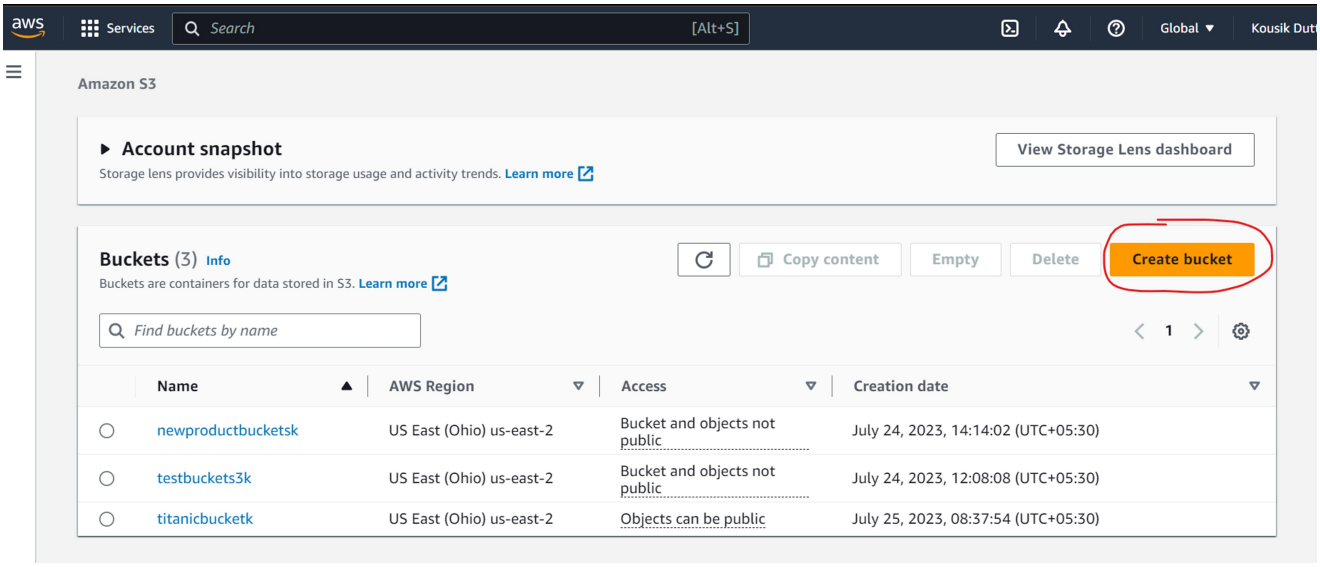
Snowflake External Stage

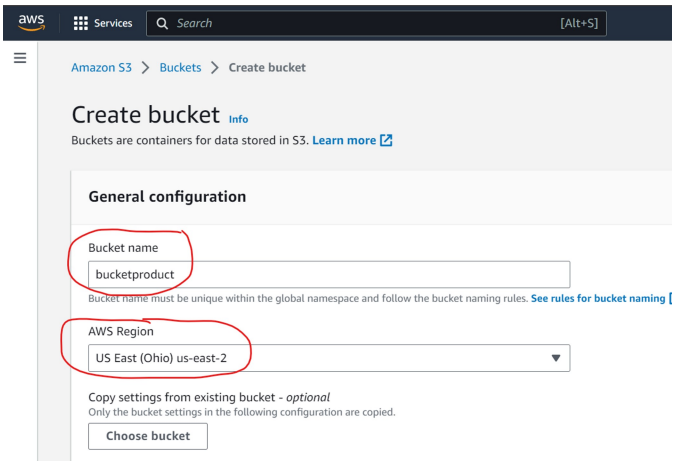
Tuesday, July 25, 2023 9:10 PM

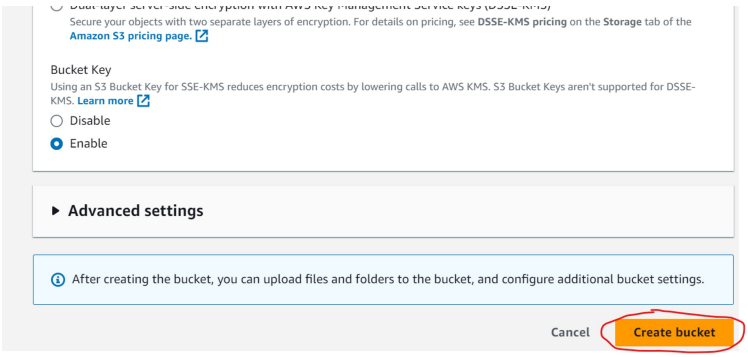
External Stage Creation

Step 1 : Open AWS account and Login to that.

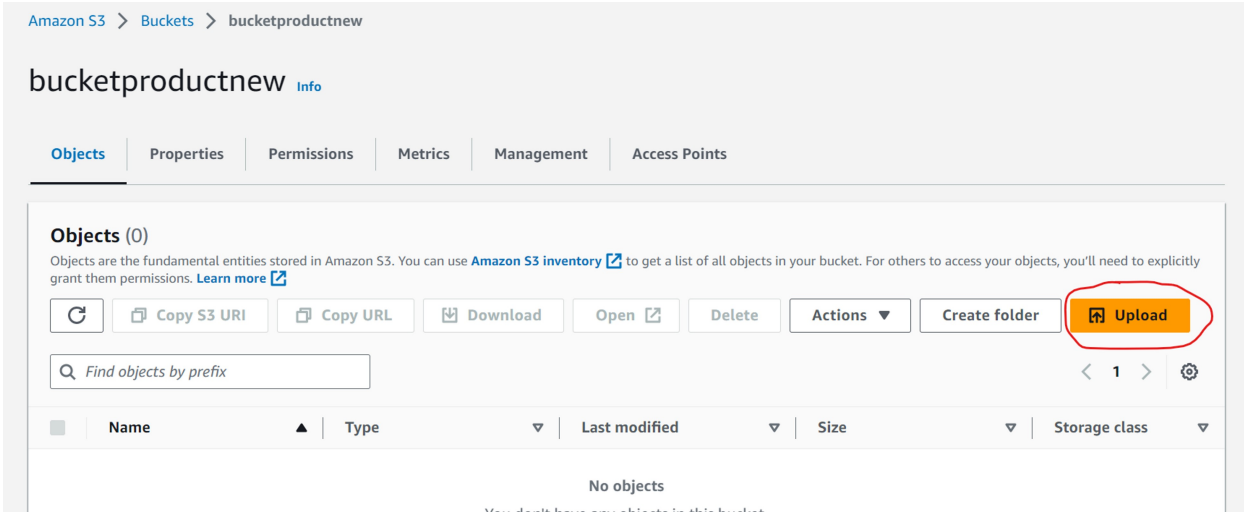
Step 2: First create a bucket :

