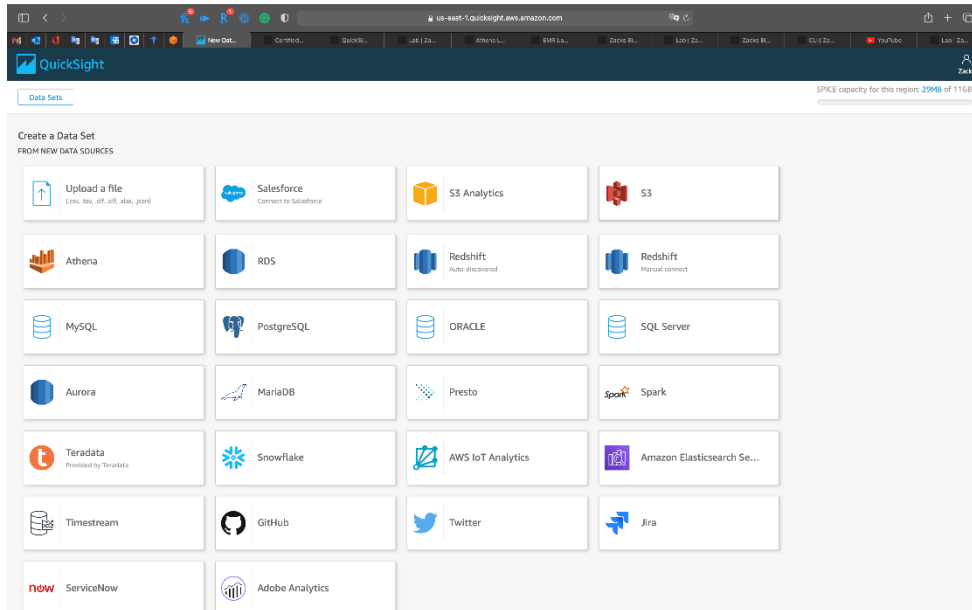
Haz clic en Datasets en el panel de navegación dentro de QuickSight



