SalesAnalysis

By,

SantanuAdhikary(2320846)

E.DivyaSree(2320625)

A.N.V.Rohit(2320408)

SanmithChetanN(2320907)

ShashankK Hulmani(2320613)

Vivechan Gowda G.C(2320755)

cognizant

Contents

Introduction

Regardless of whether you are a small local business or a national corporation, tracking your retail sales is incredibly important. The valuable activity lets you identify potential sales patterns, close more deals in much less time, and make better and informed sales forecasts. However, it is seen that most retail sales managers struggle to keep a tap on their sales performance. This, in turn, impacts their retail productivity.

The amount of data gathered by retail businesses keeps growing at an alarmingly high rate. Sales managers are unable to ensure the adequate utilization and analysis of data, contributing to the issue of data silos. Besides, it results in the most common challenges for retailers: a lack of consistency in the implementation of the sales process, limited line-of-sight for managers, and reps spending too much time not selling.

It is crucial to find a technology solution that can effectively handle all the generated data and use it to unlock meaningful insights into sales performance for the retail industry. This is where **AWS Redshift** comes into the picture. In today's blog, we will understand how using Amazon Redshift can help to significantly **improve sales performance** and tracking.

A). Amazon Redshift:

Amazon Redshift is a fully managed data warehouse service provided by Amazon Web Services (AWS). It's designed for analyzing large datasets and offers scalability, high performance, and SQL compatibility. Redshift uses a distributed architecture with massively parallel processing (MPP), making it efficient for complex queries. Data is stored in a columnar format, which enhances query speed. Whether you're running business intelligence reports or analyzing historical data, Amazon Redshift is a powerful choice for data analytics and reporting.

B). Amazon S3 Storage:

Amazon S3 (Simple Storage Service) is an object storage service provided by Amazon Web Services (AWS). It offers industry-leading scalability, data availability, security, and performance. Here are the key points about Amazon S3:

- Scalability: Amazon S3 allows you to store and retrieve any amount of data from anywhere on the web. It's designed to handle both small and large-scale applications.

- Object Storage: Instead of traditional file storage, Amazon S3 stores data as objects within resources called buckets. These objects can be photos, audio files, videos, or any other type of data.
- Durability and Availability: Amazon S3 provides 99.999999999% durability (meaning your data is highly resistant to loss) and 99.99% availability of objects over a given year.

C). Quick Sight:

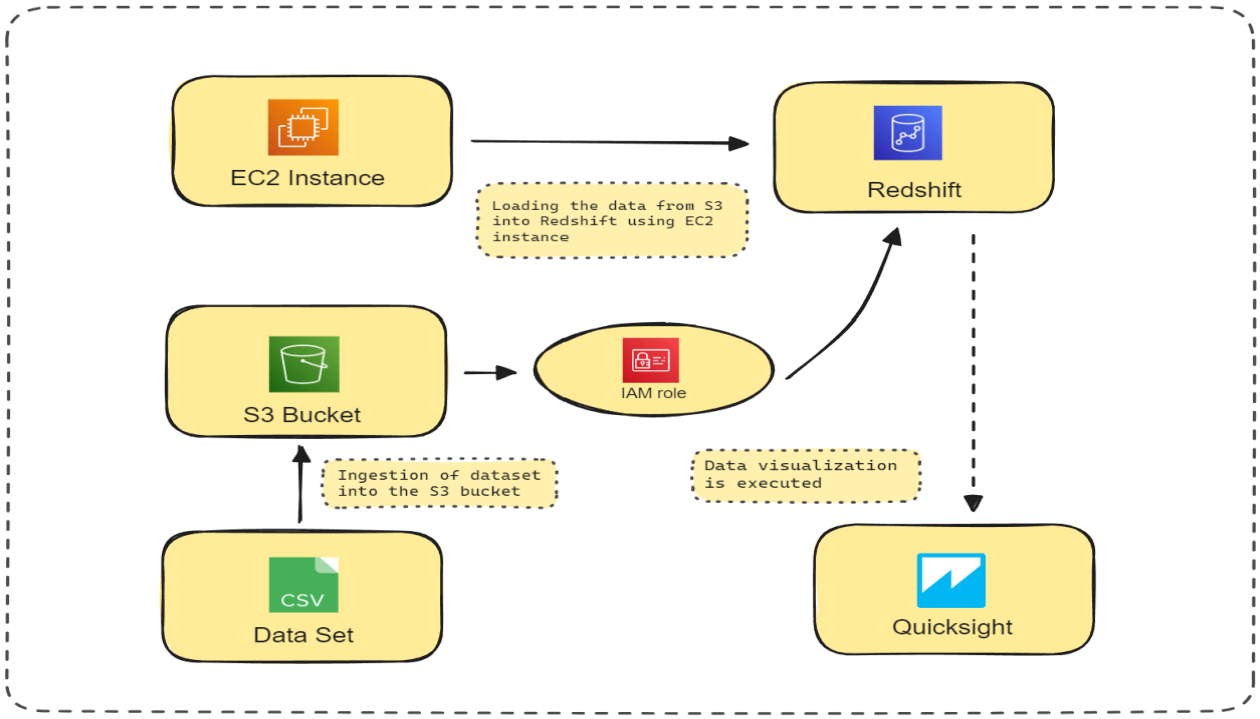
Amazon Quick Sight is a cloud-native, serverless business intelligence (BI) service provided by Amazon Web Services (AWS). Here are the key points about Quick Sight:

- Unified BI at Hyperscale: Quick Sight powers data-driven organizations with unified BI capabilities. It allows users to meet varying analytic needs from the same source of truth through:
- **Modern Interactive Dashboards**: Create interactive visualizations and explore data.
- **Paginated Reports**: Generate critical operational reports and dashboards.
    - **Embedded Analytics**: Embed dashboards or natural language query capabilities in applications.

Business Requirements

1. Import sales data into Amazon Redshift.
2. Create SQL queries to aggregate sales data by product, region, and time period.
3. Identify top-selling products by calculating total sales or units sold.
4. Determine highest revenue-generating regions by analyzing sales revenue per region.
5. Analyze trends over time by plotting sales data over different time periods (e.g., daily, monthly, quarterly).
6. Visualize the sales data using charts and graphs (e.g., bar charts, line graphs) to provide insights into sales performance.