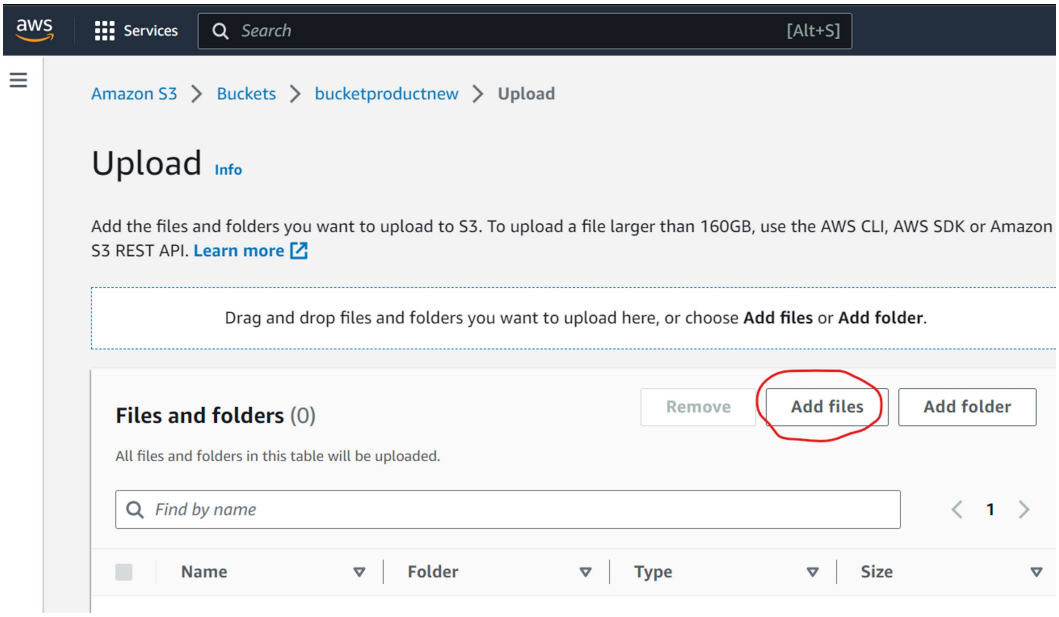


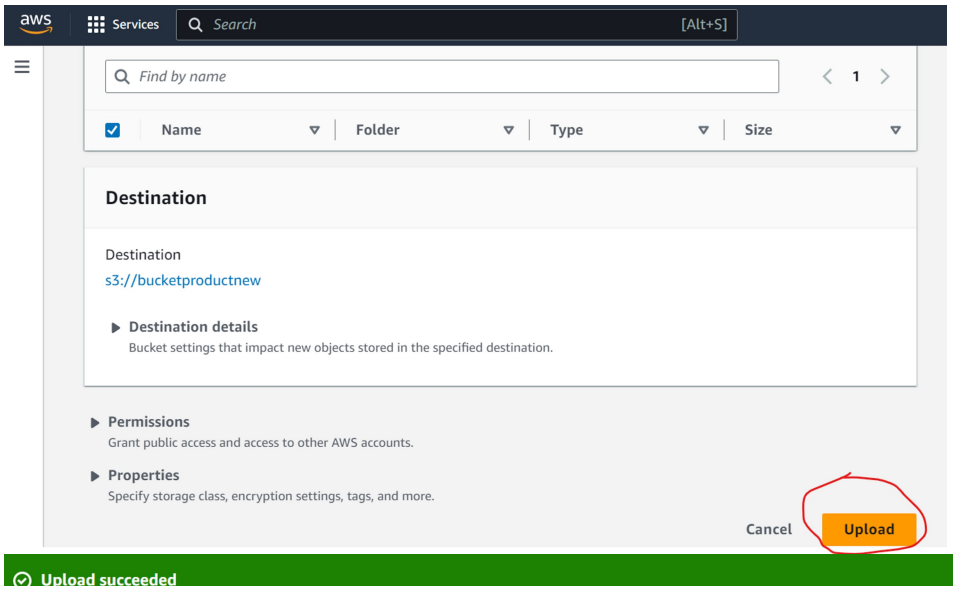


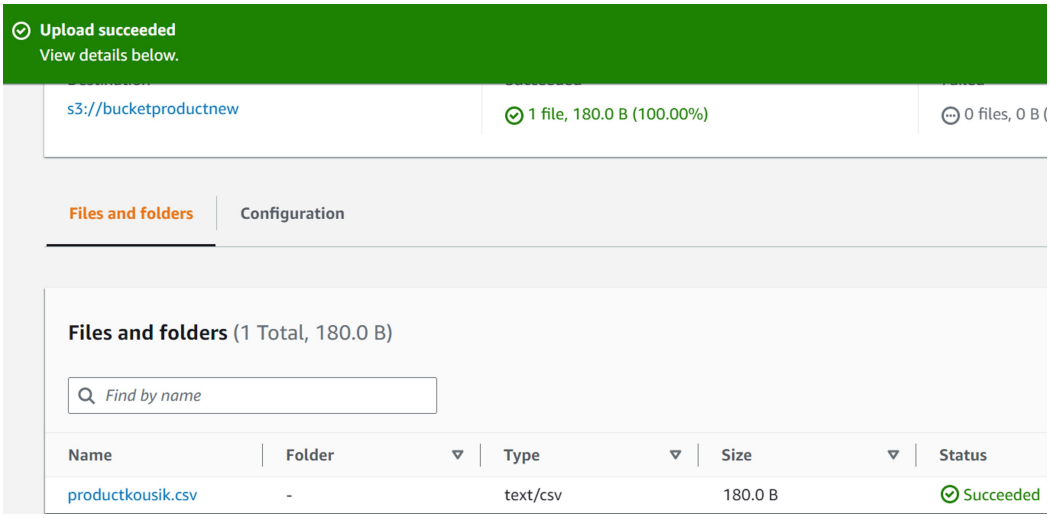


Step 3: Upload data into the bucket :

