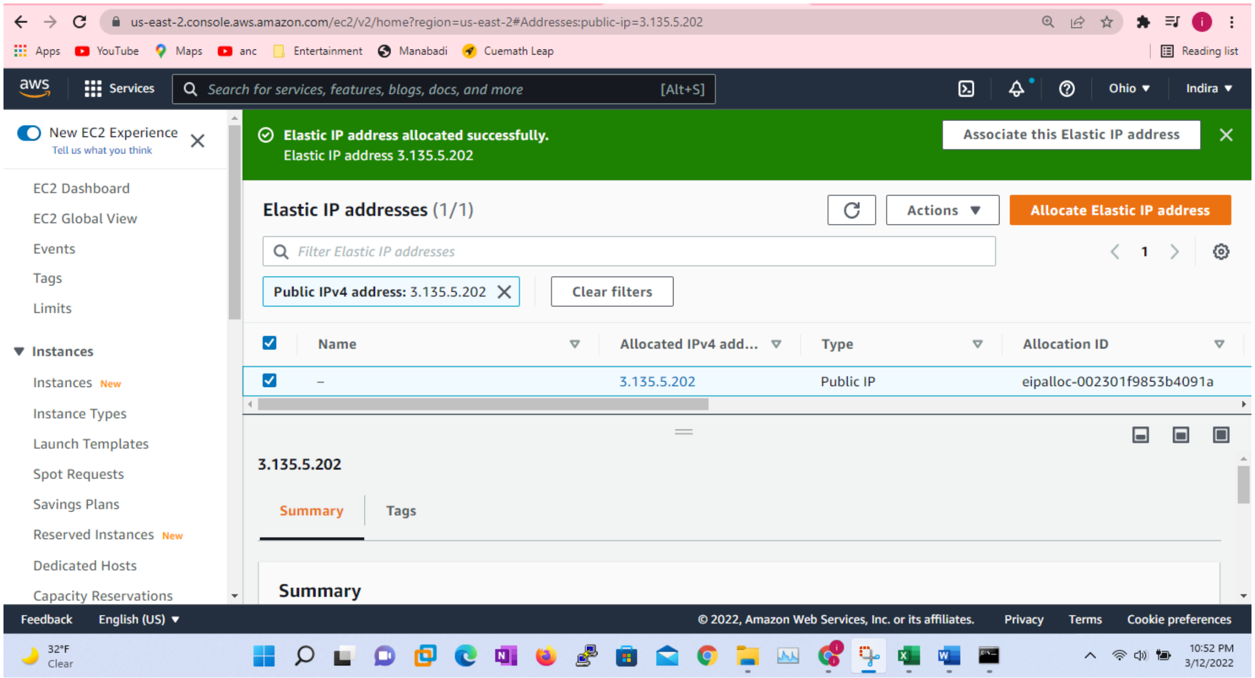
Develop Application using Redshift

To Connect the Redshift Cluster to the external world one of the ways is Elastic Ip (Or) Endpoints which is mapped to the elastic Ip address.

Go to EC2 instance Create elastic Ip



us-east-2.console.aws.amazon.com/ec2/v2/home?region=us-east-2#Addresses:public-ip=3.135.5.202