Haga clic en New dataset en la página derecha y superior

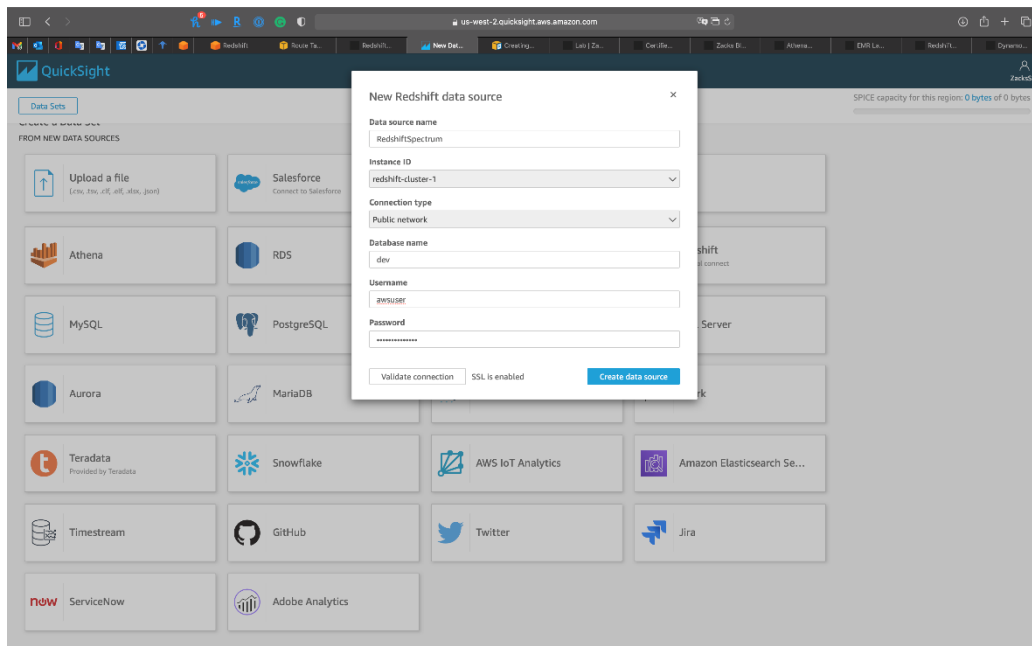


Haz clic en Redshift para crear una nueva conexion

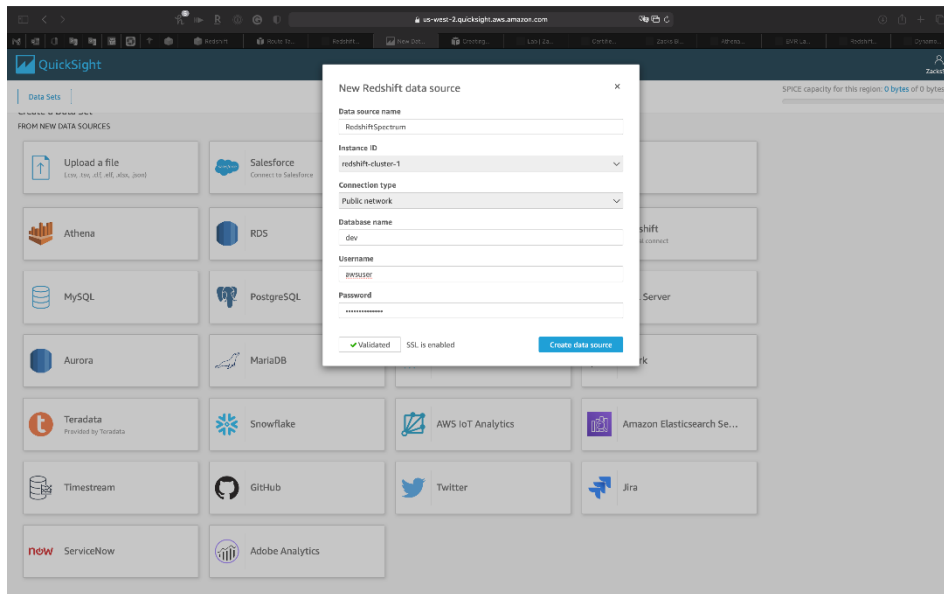


- Nombre de fuente de datos: nombre
- ID de instancia: nombre
- Tipo de conexión: Public network
- Nombre de la base de datos: dev
- Nombre de usuario: awsuser
- Contraseña