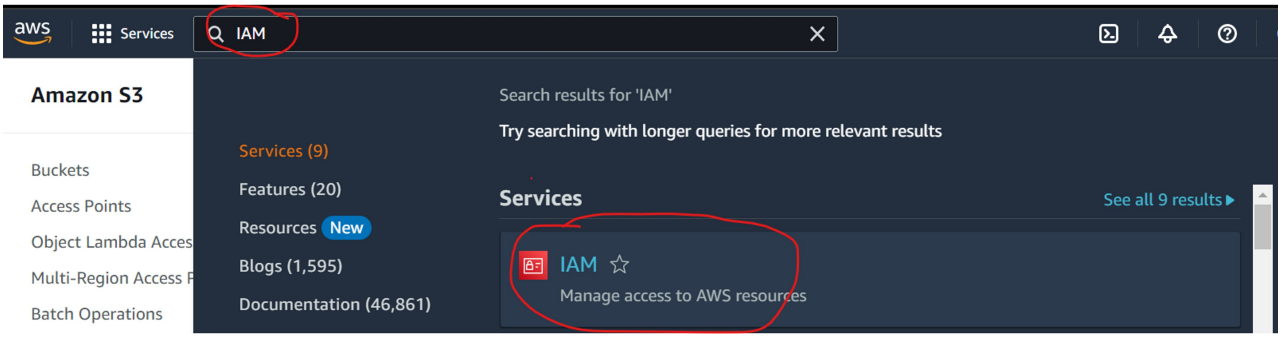




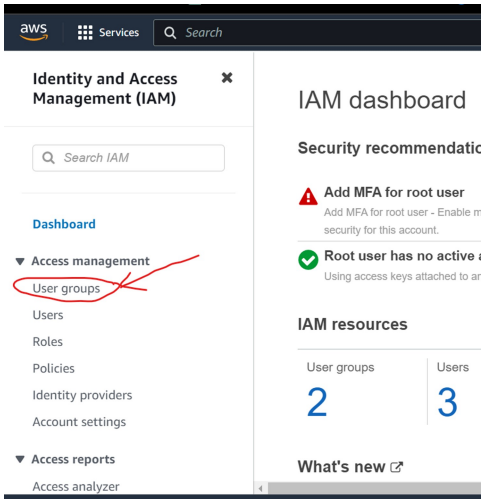




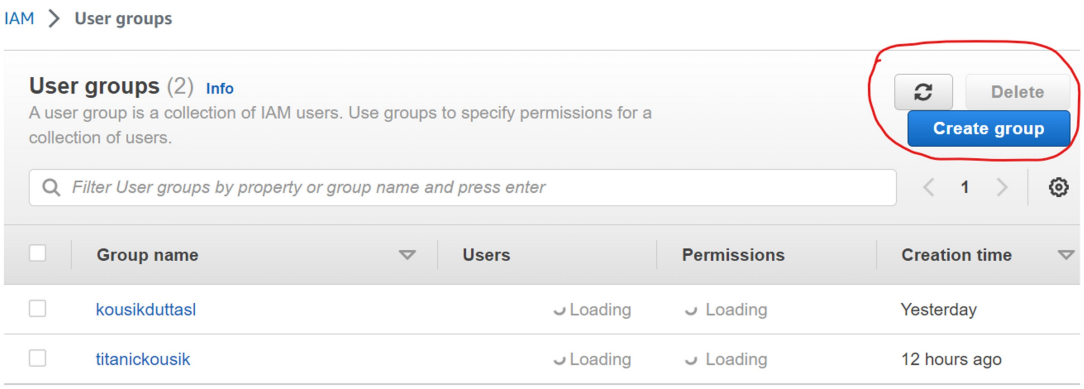
Step 4: Go to IAM :



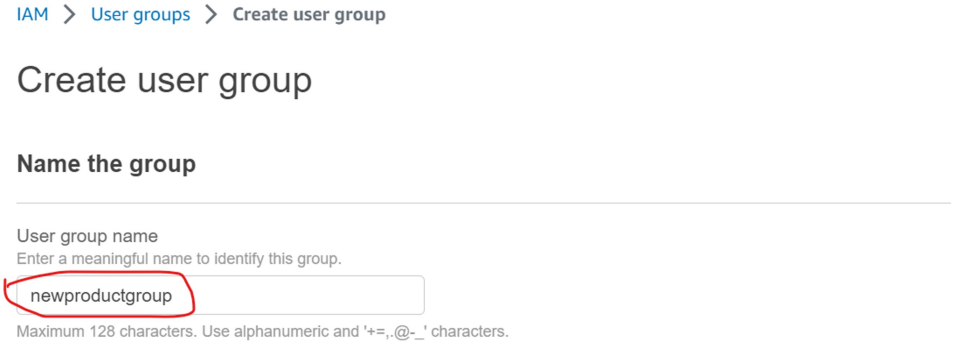
Step 5: Go to User groups:



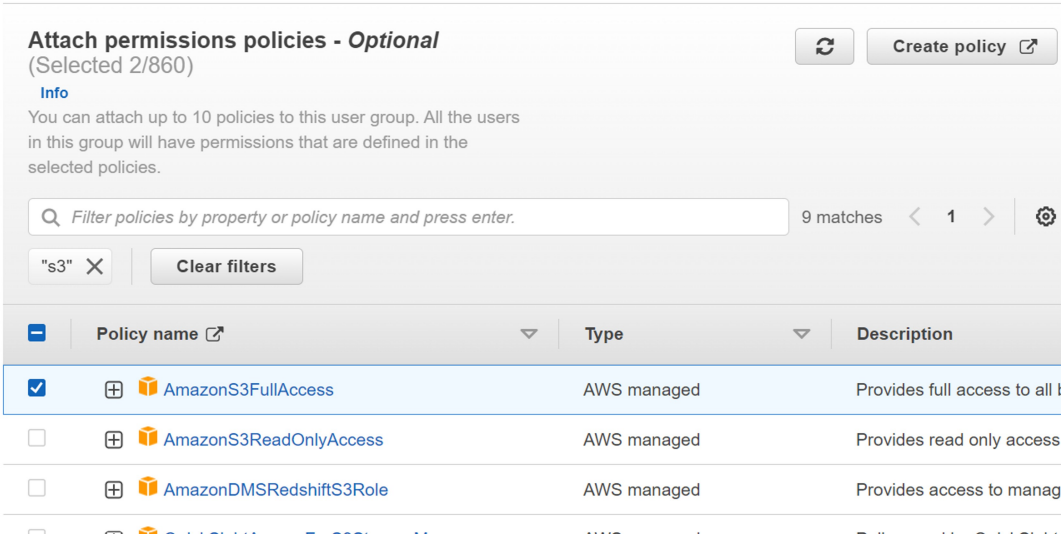
Step 6: Create User group:



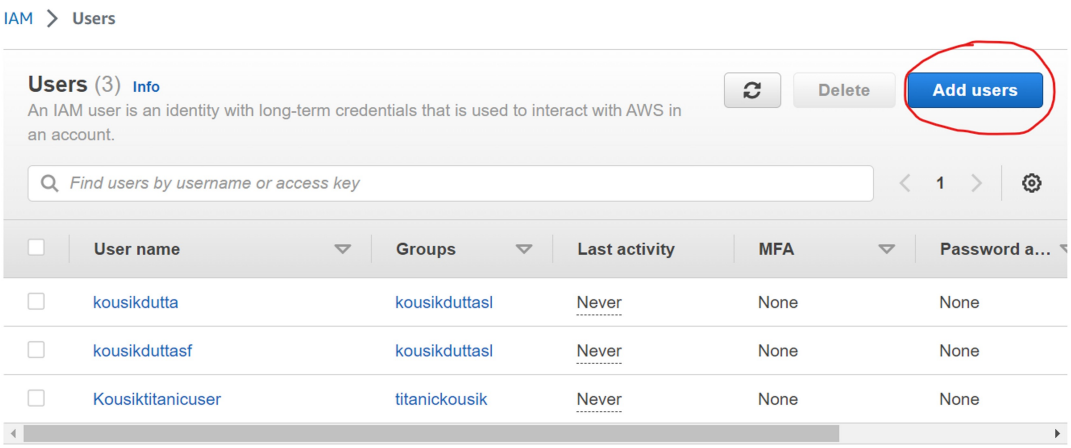
Provide a name to the group :



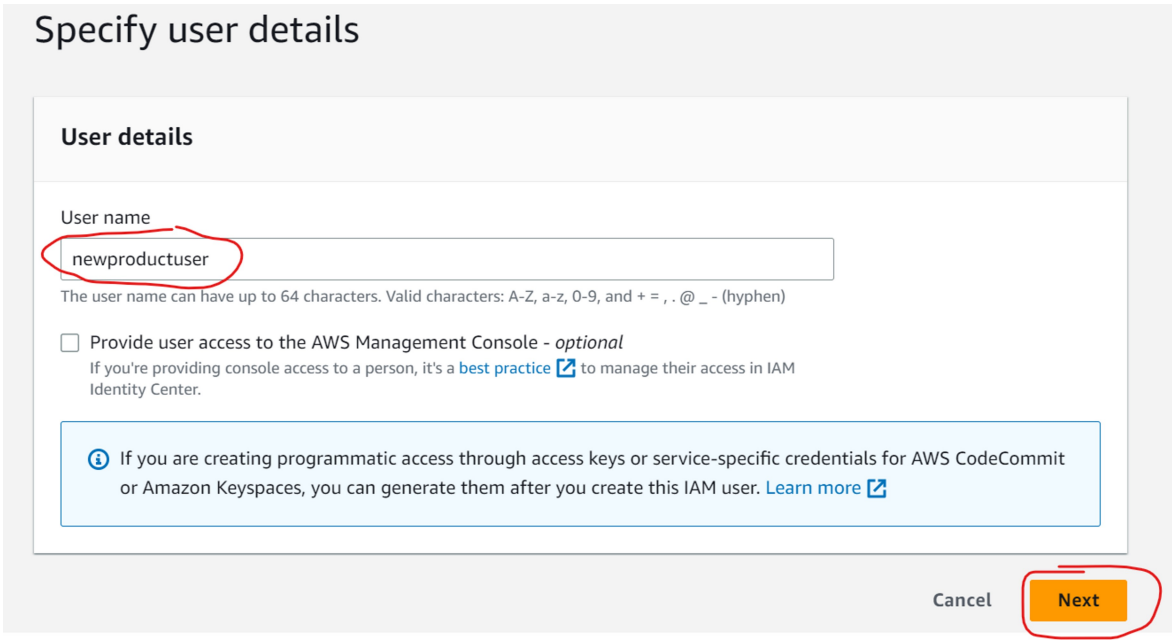
Give permission to "S3 full access" and "administrator access" :



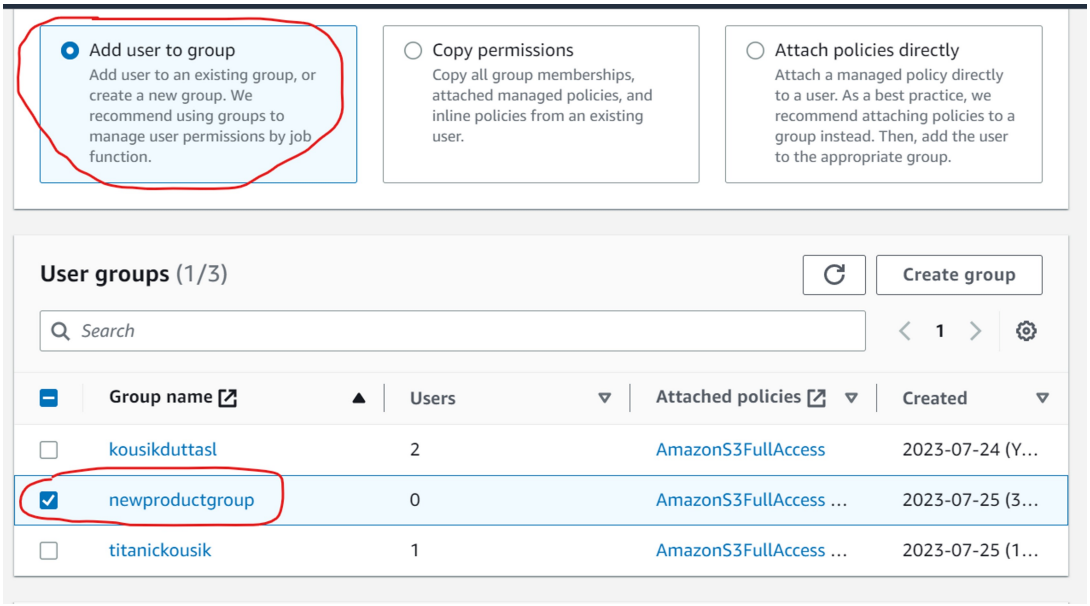
Step 6: Create User :