Elastic IP address allocated successfully.
Elastic IP address 3.135.5.202

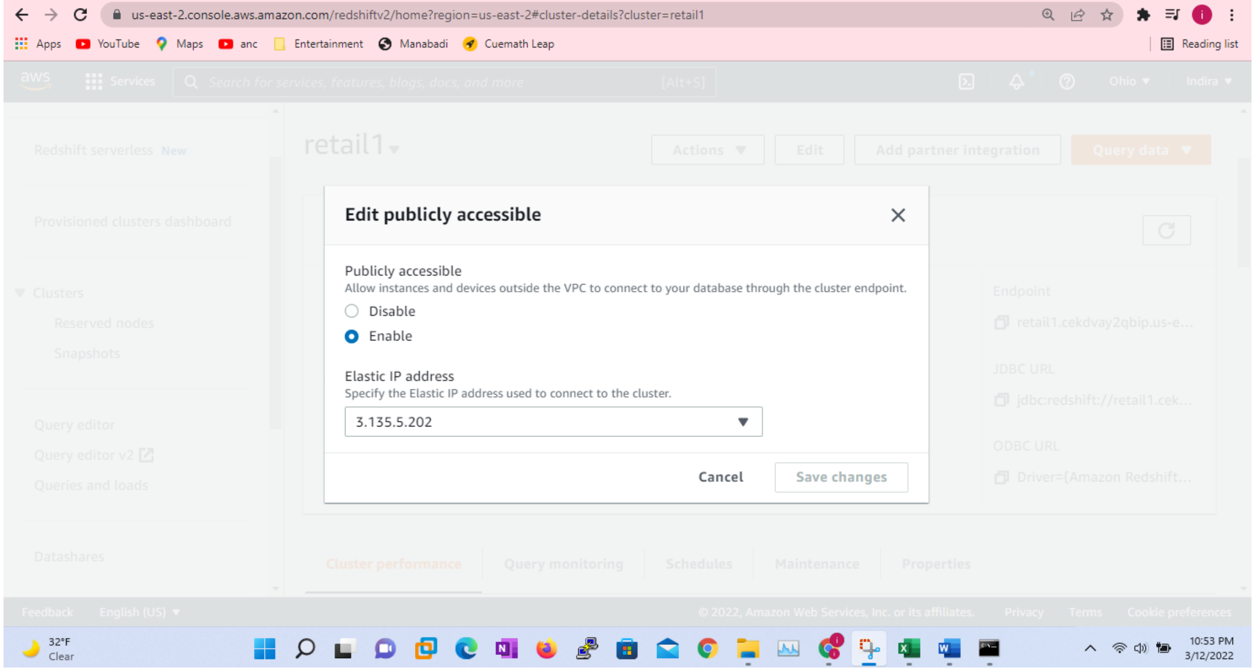
Elastic IP addresses (1/1)

Public IPv4 address: 3.135.5.202

☑	Name	Allocated IPv4 add...	Type	Allocation ID
☑	–	3.135.5.202	Public IP	eipalloc-002301f9853b4091a

3.135.5.202

Once the Elastic IP is created it is attached to the cluster in the Redshift. For that go to cluster in the Redshift and actions Modify publicity accessible setting, Attach the Elastic IP.



us-east-2.console.aws.amazon.com/redshiftv2/home?region=us-east-2#cluster-details?cluster=retail1

Edit publicly accessible

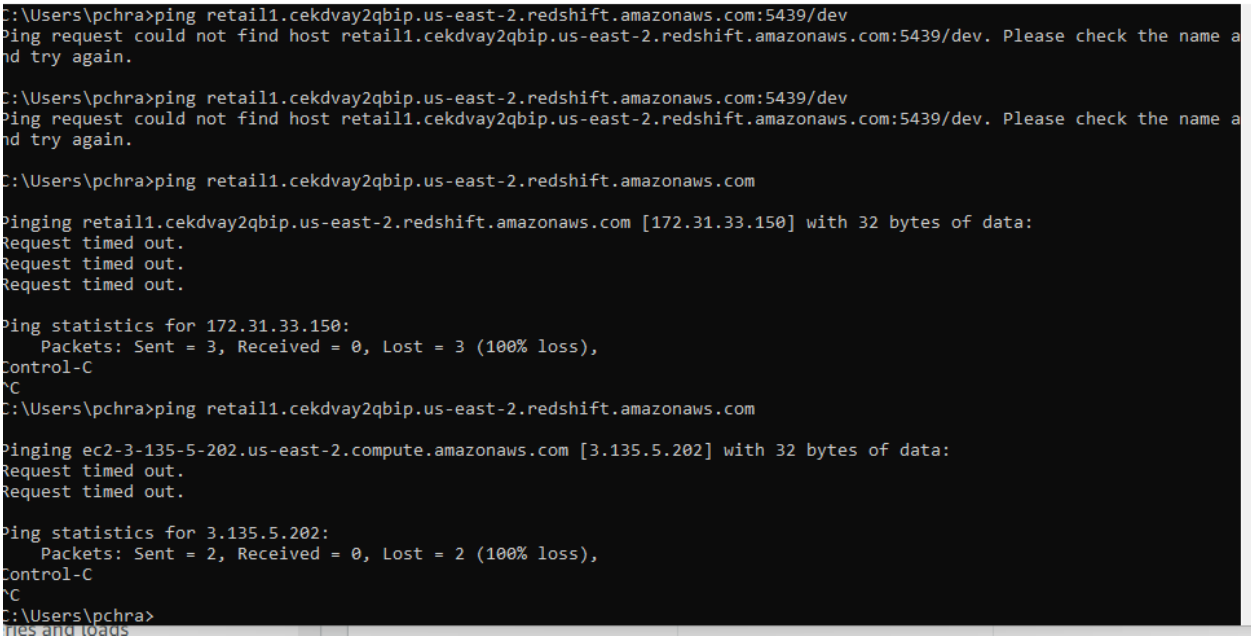
Publicly accessible
Allow instances and devices outside the VPC to connect to your database through the cluster endpoint.