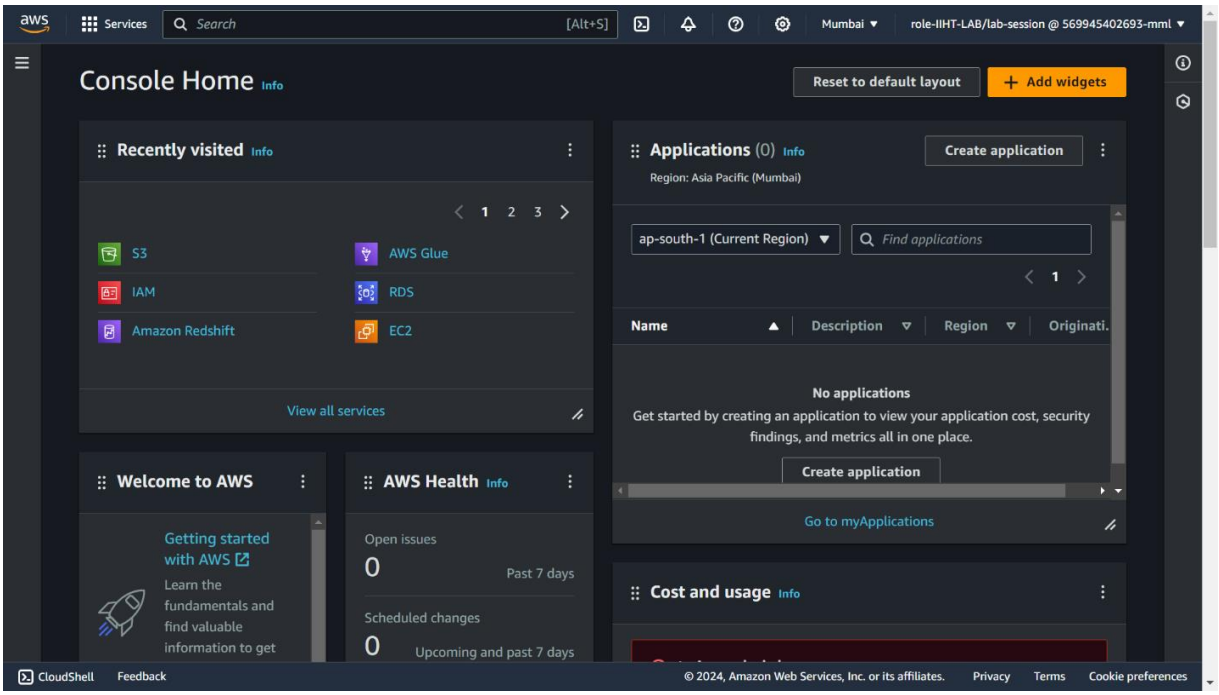
Overview of the Project



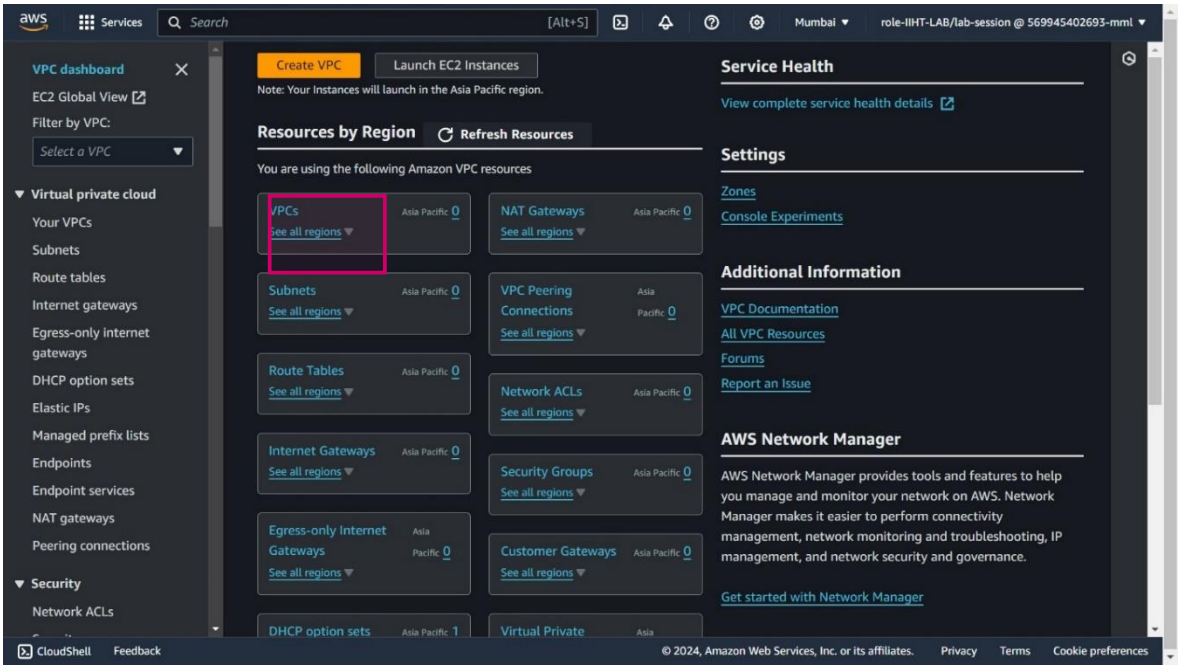
1	2	3	4	5	6
Fetching data from kaggle. And Loading the data to S3 bucket.	Creating IAM role for S3 Access from Redshift	Deploying a Redshift Cluster	Deploy EC2 instance ,connect to Redshift, Load the data from S3 to Redshift .	performing data analysis using Redshift Query Editor	Loading the data to Quick sight and identifying trends and patterns through visualization.

