



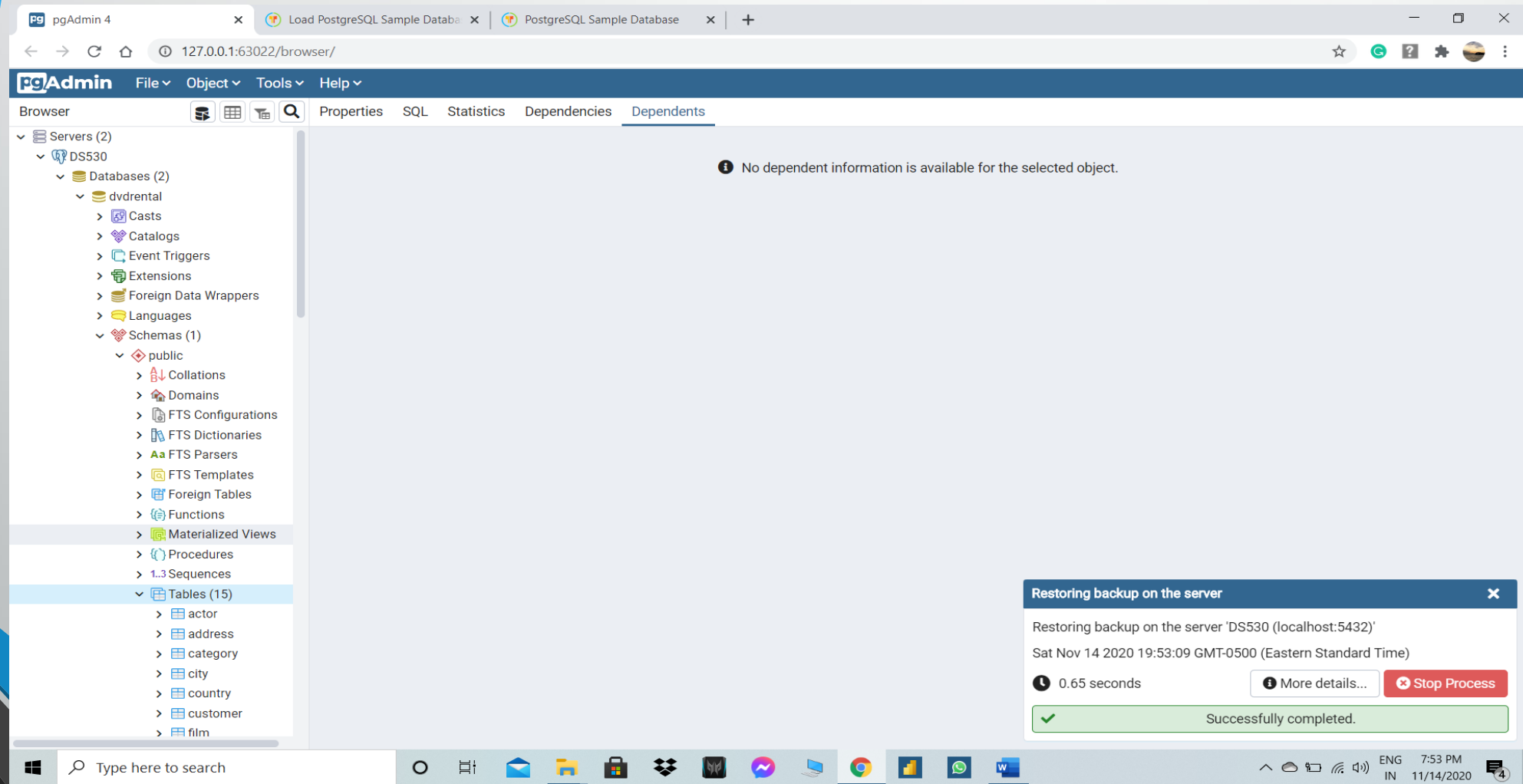
DS 530 Project

By Arpan Patel

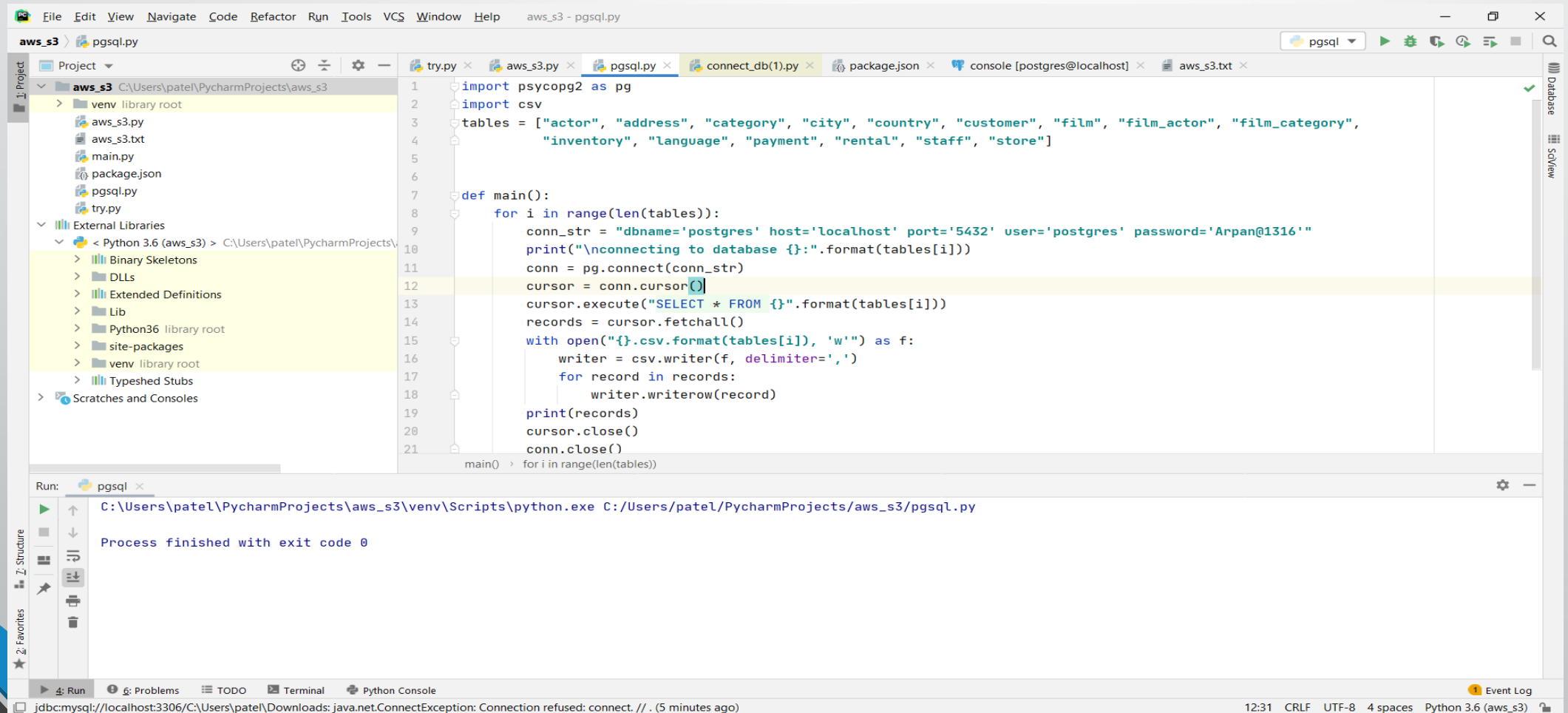
Resources

- PostgreSQL
- PgAdmin
- Pycharm
- Aws
- Aws redshift
- Aws S3

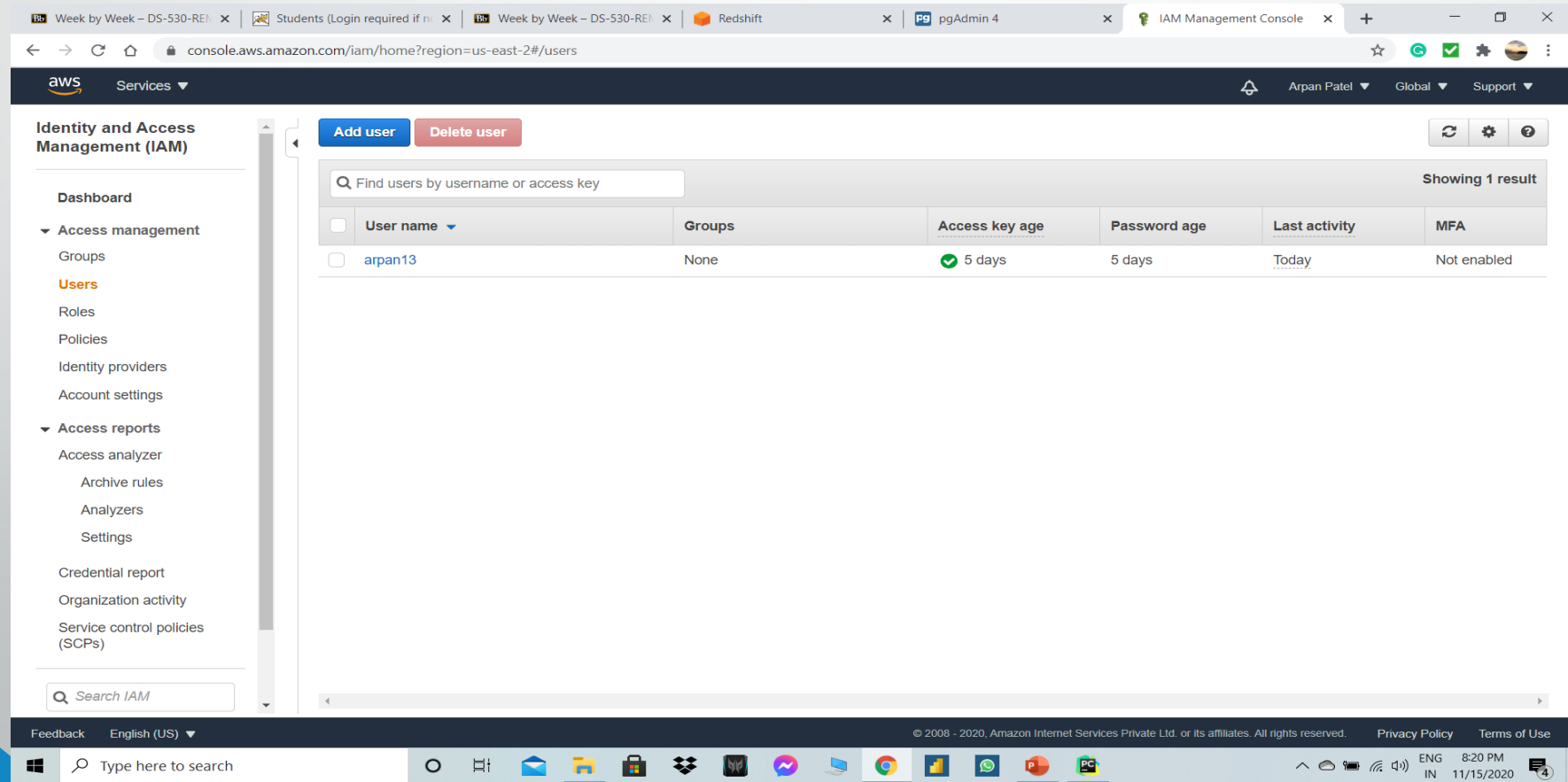
PostgreSQL pgadmin4



Pycharm code to convert tables to .csv



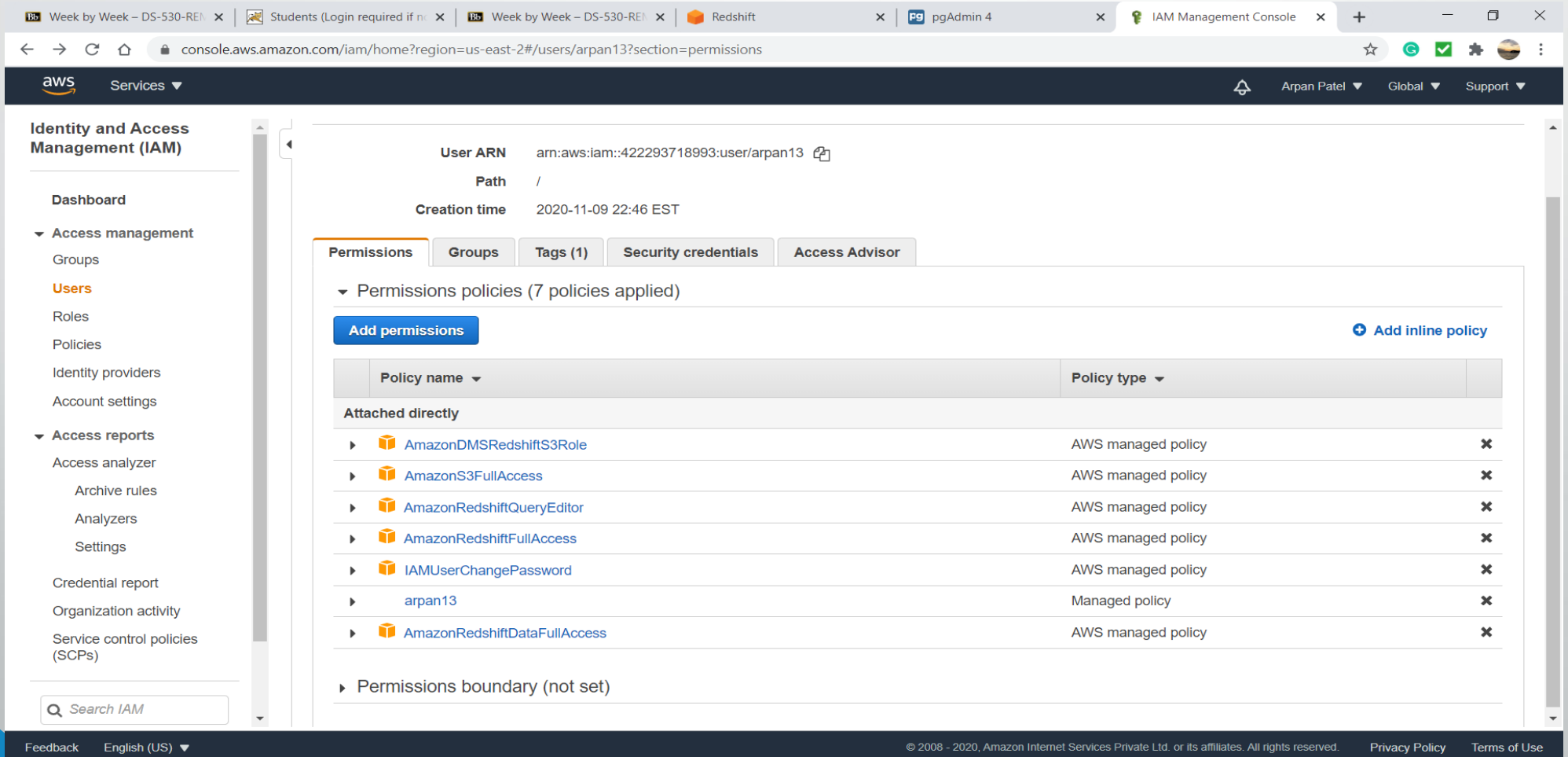
Create IAM user and assign access to s3 Bucket



The screenshot displays the AWS IAM Management Console interface. The left sidebar shows the navigation menu with 'Users' highlighted under 'Access management'. The main content area shows a table of users with one entry, 'arpan13'. The table columns are: User name, Groups, Access key age, Password age, Last activity, and MFA. The user 'arpan13' has no groups, an access key age of 5 days, a password age of 5 days, last activity today, and MFA is not enabled. The top navigation bar includes the AWS logo, 'Services' dropdown, and user profile 'Arpan Patel'. The bottom status bar shows the date '11/15/2020' and time '8:20 PM'.