○ Disable
● Enable

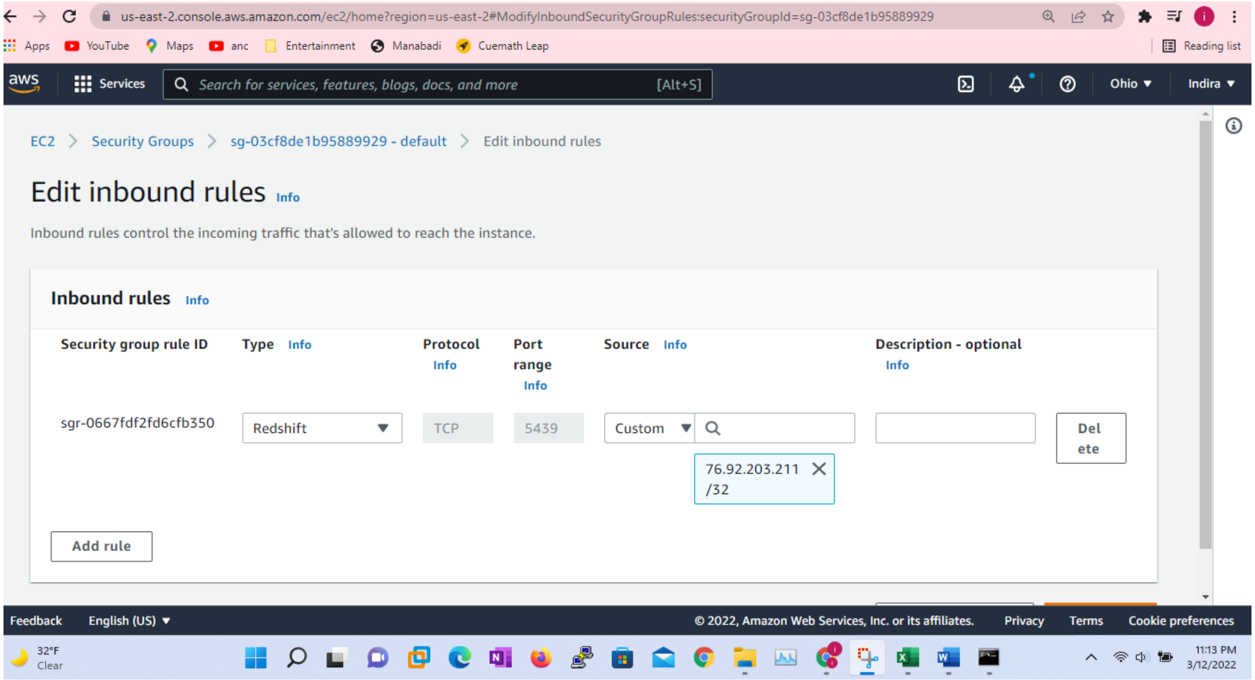
Elastic IP address
Specify the Elastic IP address used to connect to the cluster.