Implementation of Project

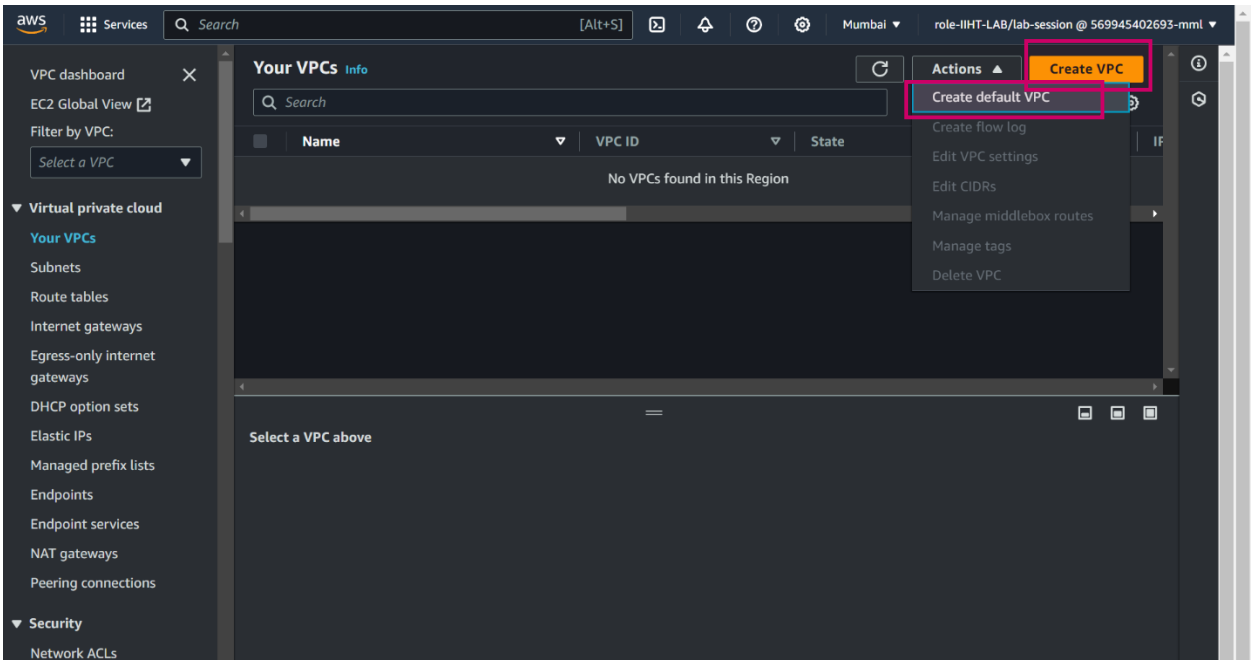
Step 1: Open AWS console

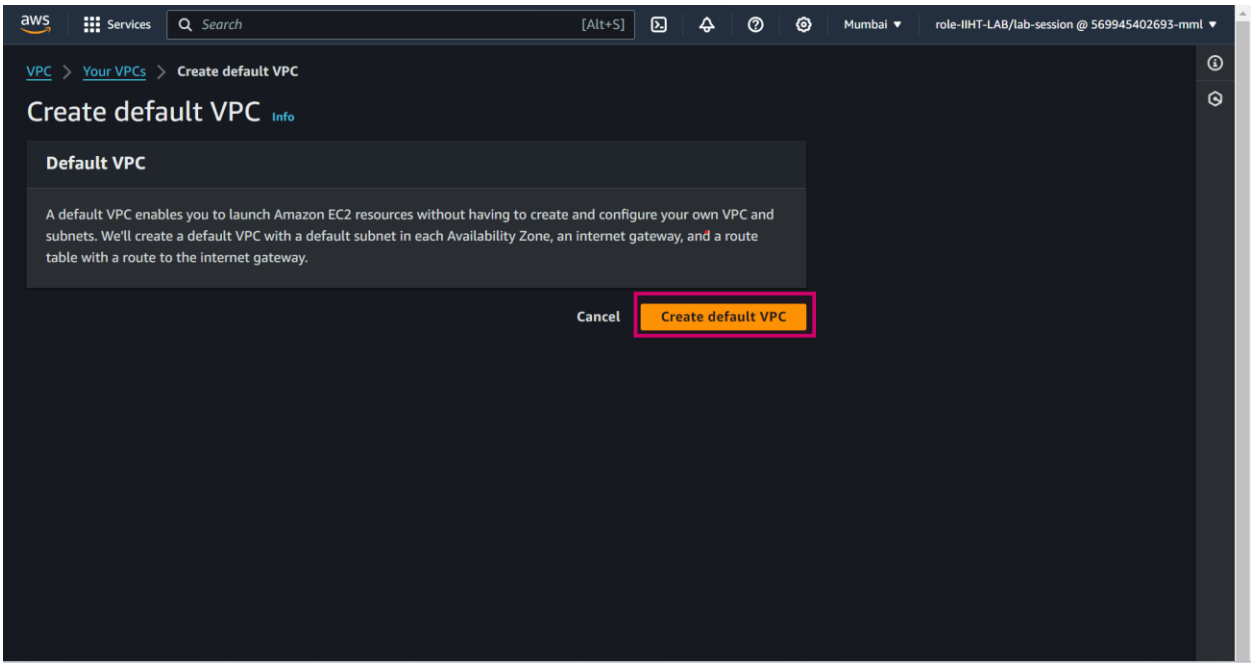


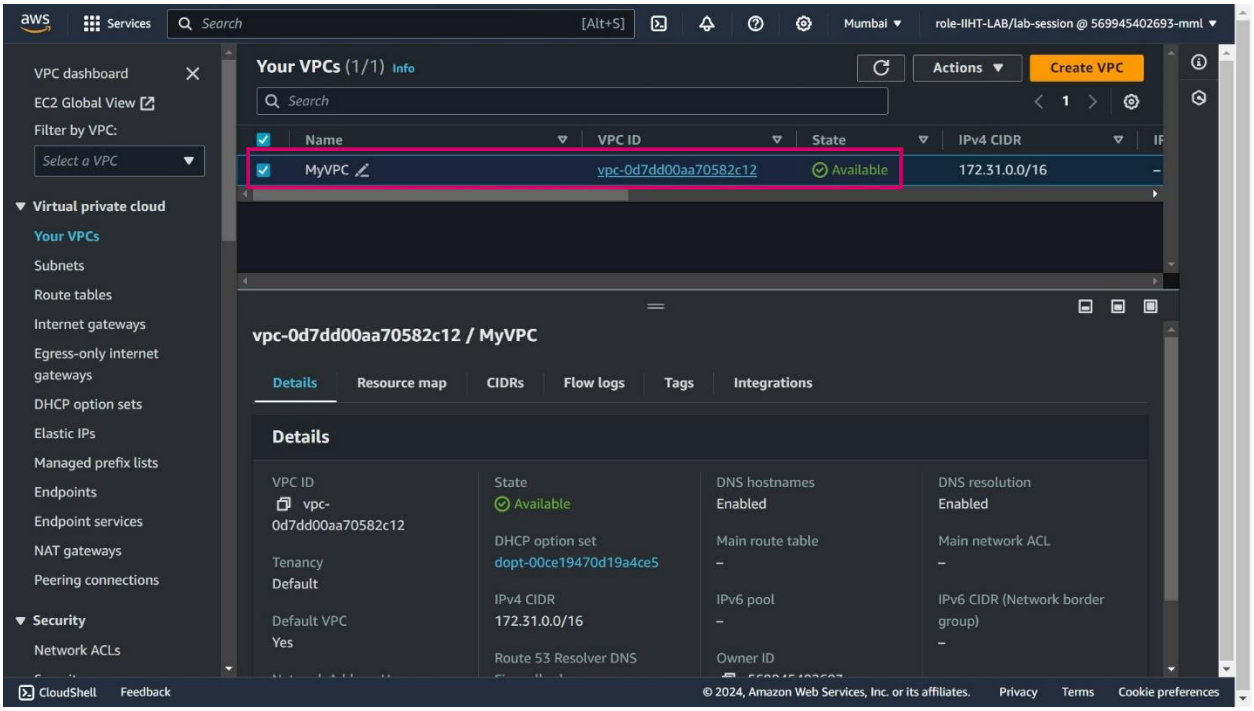
Step 2: Create VPC

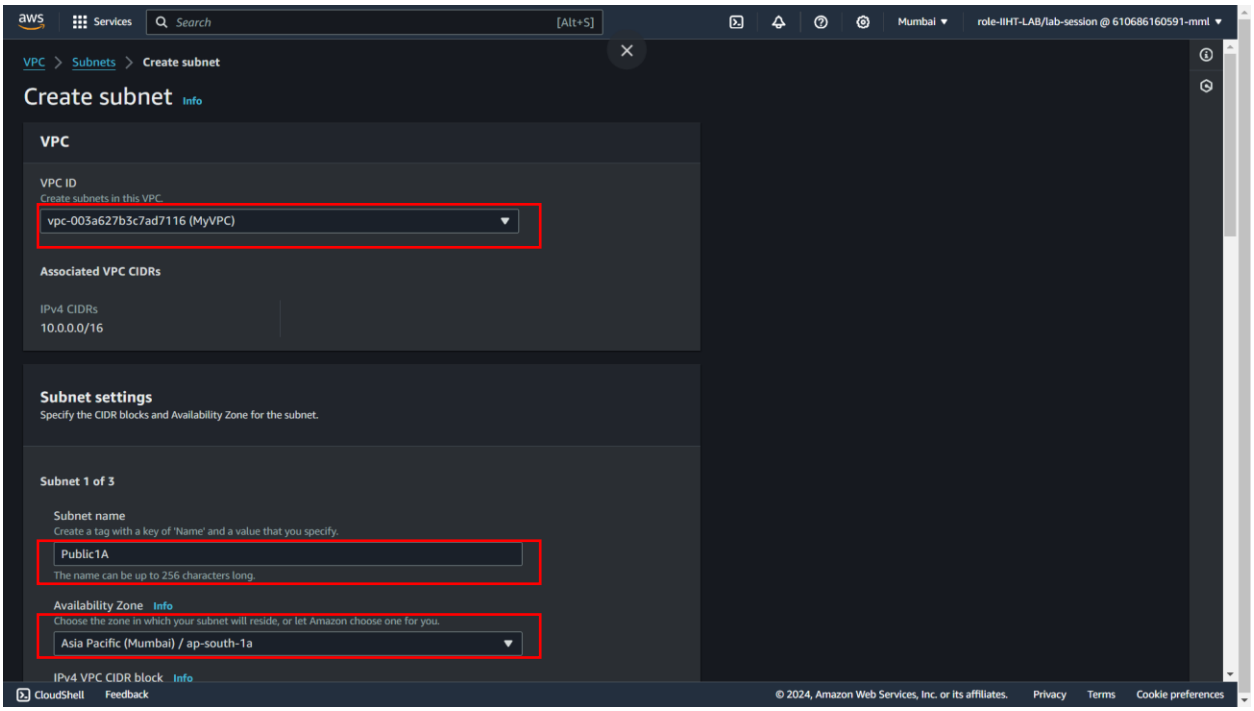


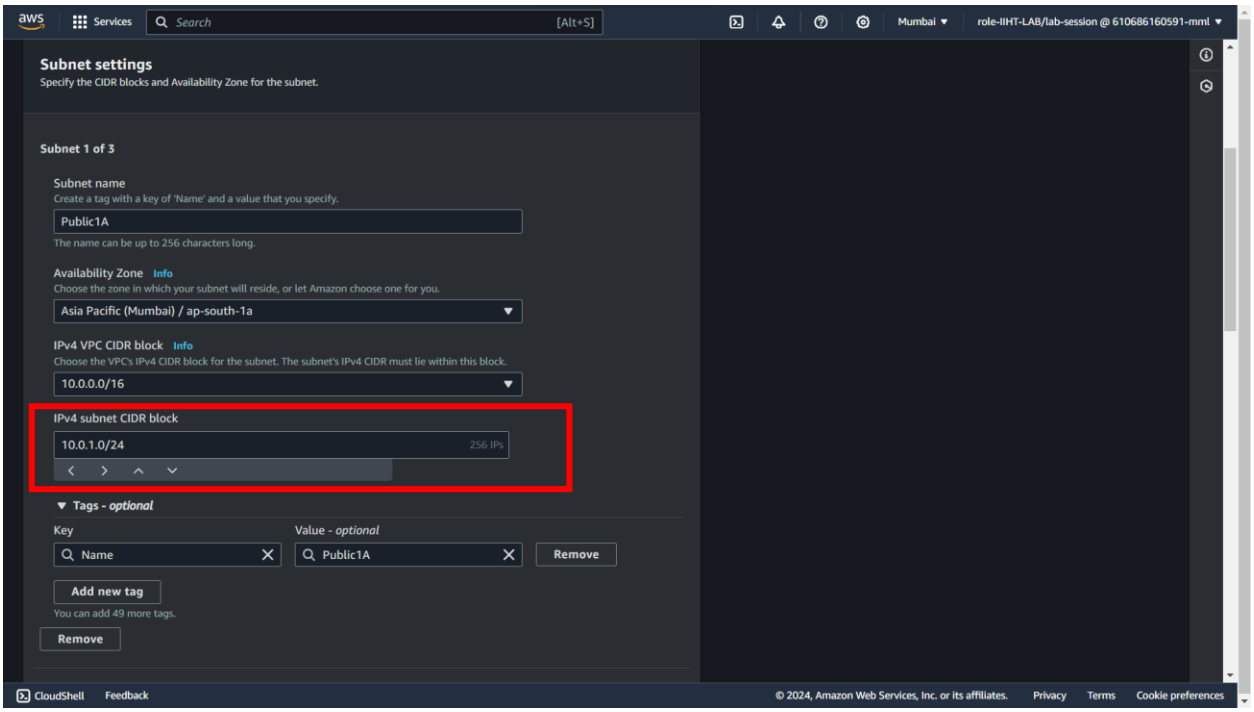
Open VPC and create default VPC and add subnets