User name	Groups	Access key age	Password age	Last activity	MFA
arpan13	None	5 days	5 days	Today	Not enabled

Assign full access to user



The screenshot shows the AWS IAM Management Console interface. The left sidebar contains navigation links for Identity and Access Management (IAM), including Dashboard, Access management, Users, Roles, Policies, Identity providers, Account settings, Access reports, Access analyzer, Archive rules, Analyzers, Settings, Credential report, Organization activity, and Service control policies (SCPs). The main content area displays the details for user **arpan13**, including the User ARN, Path, and Creation time. The **Permissions** tab is selected, showing a list of 7 policies applied to the user. The policies are categorized as 'Attached directly' and include AmazonDMSRedshiftS3Role, AmazonS3FullAccess, AmazonRedshiftQueryEditor, AmazonRedshiftFullAccess, IAMUserChangePassword, arpan13, and AmazonRedshiftDataFullAccess. The 'Permissions boundary' is set to '(not set)'.

Week by Week – DS-530-REI x Students (Login required if n x Week by Week – DS-530-REI x Redshift x pgAdmin 4 x IAM Management Console x

console.aws.amazon.com/iam/home?region=us-east-2#/users/arpan13?section=permissions

aws Services

Arpan Patel Global Support

Identity and Access Management (IAM)

- Dashboard
- Access management
 - Groups
 - Users**
 - Roles
 - Policies
 - Identity providers
 - Account settings
- Access reports
 - Access analyzer
 - Archive rules
 - Analyzers
 - Settings
 - Credential report
 - Organization activity
 - Service control policies (SCPs)

Search IAM

User ARN: arn:aws:iam::422293718993:user/arpan13

Path: /

Creation time: 2020-11-09 22:46 EST

Permissions Groups Tags (1) Security credentials Access Advisor

Permissions policies (7 policies applied)

Add permissions Add inline policy

Policy name	Policy type
Attached directly	
AmazonDMSRedshiftS3Role	AWS managed policy
AmazonS3FullAccess	AWS managed policy
AmazonRedshiftQueryEditor	AWS managed policy
AmazonRedshiftFullAccess	AWS managed policy
IAMUserChangePassword	AWS managed policy
arpan13	Managed policy
AmazonRedshiftDataFullAccess	AWS managed policy

Permissions boundary (not set)

Feedback English (US)

© 2008 - 2020, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved. Privacy Policy Terms of Use

S3 Bucket

The screenshot displays the AWS S3 Management Console interface. The browser tabs show 'pgAdmin 4' and 'S3 Management Console'. The address bar indicates the URL: 's3.console.aws.amazon.com/s3/buckets/arpan13?region=us-east-2&tab=objects'. The AWS logo and 'Services' dropdown are in the top left. The top right shows the user 'Arpan Patel', 'Global' region, and 'Support' link. A blue banner at the top right contains a message about S3 console improvements and a 'Provide feedback' button. The left sidebar shows the 'Amazon S3' menu with options: 'Buckets', 'Access points', 'Batch Operations', 'Access analyzer for S3', 'Account settings for Block Public Access', and 'Feature spotlight' (with a '2' badge). The main content area shows the 'arpan13' bucket overview. A table lists bucket details: Region (US East (Ohio) us-east-2), Amazon resource name (ARN) (arn:aws:s3:::arpan13), Creation date (November 9, 2020, 22:44 (UTC-05:00)), and Access (Objects can be public). Below the table are tabs for 'Objects', 'Properties', 'Permissions', 'Metrics', 'Management', and 'Access points'. A large dashed box prompts the user to 'Drag and drop files and folders you want to upload here, or choose Upload.' Below this is the 'Objects (13)' section, which includes a description of objects, a 'Learn more' link, and a toolbar with 'List versions', 'Delete', 'Actions', 'Create folder', and 'Upload' buttons. A search bar at the bottom of the objects section says 'Find objects by prefix'. The bottom of the console shows a footer with 'Feedback', 'English (US)' language selector, and copyright information: '© 2008 - 2020, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved.' along with 'Privacy Policy' and 'Terms of Use' links.

pgAdmin 4 x S3 Management Console x +

s3.console.aws.amazon.com/s3/buckets/arpan13?region=us-east-2&tab=objects

aws Services

Arpan Patel Global Support

Amazon S3 X

Buckets

Access points

Batch Operations

Access analyzer for S3

Account settings for Block Public Access

Feature spotlight 2

arpan13

Bucket overview

Region	Amazon resource name (ARN)	Creation date	Access
US East (Ohio) us-east-2	arn:aws:s3:::arpan13	November 9, 2020, 22:44 (UTC-05:00)	Objects can be public

Objects Properties Permissions Metrics Management Access points

Drag and drop files and folders you want to upload here, or choose **Upload**.