3.135.5.202 ▼

Validating whether the IP address is connected or not



```
C:\Users\pchra>ping retail1.cekdvay2qbip.us-east-2.redshift.amazonaws.com:5439/dev
Ping request could not find host retail1.cekdvay2qbip.us-east-2.redshift.amazonaws.com:5439/dev. Please check the name a
nd try again.

C:\Users\pchra>ping retail1.cekdvay2qbip.us-east-2.redshift.amazonaws.com:5439/dev
Ping request could not find host retail1.cekdvay2qbip.us-east-2.redshift.amazonaws.com:5439/dev. Please check the name a
nd try again.

C:\Users\pchra>ping retail1.cekdvay2qbip.us-east-2.redshift.amazonaws.com

Pinging retail1.cekdvay2qbip.us-east-2.redshift.amazonaws.com [172.31.33.150] with 32 bytes of data:
Request timed out.
Request timed out.
Request timed out.

Ping statistics for 172.31.33.150:
    Packets: Sent = 3, Received = 0, Lost = 3 (100% loss),
Control-C
^C
C:\Users\pchra>ping retail1.cekdvay2qbip.us-east-2.redshift.amazonaws.com

Pinging ec2-3-135-5-202.us-east-2.compute.amazonaws.com [3.135.5.202] with 32 bytes of data:
Request timed out.
Request timed out.

Ping statistics for 3.135.5.202:
    Packets: Sent = 2, Received = 0, Lost = 2 (100% loss),
Control-C
^C
C:\Users\pchra>
```



us-east-2.console.aws.amazon.com/ec2/home?region=us-east-2#ModifyInboundSecurityGroupRules:securityGroupId=sg-03cf8de1b95889929

EC2 > Security Groups > sg-03cf8de1b95889929 - default > Edit inbound rules

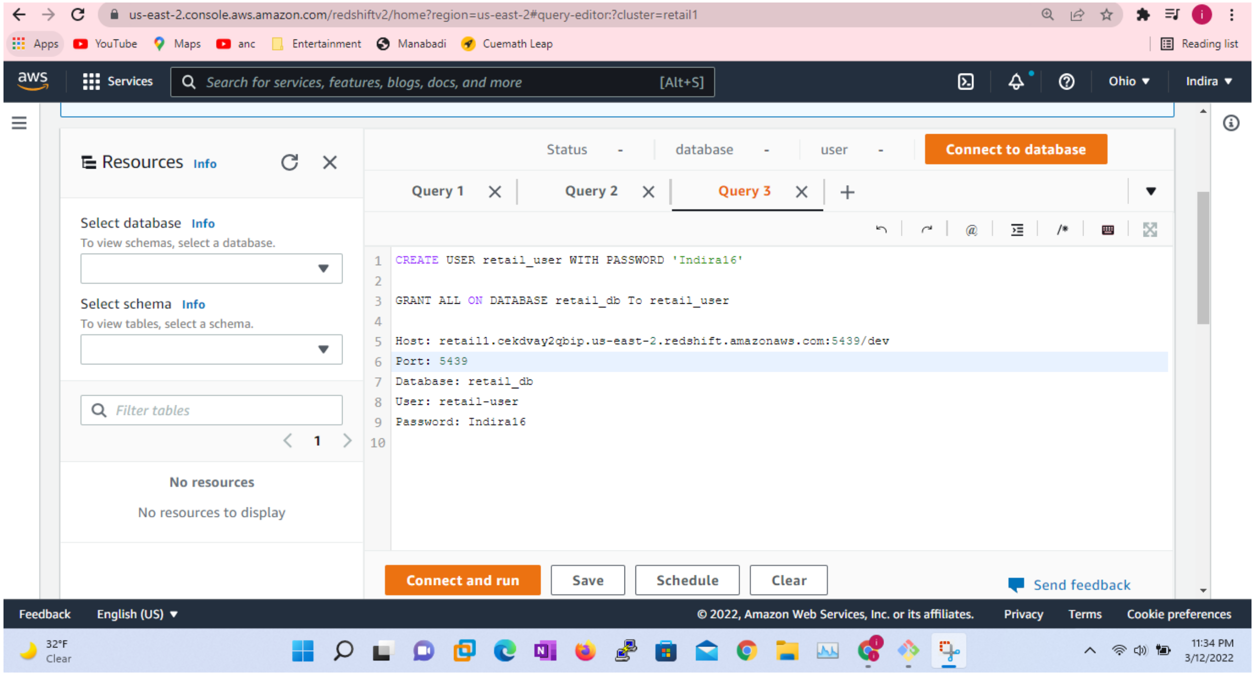
Edit inbound rules Info

Inbound rules control the incoming traffic that's allowed to reach the instance.