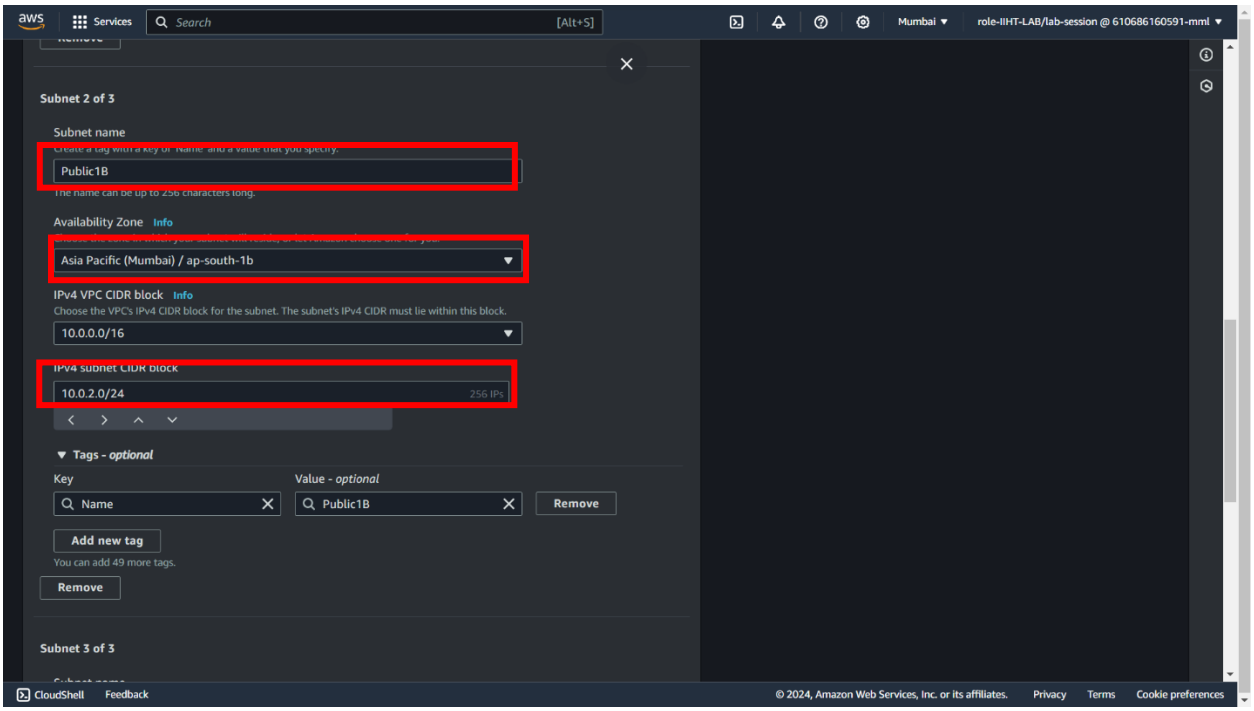


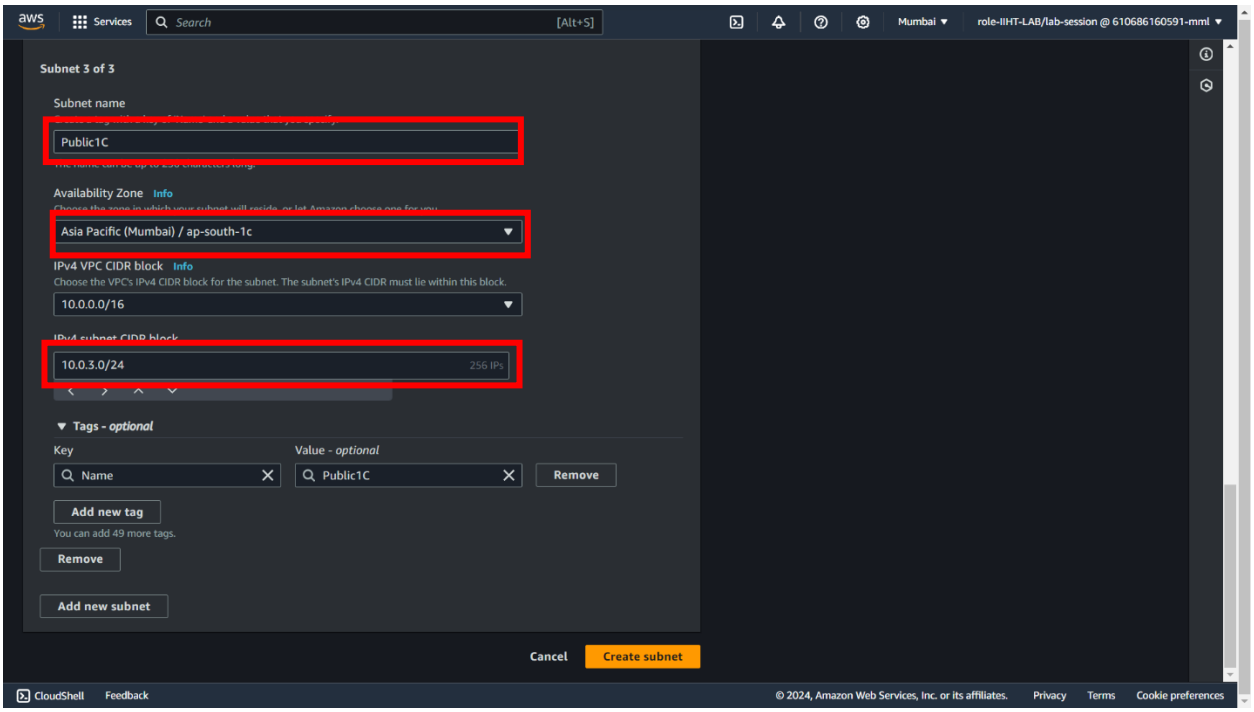


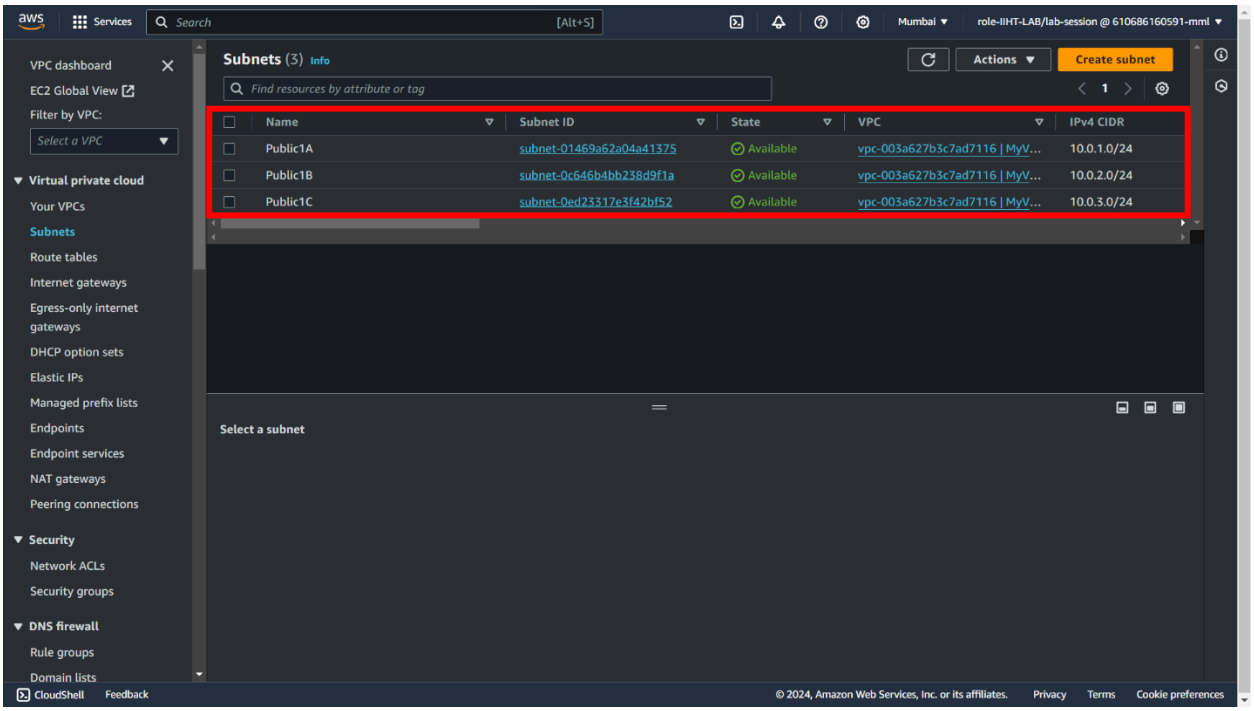

aws
Services
Search
[Alt+S]
Mumbai
role-IIHT-LAB/lab-session @ 569945402693-mml
VPC dashboard
EC2 Global View
Filter by VPC:
Select a VPC
Virtual private cloud
Your VPCs
Subnets
Route tables
Internet gateways
Egress-only internet gateways
DHCP option sets
Elastic IPs
Managed prefix lists
Endpoints
Endpoint services
NAT gateways
Peering connections
Security
Network ACLs
Your VPCs (1/1) Info
Actions
Create VPC
Search
Name
VPC ID
State
IPv4 CIDR
MyVPC
vpc-0d7dd00aa70582c12
Available
172.31.0.0/16
vpc-0d7dd00aa70582c12 / MyVPC
Details
Resource map
CIDRs
Flow logs
Tags
Integrations
Details
VPC ID
vpc-0d7dd00aa70582c12
State
Available
DNS hostnames
Enabled
DNS resolution
Enabled
Tenancy
Default
DHCP option set
dopt-00ce19470d19a4ce5
Main route table
Main network ACL
Default VPC
Yes
IPv4 CIDR
172.31.0.0/16
IPv6 pool
IPv6 CIDR (Network border group)
Route 53 Resolver DNS
Owner ID
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


Services
Search
[Alt+S]
Mumbai
role-IIHT-LAB/lab-session @ 610686160591-mml
VPC > Subnets > Create subnet
Create subnet Info
VPC
VPC ID
Create subnets in this VPC.
vpc-003a627b3c7ad7116 (MyVPC)
Associated VPC CIDRs
IPv4 CIDRs
10.0.0.0/16
Subnet settings
Specify the CIDR blocks and Availability Zone for the subnet.
Subnet 1 of 3
Subnet name
Create a tag with a key of 'Name' and a value that you specify.
Public1A
The name can be up to 256 characters long.
Availability Zone Info