Haz clic en Validate connection

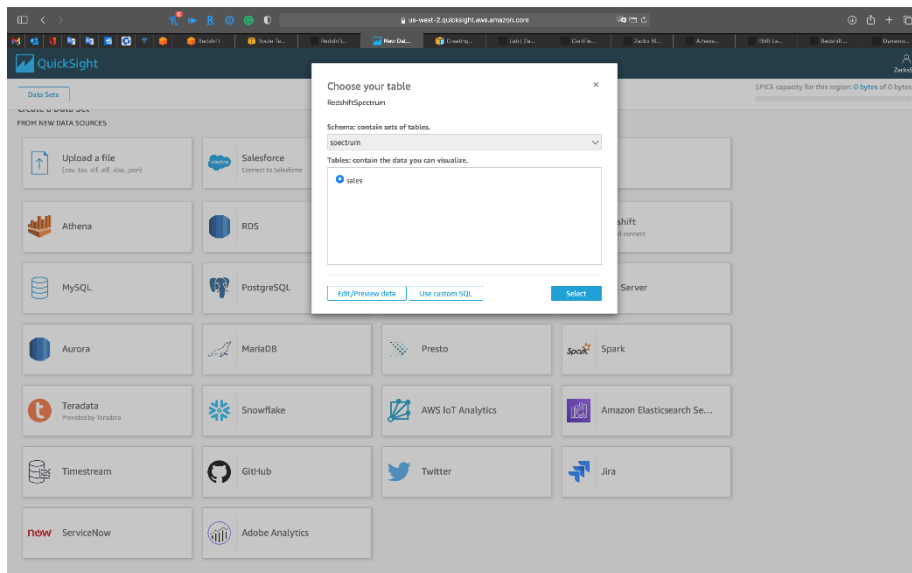


Haga clic en Create data source

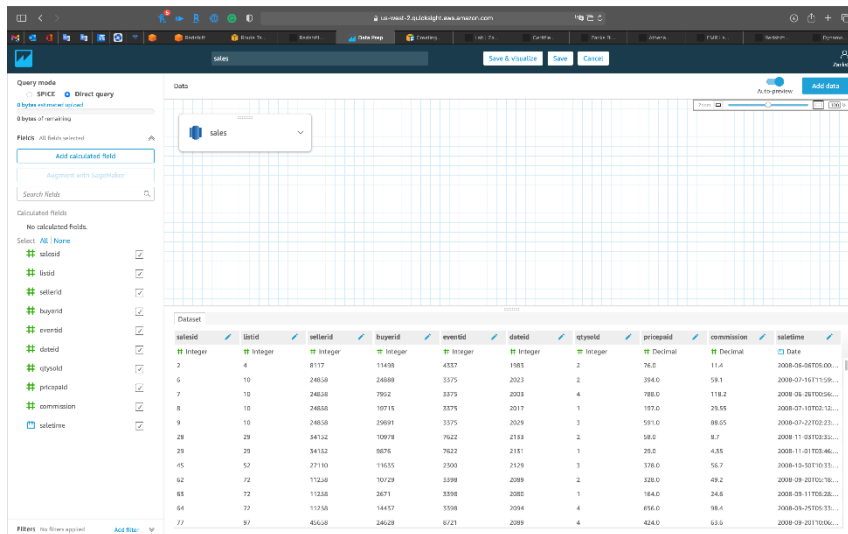


- Esquema:spectrum
- Tabla:sales

Haga clic en Edit/Preview data

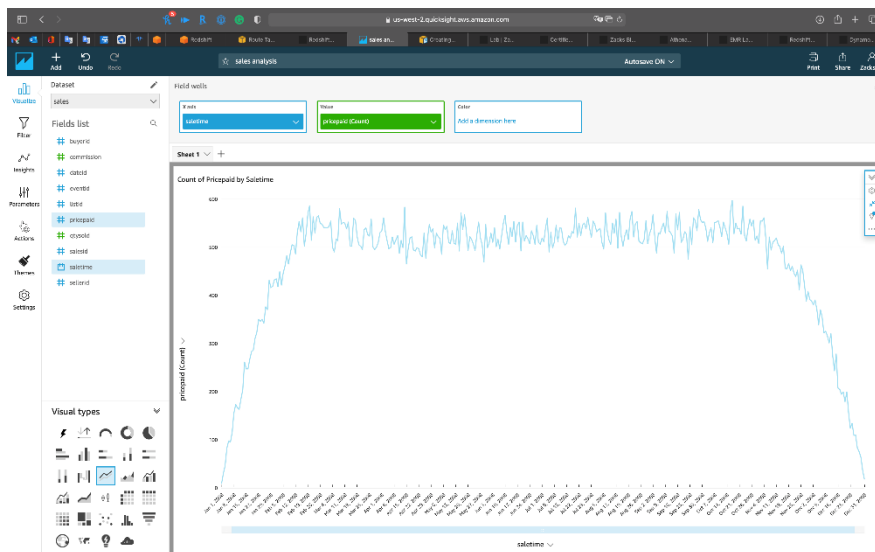


Haz clic en Save & visualize en la página superior y central



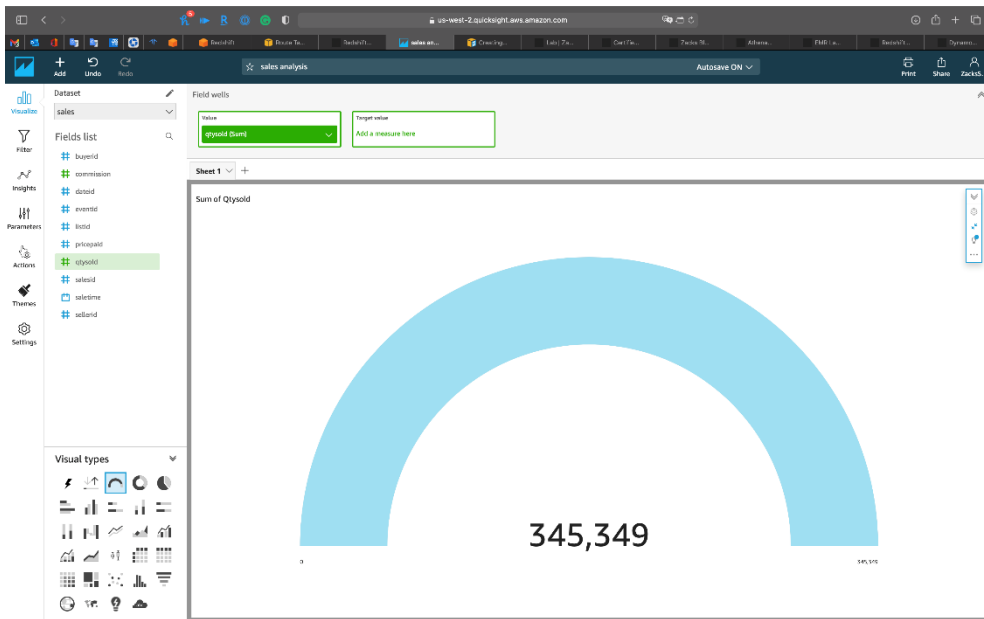
Vamos a crear ahora un grafico lineal

- Tipo visual :Line chart
- eje X:salestime
- Valor:pricepaid (count)



Tambien crearemos una table de indicadores

- Tipos visuales: Gauge chart
- Valor:qtysold (count)



Y por ultimo un gráfico de anillos

- Tipos visuales: Donut chart
- Grupo/Color: eventid
- Valor: qtysold (Average)