Inbound rules Info

Security group rule ID	Type Info	Protocol Info	Port range Info	Source Info	Description - optional Info	
sgr-0667fdf2fd6cfb350	Redshift ▼	TCP	5439	Custom ▼ 76.92.203.211/32 ✕		Delete

Create Table and User in RedShift Cluster

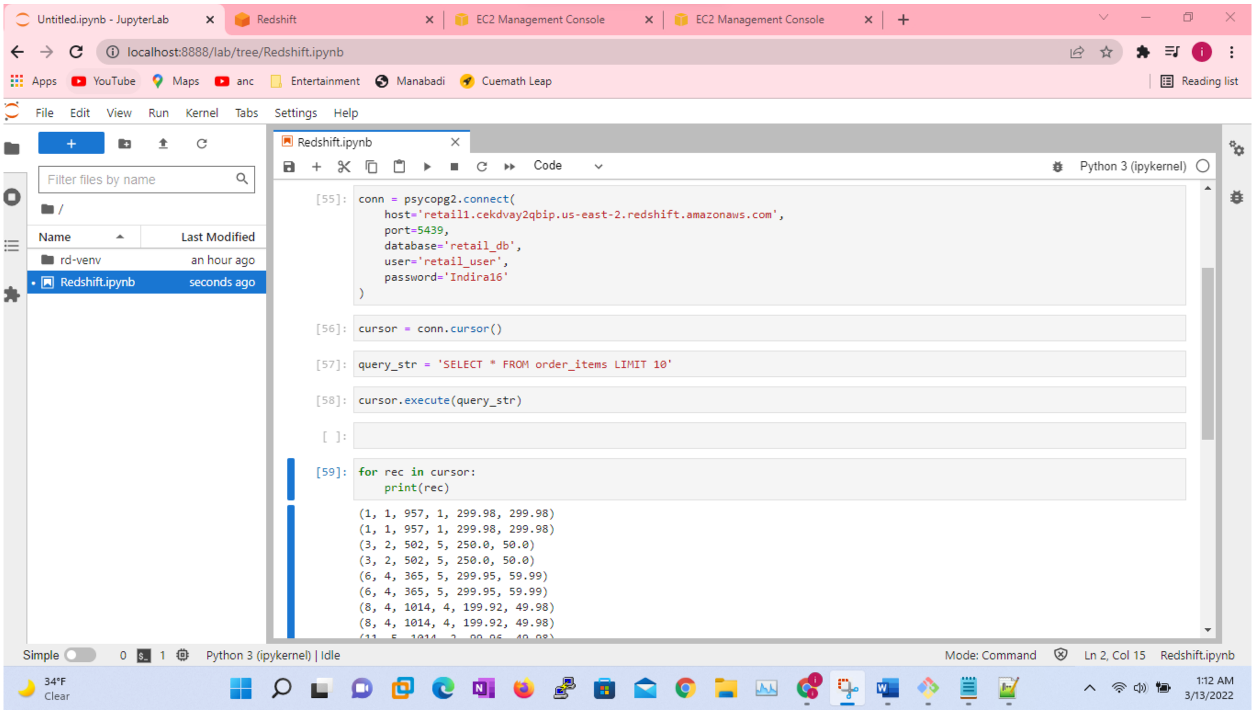


```
CREATE USER retail_user WITH PASSWORD 'Indira16'

GRANT ALL ON DATABASE retail_db To retail_user

Host: retail1.cekdvay2qbip.us-east-2.redshift.amazonaws.com:5439/dev
Port: 5439
Database: retail_db
User: retail-user
Password: Indira16
```

Psycopy is a popular python based driver which is used to connect postgresql database, As redshift is nothing but a flavor of postgres database , so we should leverage psycopy objective to connect to database that are running as part of redshift cluster as well,