Choose the zone in which your subnet will reside, or let Amazon choose one for you.
Asia Pacific (Mumbai) / ap-south-1a
IPv4 VPC CIDR block Info
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


Services
Search
[Alt+S]
Mumbai
role-IIHT-LAB/lab-session @ 610686160591-mml
Subnet settings
Specify the CIDR blocks and Availability Zone for the subnet.
Subnet 1 of 3
Subnet name
Create a tag with a key of 'Name' and a value that you specify.
Public1A
The name can be up to 256 characters long.
Availability Zone Info
Choose the zone in which your subnet will reside, or let Amazon choose one for you.
Asia Pacific (Mumbai) / ap-south-1a
IPv4 VPC CIDR block Info
Choose the VPC's IPv4 CIDR block for the subnet. The subnet's IPv4 CIDR must lie within this block.
10.0.0.0/16
IPv4 subnet CIDR block
10.0.1.0/24
256 IPs
Tags - optional
Key
Value - optional
Name
Public1A
Remove
Add new tag
You can add 49 more tags.
Remove
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

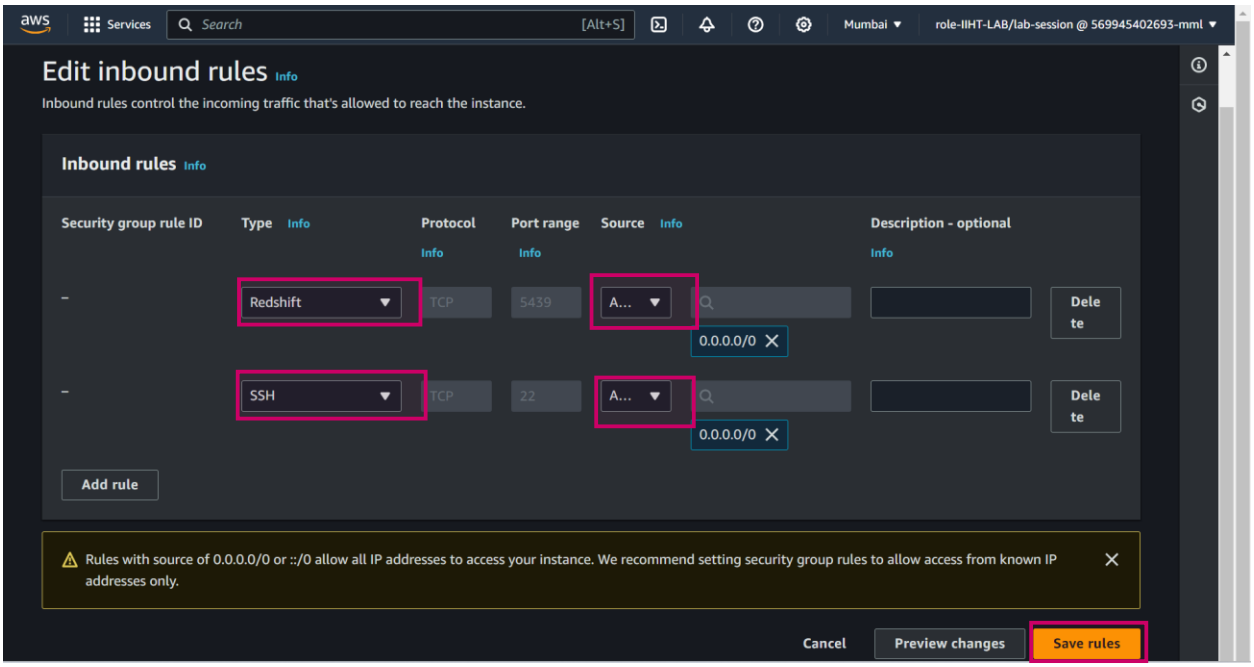

Services
Search
[Alt+S]
Mumbai
role-IIHT-LAB/lab-session @ 610686160591-mml
Subnet 2 of 3
Subnet name
Public1B
The name can be up to 256 characters long.
Availability Zone Info
Asia Pacific (Mumbai) / ap-south-1b
IPv4 VPC CIDR block Info
Choose the VPC's IPv4 CIDR block for the subnet. The subnet's IPv4 CIDR must lie within this block.
10.0.0.0/16
IPv4 subnet CIDR block
10.0.2.0/24
256 IPs
Tags - optional
Key
Value - optional
Name
Public1B
Remove
Add new tag
You can add 49 more tags.
Remove
Subnet 3 of 3
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