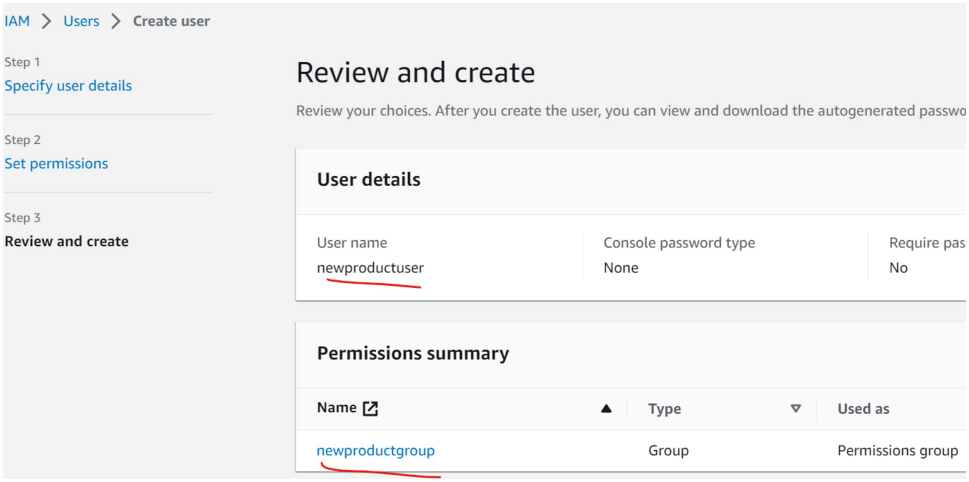
Give user name :



Map the particular user to the group :

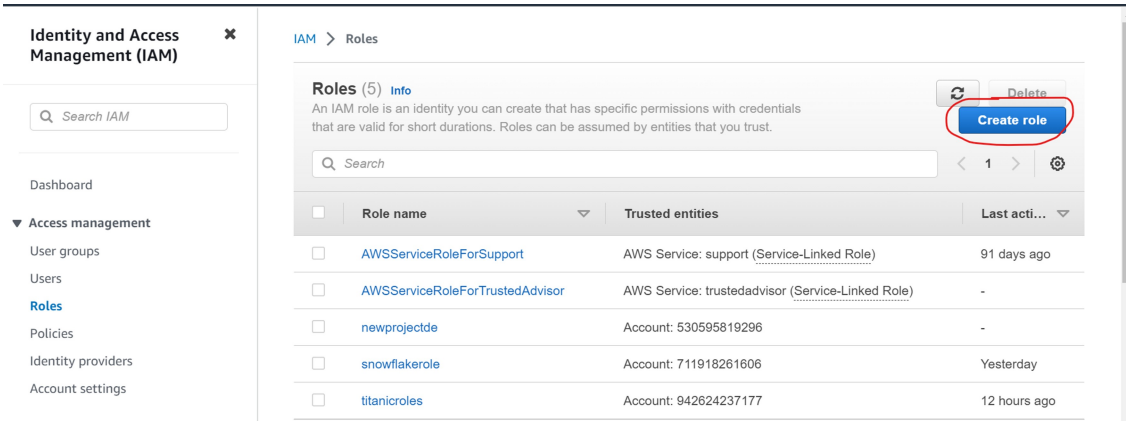


Next we will be able to see something like below:

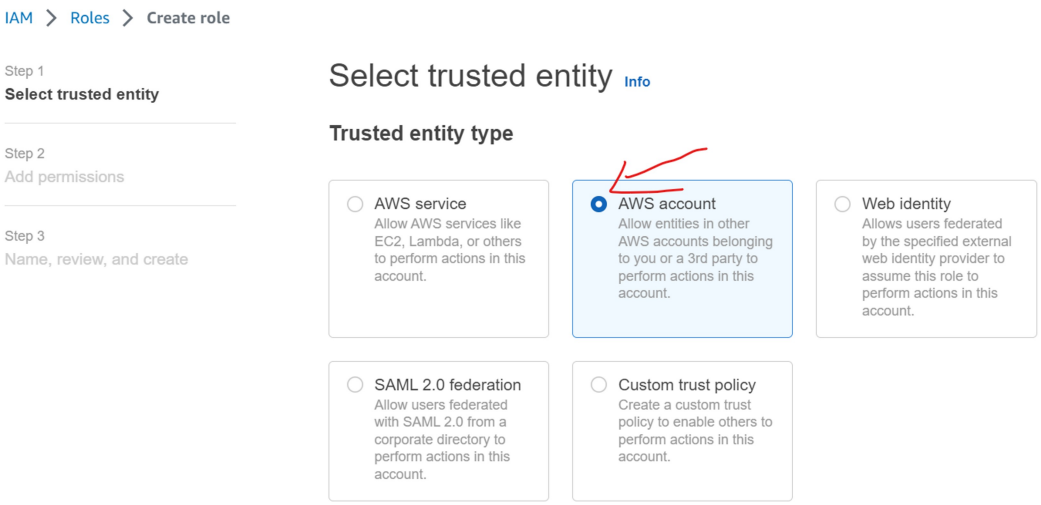


Then click on create User.

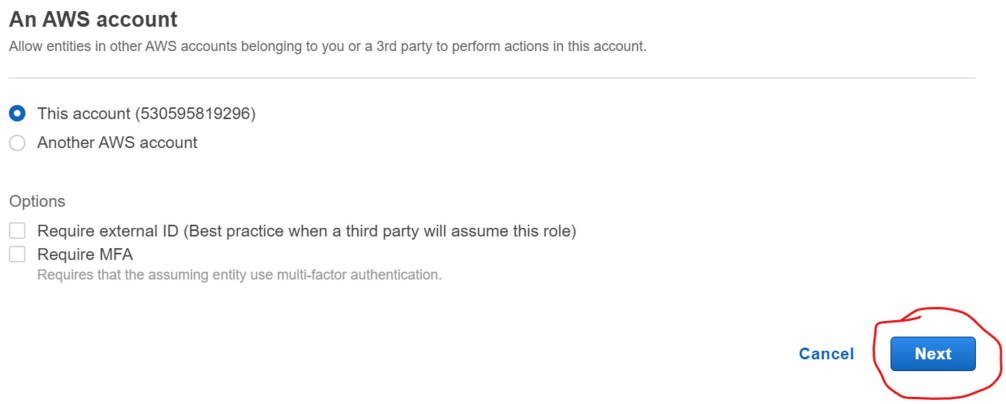
Step 7: Create Role :