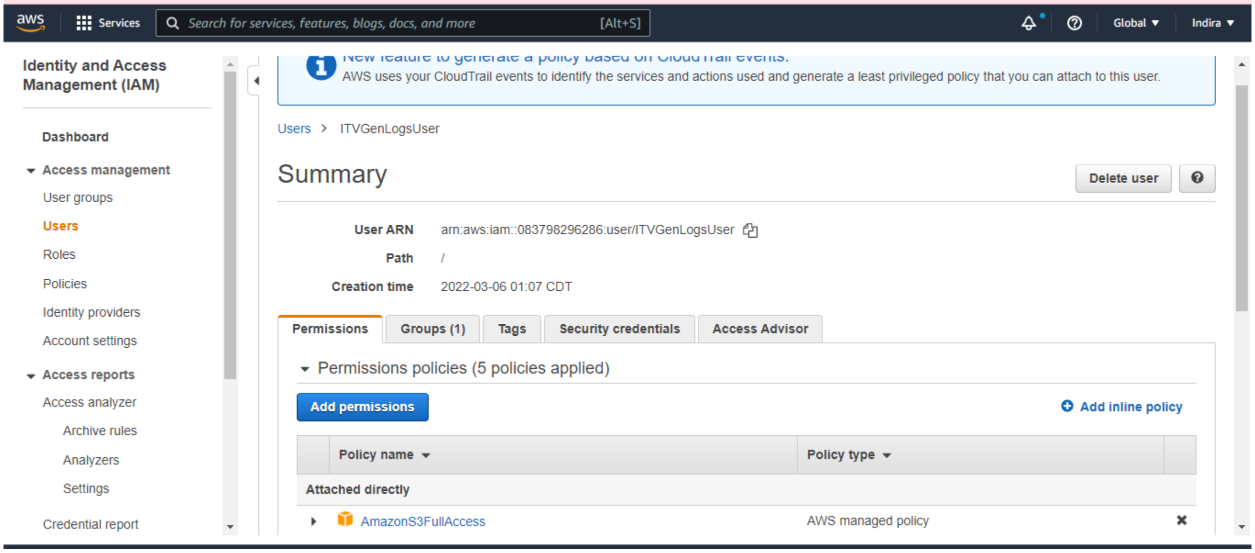
If we install psyscopy we should be able to use python as a programming language to connect to red shift databases.

Run Simple Query against RedShift Database Table using Python



```
[55]: conn = psycopg2.connect(
          host='retail1.cekdvay2qbip.us-east-2.redshift.amazonaws.com',
          port=5439,
          database='retail_db',
          user='retail_user',
          password='Indira16'
      )

[56]: cursor = conn.cursor()

[57]: query_str = 'SELECT * FROM order_items LIMIT 10'

[58]: cursor.execute(query_str)

[ ]:

[59]: for rec in cursor:
          print(rec)

      (1, 1, 957, 1, 299.98, 299.98)
      (1, 1, 957, 1, 299.98, 299.98)
      (3, 2, 502, 5, 250.0, 50.0)
      (3, 2, 502, 5, 250.0, 50.0)
      (6, 4, 365, 5, 299.95, 59.99)
      (6, 4, 365, 5, 299.95, 59.99)
      (8, 4, 1014, 4, 199.92, 49.98)
      (8, 4, 1014, 4, 199.92, 49.98)
      (11, 5, 1014, 2, 99.96, 49.98)
```

Copying data from s3 into Redshift Tables using Python as programing language. For that Copy command for external systems , We need to make sure that user have a valid permission on S3.



New feature to generate a policy based on CloudTrail events.
AWS uses your CloudTrail events to identify the services and actions used and generate a least privileged policy that you can attach to this user.