Objects (13)

Objects are the fundamental entities stored in Amazon S3. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

List versions Delete Actions Create folder Upload

Find objects by prefix

Feedback English (US)

© 2008 - 2020, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved. Privacy Policy Terms of Use

Pycharm code to upload files into s3 bucket

```
import boto3

s3 = boto3.client('s3')
s3.upload_file('main.py', 'arpan13', 'actor.csv')
s3.upload_file('address.py', 'arpan13', 'address.csv')
s3.upload_file('category.py', 'arpan13', 'category.csv')
s3.upload_file('city.py', 'arpan13', 'city.csv')
s3.upload_file('country.py', 'arpan13', 'country.csv')
s3.upload_file('customer.py', 'arpan13', 'customer.csv')
s3.upload_file('film.py', 'arpan13', 'film.csv')
s3.upload_file('film_actor.py', 'arpan13', 'film_actor.csv')
s3.upload_file('film_category.py', 'arpan13', 'film_category.csv')
s3.upload_file('inventory.py', 'arpan13', 'inventory.csv')
s3.upload_file('language.py', 'arpan13', 'language.csv')
s3.upload_file('payment.py', 'arpan13', 'payment.csv')
s3.upload_file('rental.py', 'arpan13', 'rental.csv')
```


Csv files in S3 Bucket

The screenshot displays the AWS S3 Management Console interface. The left sidebar shows the 'Amazon S3' navigation menu with options like Buckets, Access points, Batch Operations, and Account settings. The main content area is titled 'Objects (13)' and shows a list of CSV files stored in the bucket 'arpan13'. The files are listed with their names, types, last modified dates, sizes, and storage classes. The interface includes a search bar, pagination controls, and various action buttons like 'List versions', 'Delete', 'Actions', 'Create folder', and 'Upload'.

Amazon S3

Buckets

Access points

Batch Operations

Access analyzer for S3

Account settings for Block Public Access

Feature spotlight **2**

Objects (13)

Objects are the fundamental entities stored in Amazon S3. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

☐ List versions

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	actor.csv	csv	November 15, 2020, 14:01 (UTC-05:00)	8.6 KB	Standard
<input type="checkbox"/>	address.csv	csv	November 15, 2020, 14:01 (UTC-05:00)	51.0 KB	Standard
<input type="checkbox"/>	category.csv	csv	November 15, 2020, 14:01 (UTC-05:00)	590.0 B	Standard
<input type="checkbox"/>	city.csv	csv	November 15, 2020, 14:01 (UTC-05:00)	25.3 KB	Standard
<input type="checkbox"/>	country.csv	csv	November 15, 2020, 14:01 (UTC-05:00)	4.3 KB	Standard
<input type="checkbox"/>	customer.csv	csv	November 15, 2020, 14:01 (UTC-05:00)	59.3 KB	Standard
<input type="checkbox"/>	film_actor.csv	csv	November 15, 2020, 14:01 (UTC-05:00)	183.3 KB	Standard
<input type="checkbox"/>	film_category.csv	csv	November 15, 2020, 14:01 (UTC-05:00)	32.5 KB	Standard
<input type="checkbox"/>	film.csv	csv	November 15, 2020, 14:01 (UTC-05:00)	327.8 KB	Standard
<input type="checkbox"/>	inventory.csv	csv	November 15, 2020, 14:01 (UTC-05:00)	168.4 KB	Standard
<input type="checkbox"/>	language.csv	csv	November 15, 2020, 14:01 (UTC-05:00)	300.0 B	Standard

Feedback English (US) © 2008 - 2020, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved. Privacy Policy Terms of Use

Amazon Redshift Cluster

The screenshot displays the Amazon Redshift console interface. The top navigation bar includes the AWS logo, a 'Services' dropdown, and user information for 'Arpan Patel' in the 'Ohio' region. The left sidebar contains navigation links for Dashboard, Clusters (highlighted), Queries, Editor, Config, Marketplace, Alarms, Events, and What's New. The main content area is titled 'Amazon Redshift > Clusters' and shows a list of clusters under the heading 'Clusters (1)'. A search bar and a status filter set to 'All status' are present. The cluster list table has columns for Cluster, Status, Storage capacity, CPU utilization, Snapshots, Notifications, and Tags. One cluster, 'redshift-cluster-3', is listed with a status of 'Available', storage capacity of '< 1%', CPU utilization of '8%', and '3 snapshots'. The bottom of the page features a footer with 'Feedback', 'English (US)' language selection, and copyright information for 2008-2020.

pgAdmin 4 x Redshift x +

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

aws Services

Arpan Patel Ohio Support

Amazon Redshift > Clusters

Clusters (1)

Search All status < 1 > ⚙

<input type="checkbox"/>	Cluster ▲	Status ▼	Storage capacity us... ▼	CPU utilization ▼	Snapsh... ▼	Notificati... ▼	Tags ▼
<input type="checkbox"/>	redshift-cluster-3 dc2.large 1 node 160 GB	Available	< 1%	8%	3 snapshots		

Feedback English (US) © 2008 - 2020, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved. Privacy Policy Terms of Use