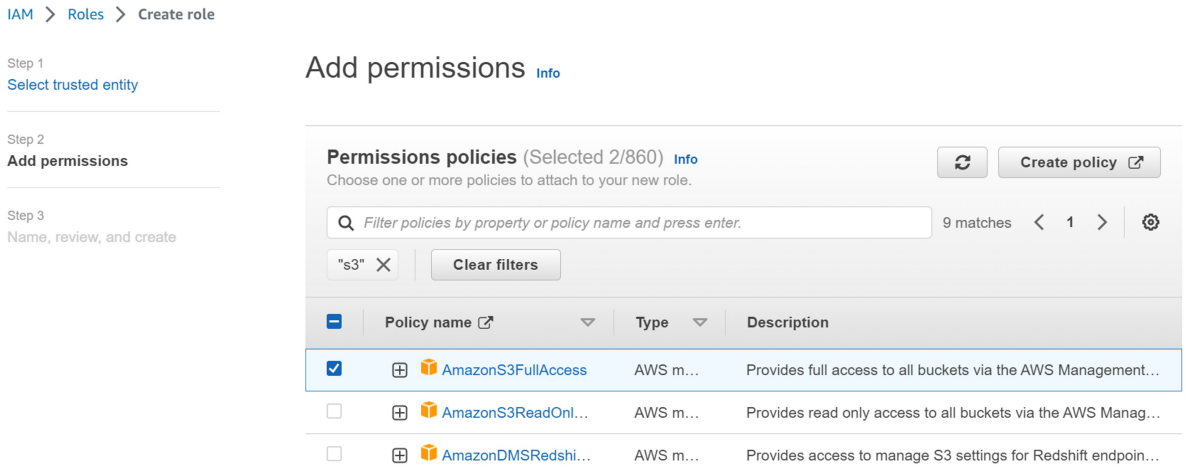
Click on AWS Account :



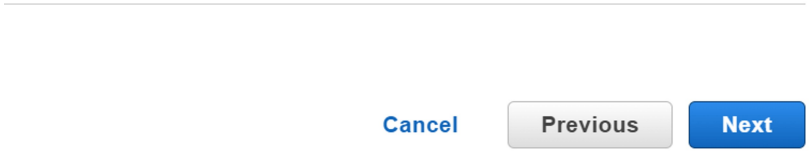
Click on next :



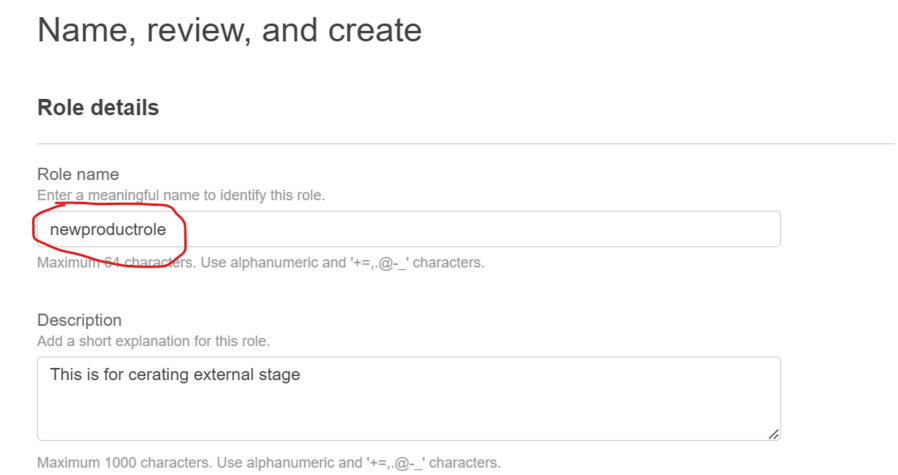
Give permission to "S3 full access" and "administrator access" :



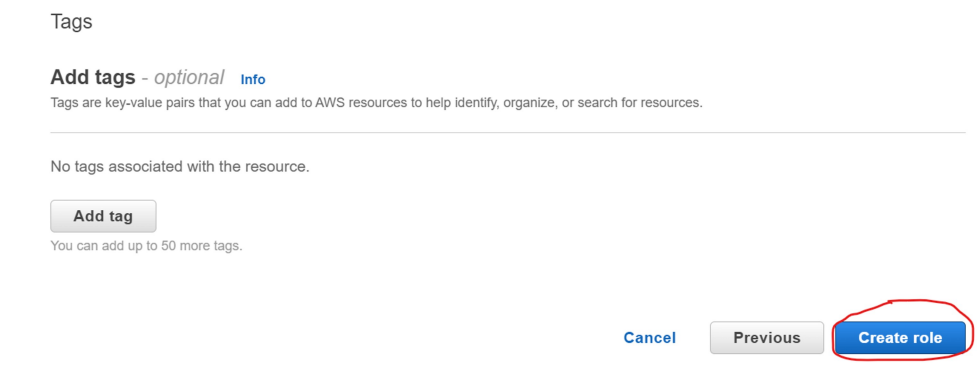
Click on next :



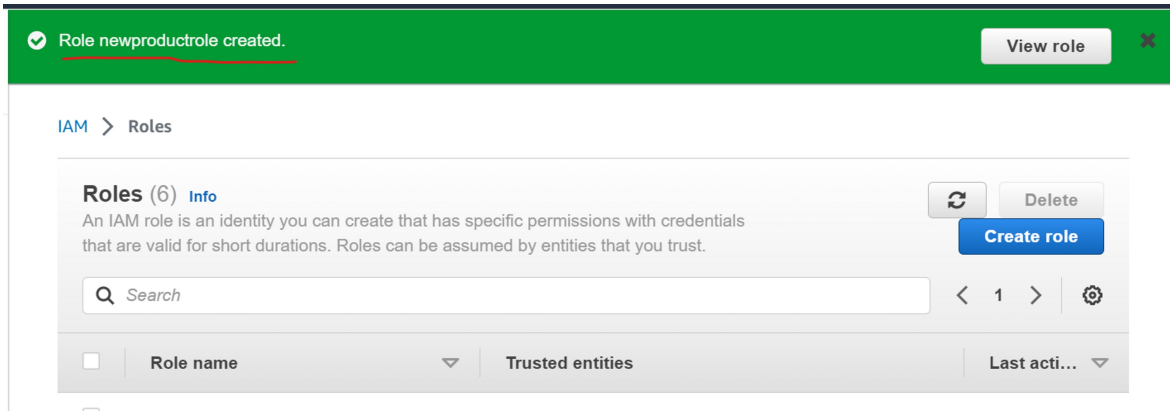
Give the name of the role :



Click on create role :

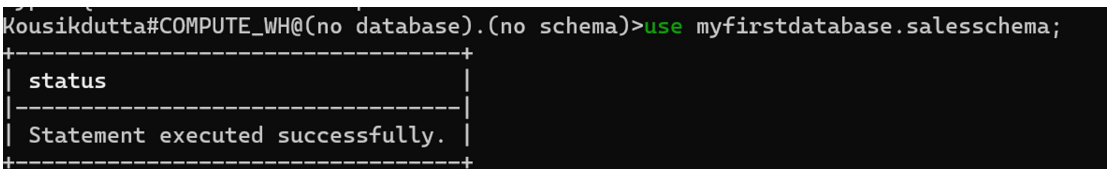


It will show like below :