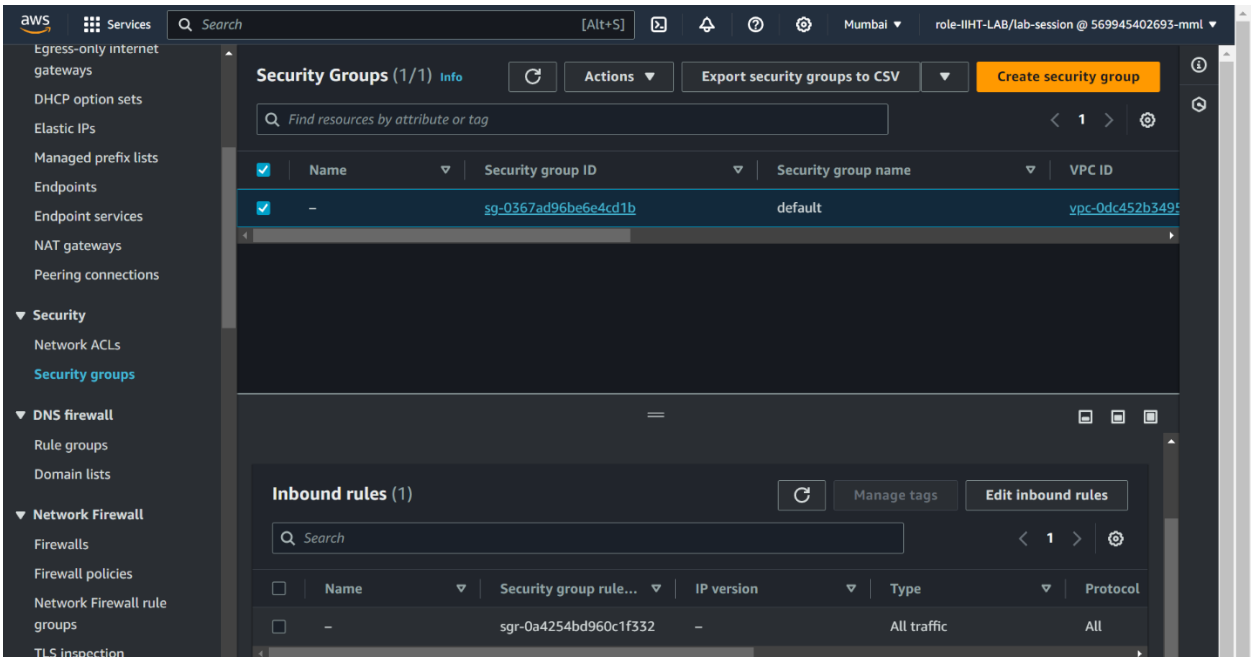

Services
Search
[Alt+S]
Mumbai
role-IIHT-LAB/lab-session @ 610686160591-mml
Subnet 3 of 3
Subnet name
Public1C
Availability Zone Info
Asia Pacific (Mumbai) / ap-south-1c
IPv4 VPC CIDR block Info
Choose the VPC's IPv4 CIDR block for the subnet. The subnet's IPv4 CIDR must lie within this block.
10.0.0.0/16
IPv4 subnet CIDR block
10.0.3.0/24
256 IPs
Tags - optional
Key
Value - optional
Name
Public1C
Remove
Add new tag
You can add 49 more tags.
Remove
Add new subnet
Cancel
Create subnet
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences



Step3: Creating Security Groups

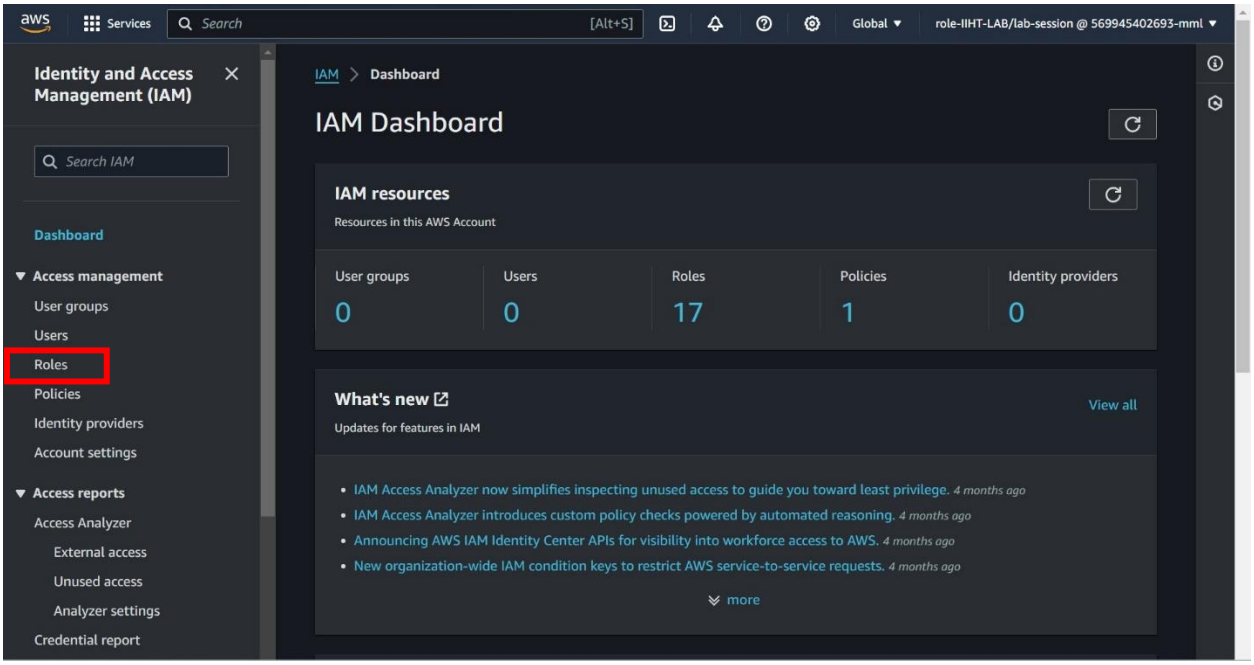
There is a default security group we need to change the inbound rules to give access to redshift from S3