Users > ITVGenLogsUser

Summary

User ARN arn:aws:iam::083798296286:user/ITVGenLogsUser

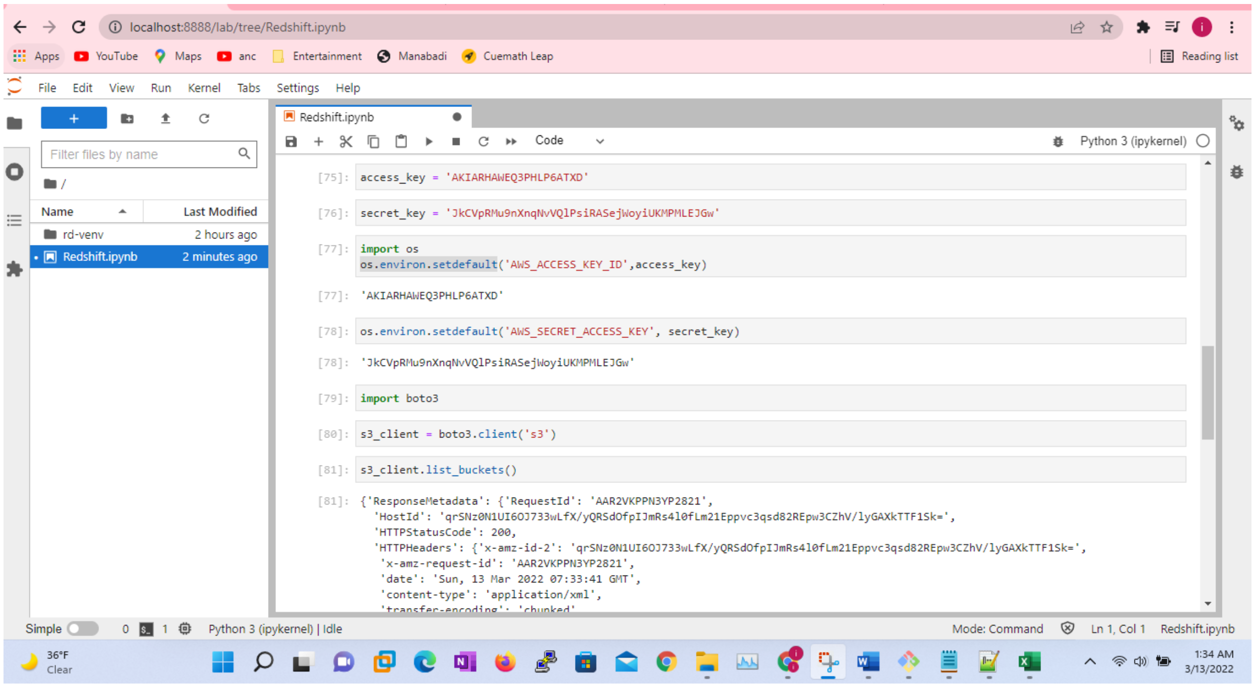
Path /

Creation time 2022-03-06 01:07 CDT

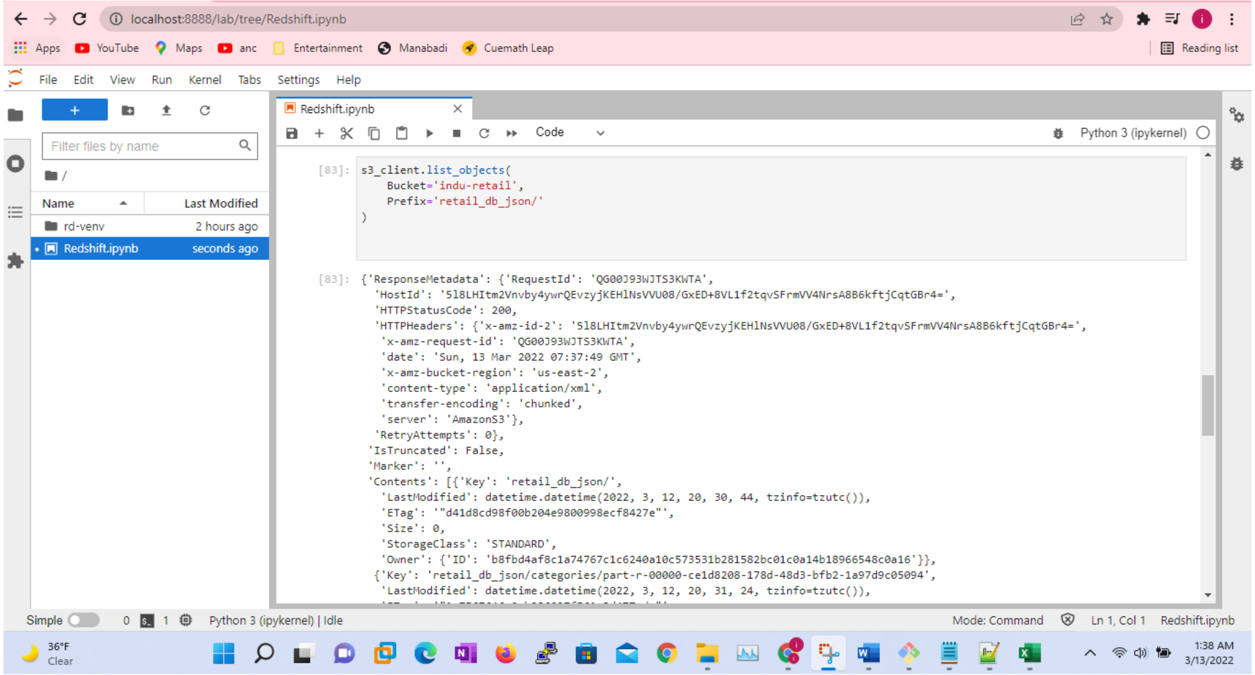
Permissions policies (5 policies applied)