Step 7: Next go to Snowflake CLI and execute below commands:

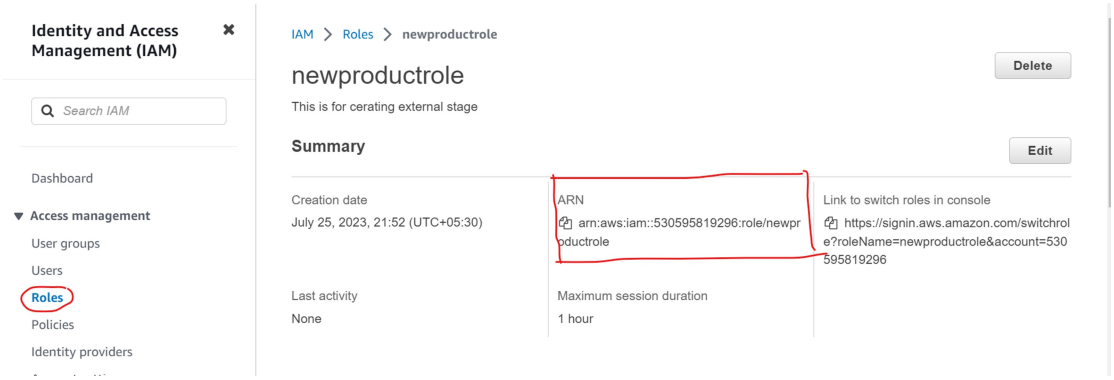
Select a particular database and schema



Create CSV Format :

```
create or replace file format MYFIRSTDATABASE.SALESSCHEMA.my_csv_format
        type = csv
        field_delimiter = ','
        skip_header = 1
        null_if = ('NULL', 'null')
        empty_field_as_null = true;
```

Next go to the role and copy the ARN :



Then , copy the bucket name which you have created and ARN and write the below code and execute:

```
Kousikdutta#COMPUTE_WH@MYFIRSTDATABASE.SALESSCHEMA>create or replace storage integration s3_int2
                                                   type = external_stage
                                                   storage_provider = s3
                                                   enabled = true                storage_aws_role_arn = 'arn:aws
                                          :iam::530595819296:role/newproductrole'
                                                   storage_allowed_locations = ('s3://bucketproductnew/');
+-----------------------------------------+
| status                                  |
|-----------------------------------------|
| Integration S3_INT2 successfully created. |
```

Next create the below stage :

```
Kousikdutta#COMPUTE_WH@MYFIRSTDATABASE.SALESSCHEMA>CREATE STAGE my_s3_stage3
                                                   STORAGE_INTEGRATION = s3_int2
                                                   URL = 's3://bucketproductnew/productkousik.csv'
                                                   file_format =my_csv_format;
+-------------------------------------------------+
| status                                          |
|-------------------------------------------------|
| Stage area MY_S3_STAGE3 successfully created. |
```

Don't forget to check the integration description :

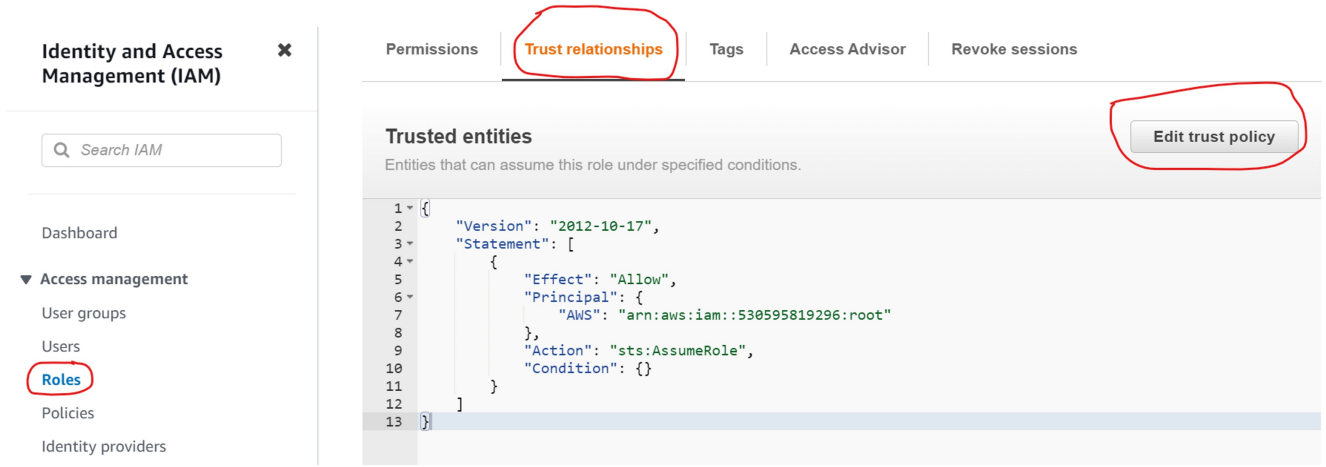
```
Kousikdutta#COMPUTE_WH@MYFIRSTDATABASE.SALESSCHEMA>desc integration s3_int2;
+---------------------------+---------------+-------------------------------------------------+------------------+
| property                  | property_type | property_value                                  | property_default |
|---------------------------+---------------+-------------------------------------------------+------------------|
| ENABLED                   | Boolean       | true                                            | false            |
| STORAGE_PROVIDER          | String        | S3                                              |                  |
| STORAGE_ALLOWED_LOCATIONS | List          | s3://bucketproductnew/                          | []               |
| STORAGE_BLOCKED_LOCATIONS | List          |                                                 | []               |
| STORAGE_AWS_IAM_USER_ARN  | String        | arn:aws:iam::942624237177:user/5psa0000-s       |                  |
| STORAGE_AWS_ROLE_ARN      | String        | arn:aws:iam::530595819296:role/newproductrole   |                  |
| STORAGE_AWS_EXTERNAL_ID   | String        | ZC56539_SFCRole=2_uXHmkGGm4cLVAW+GKXzERb/of1s=  |                  |
| COMMENT                   | String        |                                                 |                  |
+---------------------------+---------------+-------------------------------------------------+------------------+
8 Row(s) produced. Time Elapsed: 0.324s
Kousikdutta#COMPUTE_WH@MYFIRSTDATABASE.SALESSCHEMA>
```