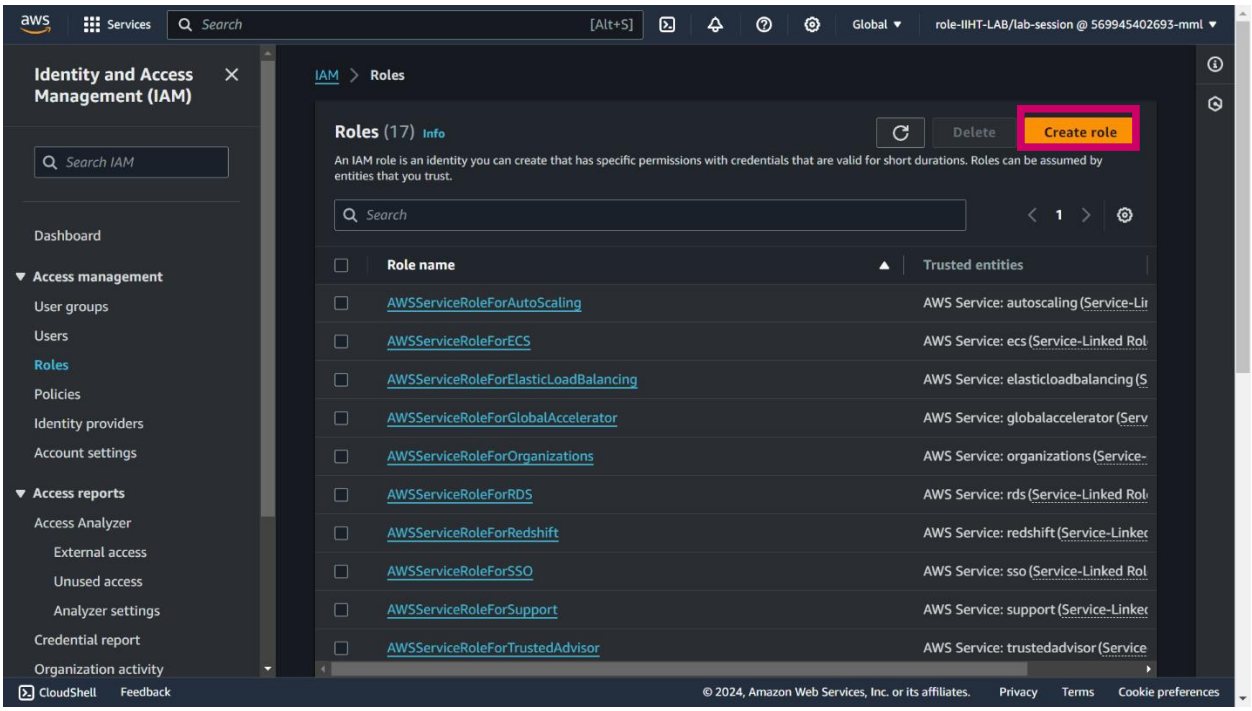


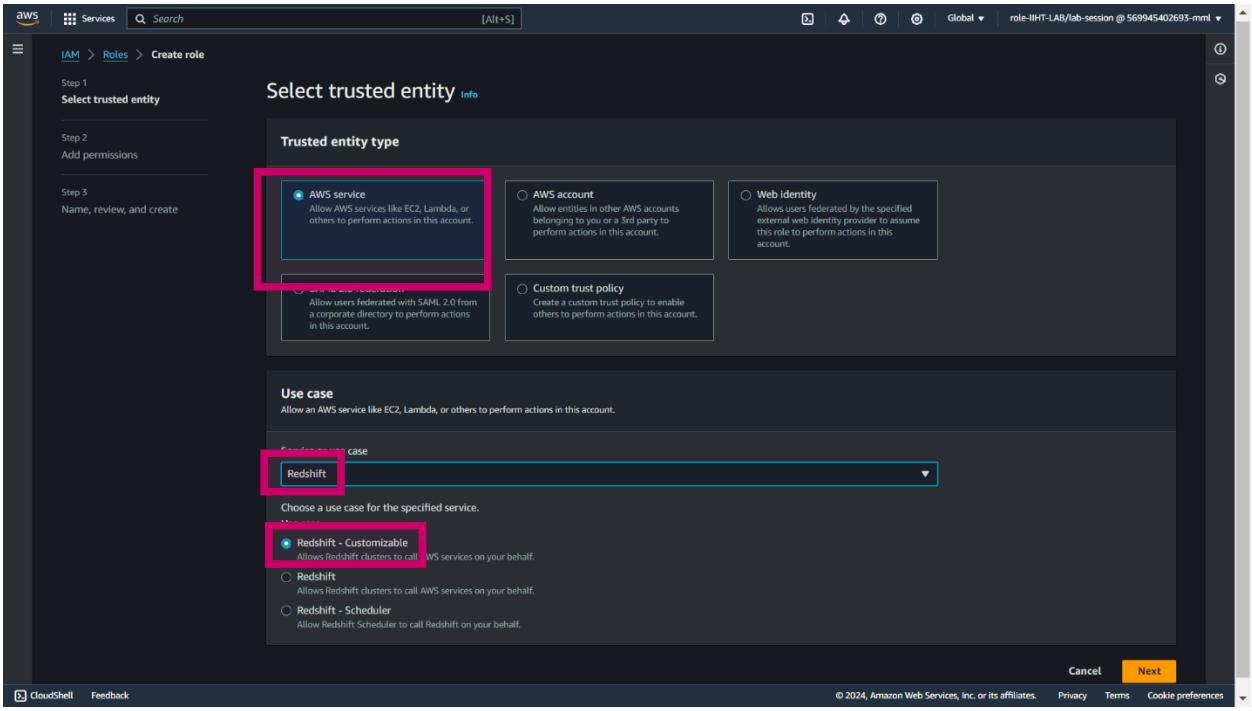


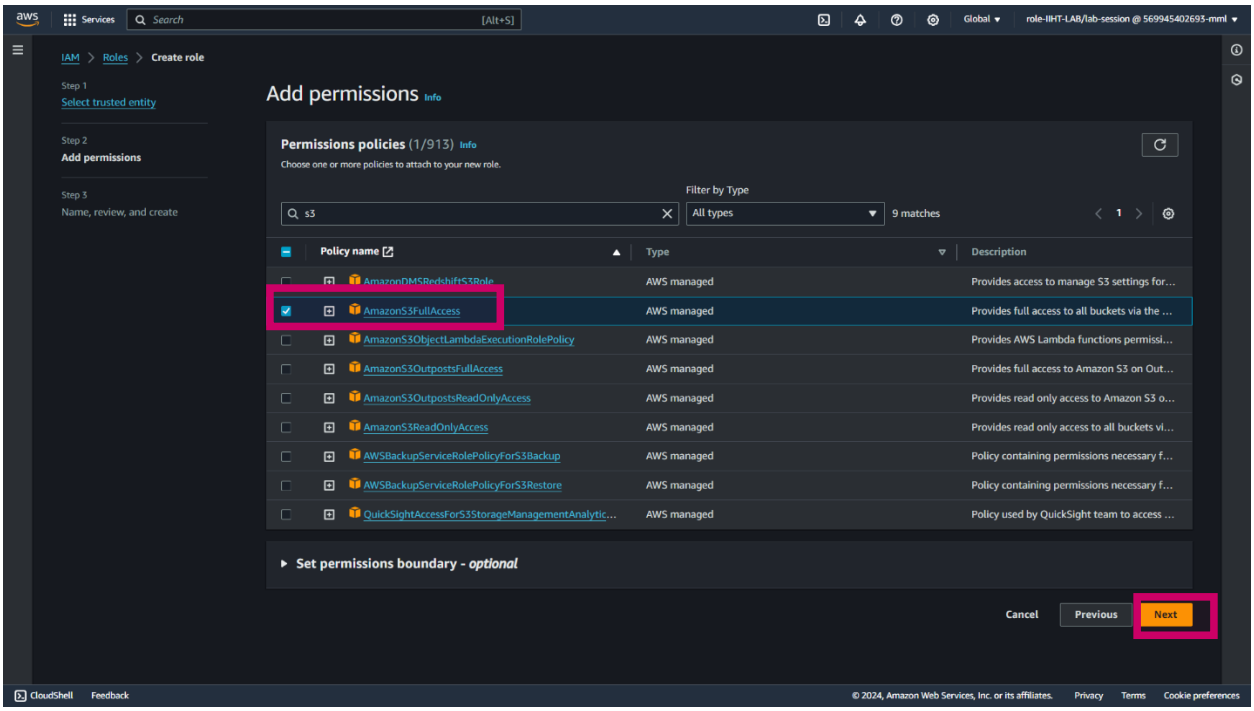
Step 4 : Open IAM Dashboard to create role for Redshift to access the S3

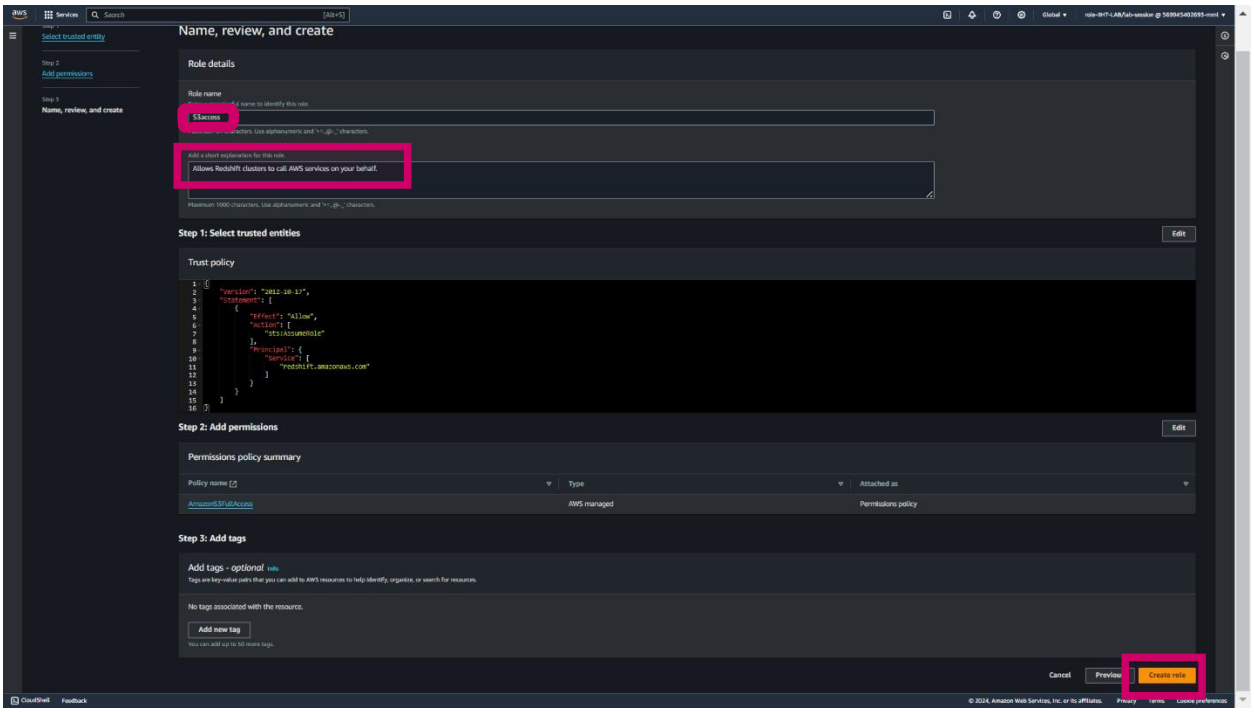
It is used to securely access the AWS services and resources. It helps to enhance the security by ensuring that only authorized individuals or system can interact with our AWS.