Policy name	Policy type
Attached directly	
AmazonS3FullAccess	AWS managed policy

Validate Access of IAM User Using Boto3



```
[75]: access_key = 'AKIARHAWEQ3PHLP6ATXD'

[76]: secret_key = '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'

[77]: import os
      os.environ.setdefault('AWS_ACCESS_KEY_ID',access_key)

[77]: 'AKIARHAWEQ3PHLP6ATXD'

[78]: os.environ.setdefault('AWS_SECRET_ACCESS_KEY', secret_key)

[78]: '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'

[79]: import boto3

[80]: s3_client = boto3.client('s3')

[81]: s3_client.list_buckets()

[81]: {'ResponseMetadata': {'RequestId': 'AAR2VKPPN3YP2821',
        'HostId': 'qrSNz0N1UI6OJ733wLfX/yQRSdOfpIJmRs4l0fLm21Eppvc3qsd82REpw3CZhV/lyGAXkTTF1Sk=',
        'HTTPStatusCode': 200,
        'HTTPHeaders': {'x-amz-id-2': 'qrSNz0N1UI6OJ733wLfX/yQRSdOfpIJmRs4l0fLm21Eppvc3qsd82REpw3CZhV/lyGAXkTTF1Sk=',
         'x-amz-request-id': 'AAR2VKPPN3YP2821',
         'date': 'Sun, 13 Mar 2022 07:33:41 GMT',
         'content-type': 'application/xml',
         'transfer-encoding': 'chunked',
```



```
[83]: s3_client.list_objects(
          Bucket='indu-retail',
          Prefix='retail_db_json/'
      )

[83]: {'ResponseMetadata': {'RequestId': 'QG00J93WJTS3KWTA',
        'HostId': '5l8LHItm2Vnvby4ywrQEvzyjKEHlNsVVU08/GxED+8VL1f2tqvSFrmVV4NrsA8B6kftjCqtGBr4=',
        'HTTPStatusCode': 200,
        'HTTPHeaders': {'x-amz-id-2': '5l8LHItm2Vnvby4ywrQEvzyjKEHlNsVVU08/GxED+8VL1f2tqvSFrmVV4NrsA8B6kftjCqtGBr4=',
         'x-amz-request-id': 'QG00J93WJTS3KWTA',
         'date': 'Sun, 13 Mar 2022 07:37:49 GMT',
         'x-amz-bucket-region': 'us-east-2',
         'content-type': 'application/xml',
         'transfer-encoding': 'chunked',
         'server': 'AmazonS3'},
        'RetryAttempts': 0},
       'IsTruncated': False,
       'Marker': '',
       'Contents': [{'Key': 'retail_db_json/',
         'LastModified': datetime.datetime(2022, 3, 12, 20, 30, 44, tzinfo=tzutc()),
         'ETag': '"d41d8cd98f00b204e9800998ecf8427e"',
         'Size': 0,
         'StorageClass': 'STANDARD',
         'Owner': {'ID': 'b8fbd4af8c1a74767c1c6240a10c573531b281582bc01c0a14b18966548c0a16'}},
        {'Key': 'retail_db_json/categories/part-r-00000-ce1d8208-178d-48d3-bfb2-1a97d9c05094',
         'LastModified': datetime.datetime(2022, 3, 12, 20, 31, 24, tzinfo=tzutc()),
```

By the Copy Command The data from s3 is copied into the Redshift Using Python

Copy_stmt = f”””
COPY {table_name}

```
FROM ‘{s3_location}’
CREDENTIALS ‘aws_access_key_id={access_key};aws_secret_access_key={seret_key}’
JSON AS ‘auto’
“””

Cursor = conn.cursor()
Cursor.execute(copy_stmt)
Query_stmt = ‘SELECT count(*) FROM order_items’
Cursor.execute(query_stmt)
Cursor.fetchall()
```