Step 8: Next change the Trust relationship in role:

```
Kousikdutta#COMPUTE_WH@MYFIRSTDATABASE.SALESSCHEMA>desc integration s3_int2;
+---------------------------+---------------+-------------------------------------------------+------------------+
| property                  | property_type | property_value                                  | property_default |
|---------------------------+---------------+-------------------------------------------------+------------------|
| ENABLED                   | Boolean       | true                                            | false            |
| STORAGE_PROVIDER          | String        | S3                                              |                  |
| STORAGE_ALLOWED_LOCATIONS | List          | s3://bucketproductnew/                          | []               |
| STORAGE_BLOCKED_LOCATIONS | List          |                                                 | []               |
| STORAGE_AWS_IAM_USER_ARN  | String        | arn:aws:iam::942624237177:user/5psa0000-s       |                  |
| STORAGE_AWS_ROLE_ARN      | String        | arn:aws:iam::530595819296:role/newproductrole   |                  |
| STORAGE_AWS_EXTERNAL_ID   | String        | ZC56539_SFCRole=2_uXHmkGGm4cLVAW+GKXzERb/of1s=  |                  |
| COMMENT                   | String        |                                                 |                  |
+---------------------------+---------------+-------------------------------------------------+------------------+
8 Row(s) produced. Time Elapsed: 0.324s
Kousikdutta#COMPUTE_WH@MYFIRSTDATABASE.SALESSCHEMA>
```

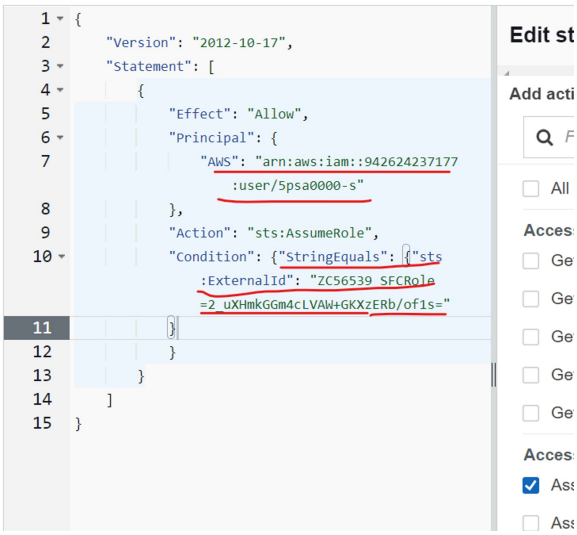
In the above " Storage_AWS_IAM_User_ARN " and " External ID " <- copy both of them and paste it in trust relationship :

Go to : -- > role > > Trust relationship --> Edit trust policy

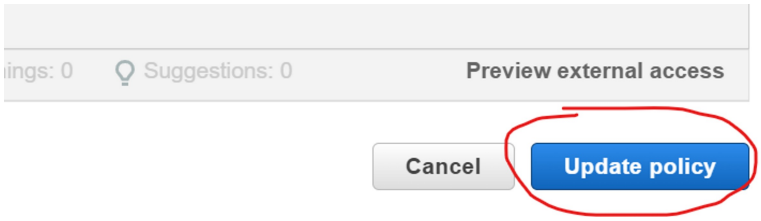


Let's edit that :

Below are the "Storage_AWS_IAM_User_ARN " and " External ID " values which we have got from the above, change them accordingly as shown below :



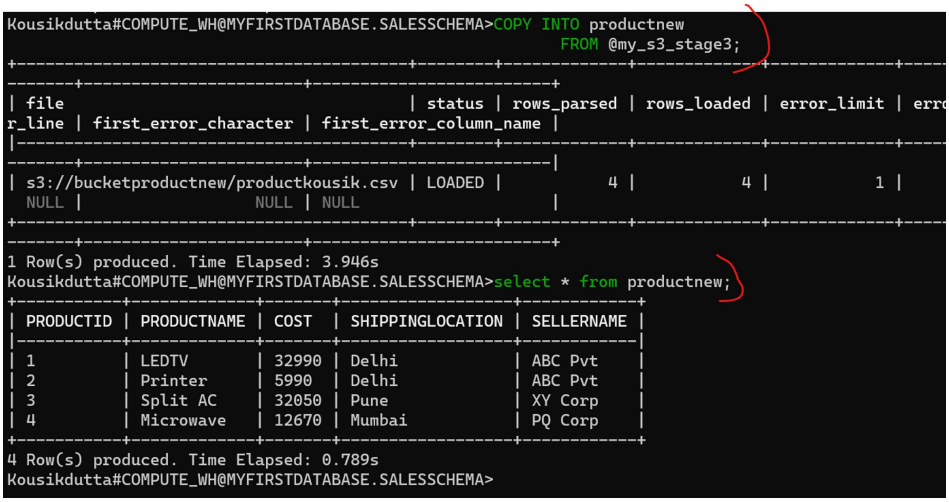
Click on update policy :



Step 9: Next create the schema in snowflake :

```
create table productnew (
ProductID varchar(512),
ProductName varchar(512),
Cost varchar(512),
ShippingLocation varchar(512),
SellerName varchar(512)
)
```

Step 10 : Load the data into that schema with the help of external stage :



```
Kousikdutta#COMPUTE_WH@MYFIRSTDATABASE.SALESSCHEMA>COPY INTO productnew
                                                    FROM @my_s3_stage3;
+-----------------------------------------+--------+-------------+-------------+-------------+-----
-------+----------------------+-------------------------+
| file                                    | status | rows_parsed | rows_loaded | error_limit | erro
r_line | first_error_character | first_error_column_name |
|-----------------------------------------+--------+-------------+-------------+-------------+-----
-------+----------------------+-------------------------|
| s3://bucketproductnew/productkousik.csv | LOADED |           4 |           4 |           1 |
  NULL |                 NULL | NULL                    |
+-----------------------------------------+--------+-------------+-------------+-------------+-----
-------+----------------------+-------------------------+
1 Row(s) produced. Time Elapsed: 3.946s
Kousikdutta#COMPUTE_WH@MYFIRSTDATABASE.SALESSCHEMA>select * from productnew;
+-----------+-------------+-------+------------------+------------+
| PRODUCTID | PRODUCTNAME | COST  | SHIPPINGLOCATION | SELLERNAME |
|-----------+-------------+-------+------------------+------------|
| 1         | LEDTV       | 32990 | Delhi            | ABC Pvt    |
| 2         | Printer     | 5990  | Delhi            | ABC Pvt    |
| 3         | Split AC    | 32050 | Pune             | XY Corp    |
| 4         | Microwave   | 12670 | Mumbai           | PQ Corp    |
+-----------+-------------+-------+------------------+------------+
4 Row(s) produced. Time Elapsed: 0.789s
Kousikdutta#COMPUTE_WH@MYFIRSTDATABASE.SALESSCHEMA>
```

Yes!!! We have received the data in snowflake.