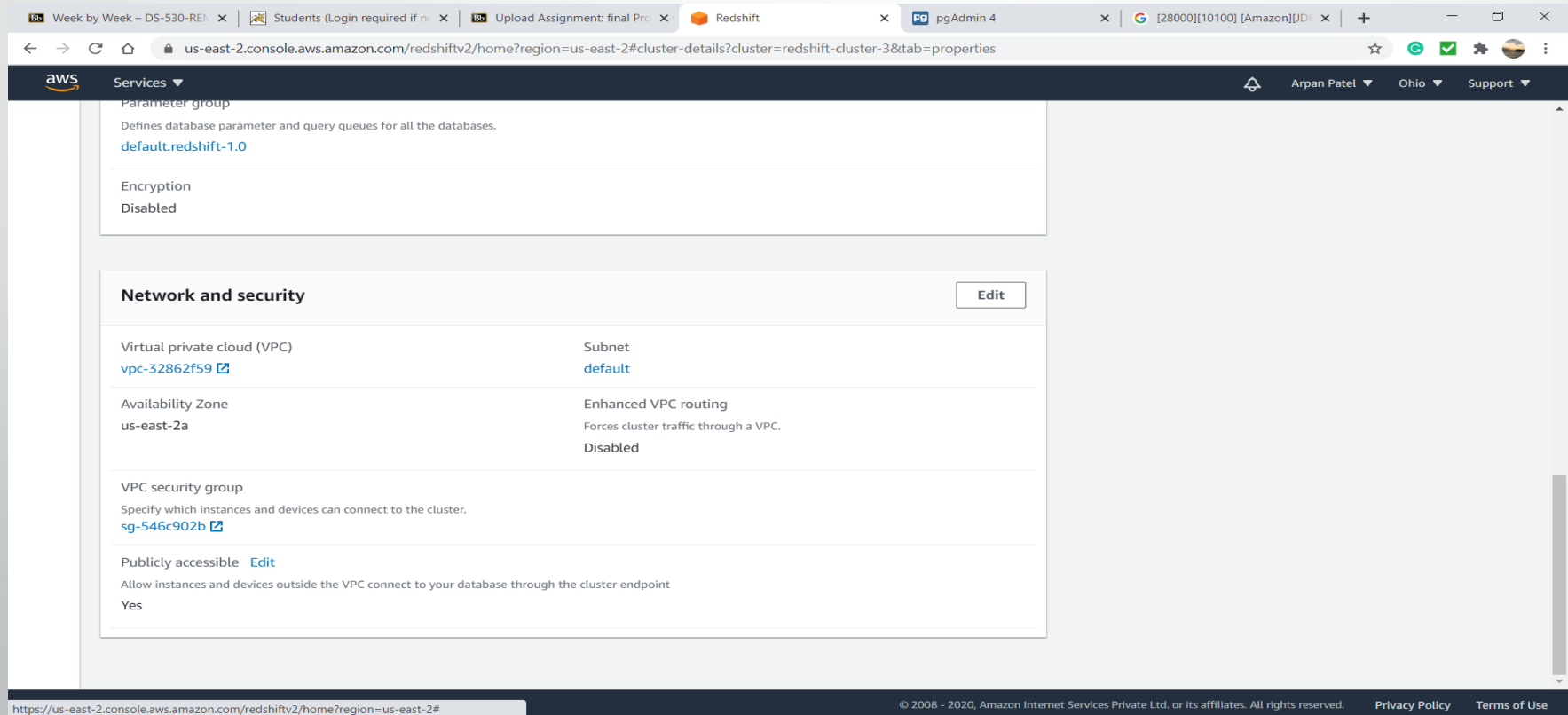
Modifying cluster for public access to yes

The screenshot displays the AWS Redshift console interface. The browser's address bar shows the URL `us-east-2.console.aws.amazon.com/redshiftv2/home?region=us-east-2#clusters`. The left-hand navigation menu includes options for Dashboard, Clusters (which is highlighted), Queries, Editor, Config, Marketplace, Alarms, Events, and What's New. The main content area is titled 'Amazon Redshift > Clusters' and features a 'Clusters (1)' section with a search bar and a status filter set to 'All status'. Below this is a table listing the cluster details:

	Cluster	Status	Storage capacity us...	CPU utilization	Snapsh...	Notificati...	Tags
<input type="checkbox"/>	redshift-cluster-3 dc2.large 1 node 160 GB	Modifying Modifying	< 1%	12%	3 snapshots		

At the top of the cluster list, there are buttons for 'Query cluster', 'Actions', and 'Create cluster'. The bottom of the screen shows the Windows taskbar with the search bar and various application icons. The footer of the AWS console includes a 'Feedback' link, the language 'English (US)', and copyright information for Amazon Internet Services Private Ltd. (2008-2020).

Now publicly accessible to yes if you need to access from internet gateway

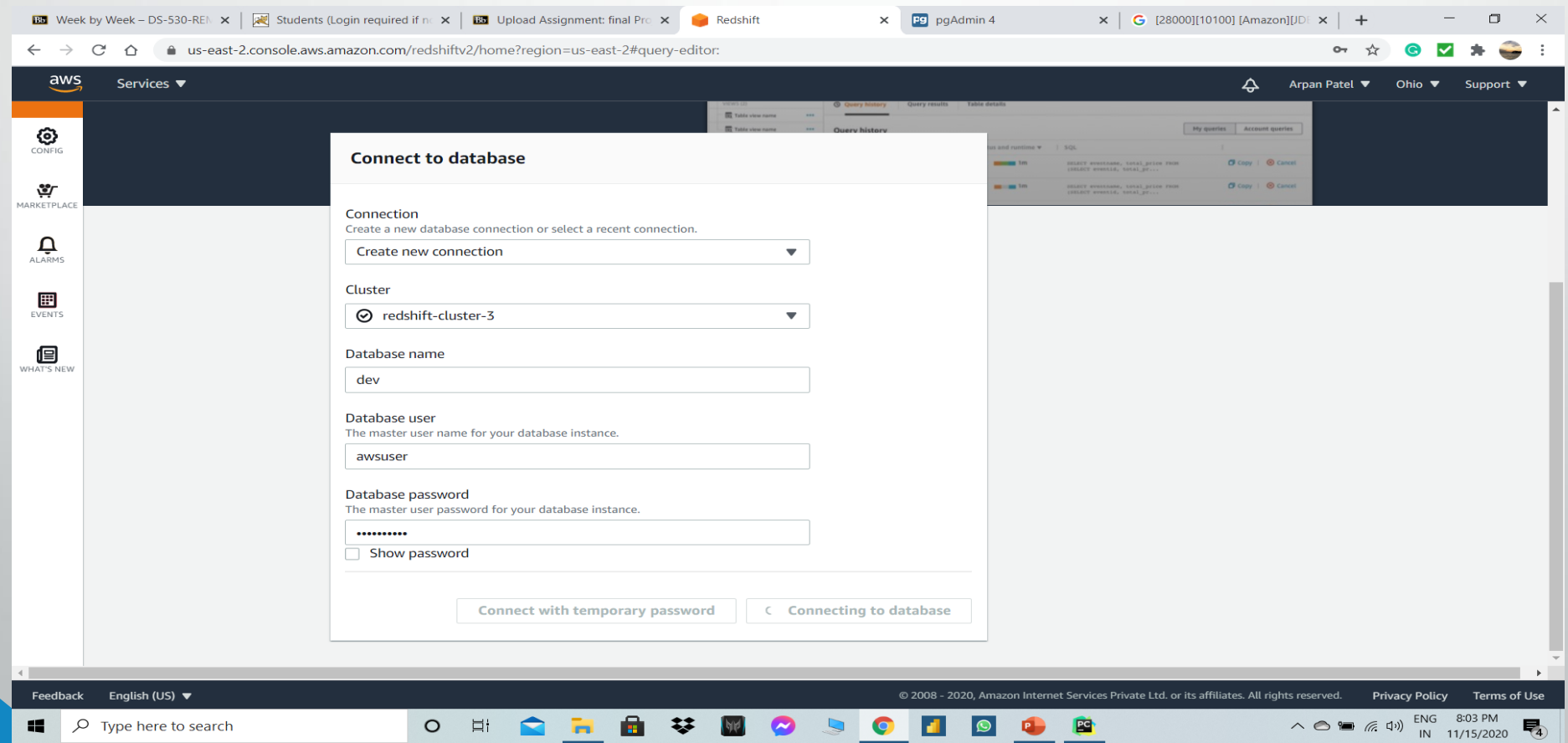


The screenshot displays the AWS Redshift console interface. The browser's address bar shows the URL: `us-east-2.console.aws.amazon.com/redshiftv2/home?region=us-east-2#cluster-details?cluster=redshift-cluster-3&tab=properties`. The console header includes the AWS logo, a 'Services' dropdown, and user information for 'Arpan Patel' in the 'Ohio' region. The main content area is titled 'Parameter group' and shows details for 'default.redshift-1.0', including 'Encryption' which is 'Disabled'. Below this, the 'Network and security' section is expanded, showing an 'Edit' button. This section contains the following details:

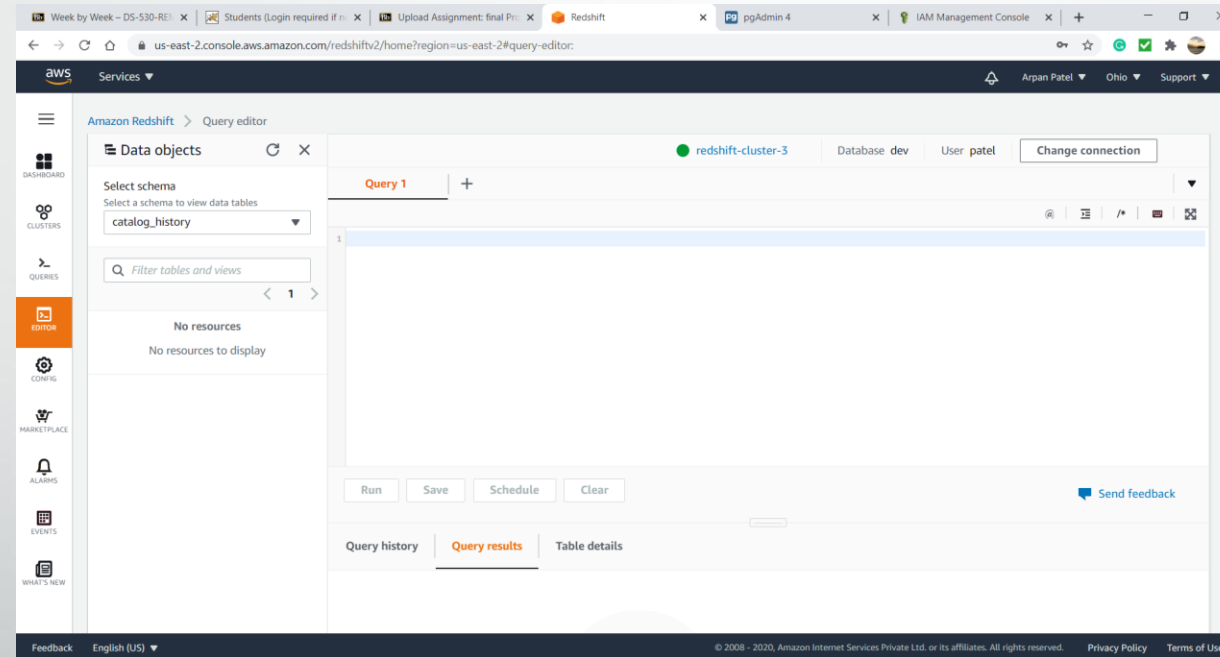
- Virtual private cloud (VPC):** `vpc-32862f59` with a link icon.
- Subnet:** `default` with a link icon.
- Availability Zone:** `us-east-2a`.
- Enhanced VPC routing:** Forces cluster traffic through a VPC.
- Disabled:** (Status for Enhanced VPC routing)
- VPC security group:** Specify which instances and devices can connect to the cluster. `sg-546c902b` with a link icon.
- Publicly accessible:** `Yes` (with an 'Edit' link).

The footer of the console shows the URL `https://us-east-2.console.aws.amazon.com/redshiftv2/home?region=us-east-2#` and copyright information: '© 2008 - 2020, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved.' along with links for 'Privacy Policy' and 'Terms of Use'.

Connect to database user in aws redshift for query editor



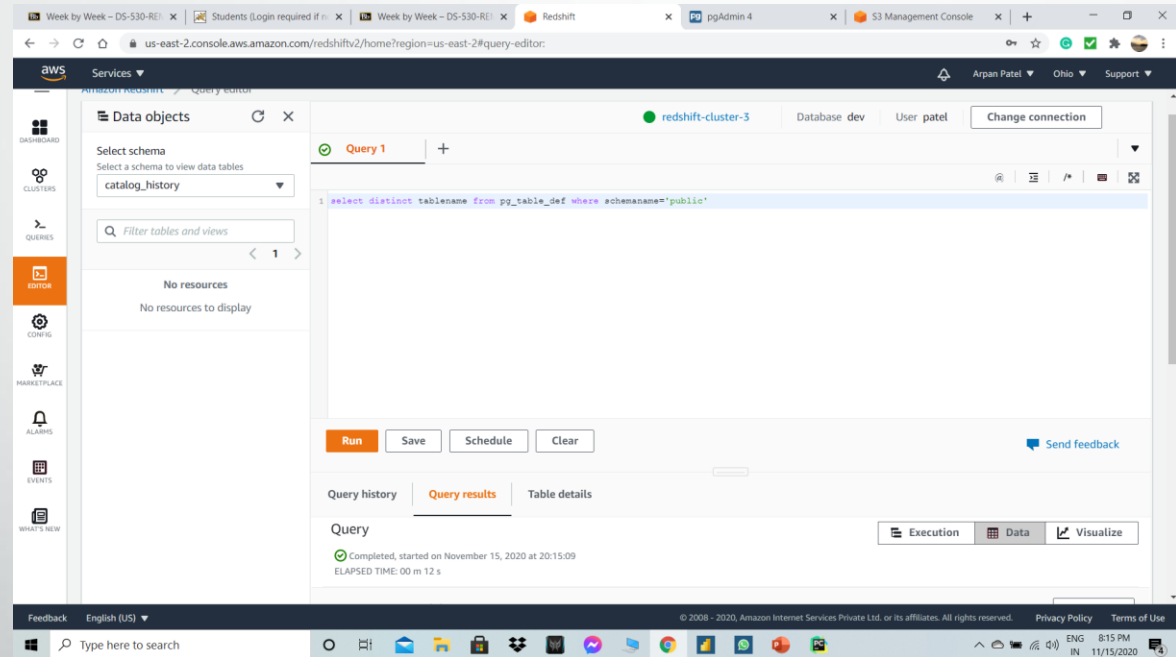
Redshift query editor



Query pycharm

```
import psycopg2
import pprint
#conn_string = "host='localhost' dbname='postgres' user='postgres' password='rootpassword'"
conn_string = "host='localhost' dbname='postgres' user='postgres' password='Arpan@1316'"
conn = psycopg2.connect(conn_string)
cursor = conn.cursor()
#cursor.execute("SELECT relname, relpages FROM pg_class ORDER BY relpages DESC")
cursor.execute("SELECT * FROM public.address")
records = cursor.fetchall()
pprint.pprint(records)
|
```

Sample Queries in redshift





THANK YOU