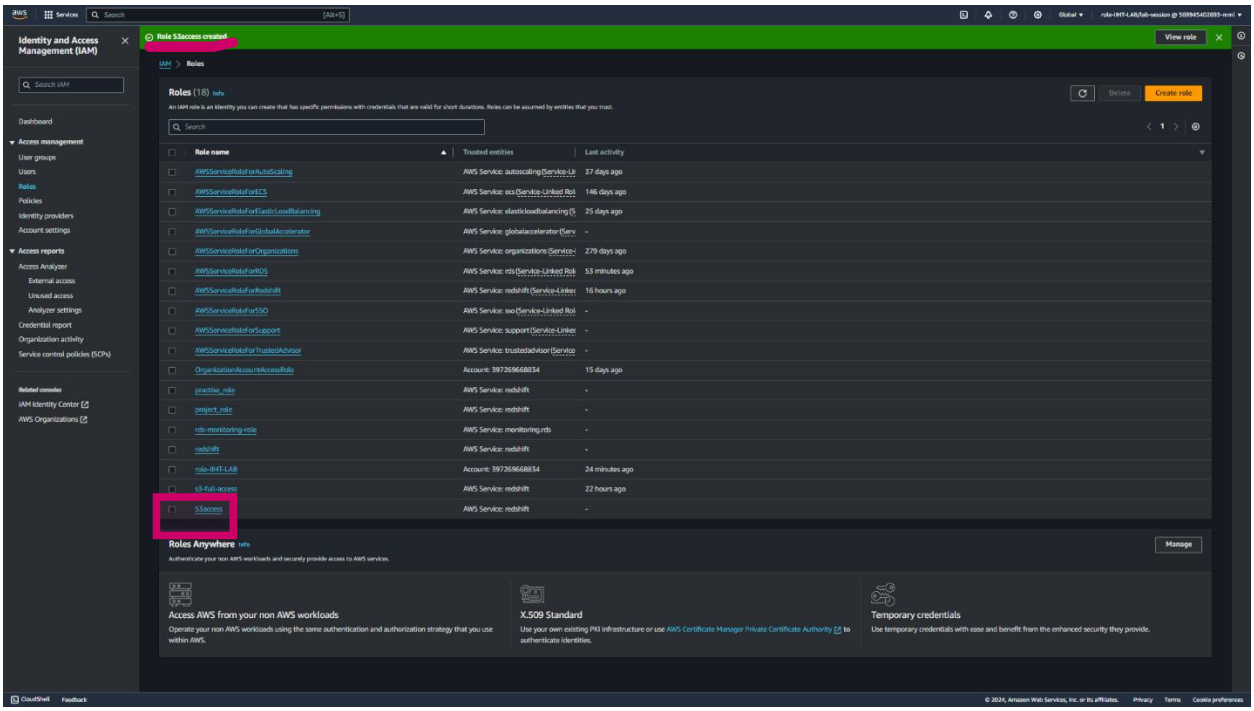



Identity and Access Management (IAM)
Search IAM
Dashboard
Access management
User groups
Users
Roles
Policies
Identity providers
Account settings
Access reports
Access Analyzer
External access
Unused access
Analyzer settings
Credential report
Organization activity
IAM > Roles
Roles (17) Info
Delete
Create role
An IAM role is an identity you can create that has specific permissions with credentials that are valid for short durations. Roles can be assumed by entities that you trust.
Search
Role name
Trusted entities
AWSServiceRoleForAutoScaling
AWS Service: autoscaling (Service-Li
AWSServiceRoleForECS
AWS Service: ecs (Service-Linked Rol
AWSServiceRoleForElasticLoadBalancing
AWS Service: elasticloadbalancing (S
AWSServiceRoleForGlobalAccelerator
AWS Service: globalaccelerator (Serv
AWSServiceRoleForOrganizations
AWS Service: organizations (Service-
AWSServiceRoleForRDS
AWS Service: rds (Service-Linked Rol
AWSServiceRoleForRedshift
AWS Service: redshift (Service-Linkec
AWSServiceRoleForSSO
AWS Service: sso (Service-Linked Rol
AWSServiceRoleForSupport
AWS Service: support (Service-Linkec
AWSServiceRoleForTrustedAdvisor
AWS Service: trustedadvisor (Service
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

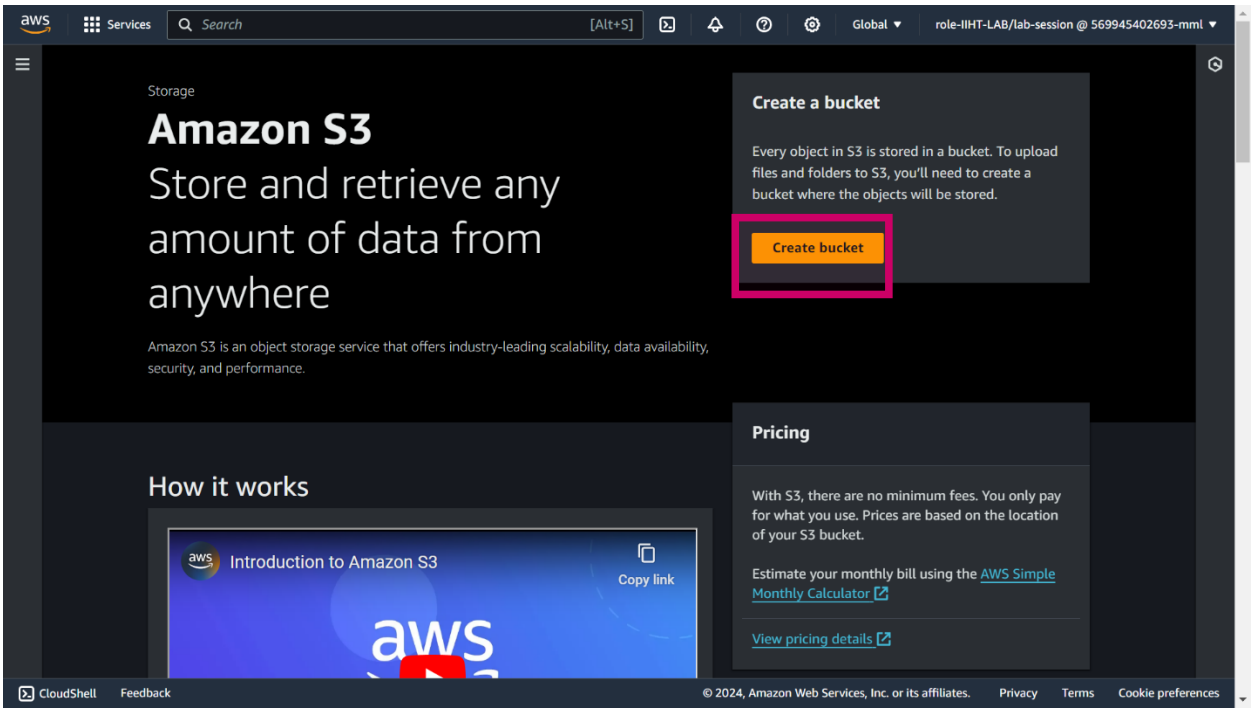

IAM > Roles > Create role
Step 1
Select trusted entity
Step 2
Add permissions
Step 3
Name, review, and create
Select trusted entity Info
Trusted entity type
AWS service
Allow AWS services like EC2, Lambda, or others to perform actions in this account.
AWS account
Allow entities in other AWS accounts belonging to you or a 3rd party to perform actions in this account.
Web identity
Allows users federated by the specified external web identity provider to assume this role to perform actions in this account.
Custom trust policy
Create a custom trust policy to enable others to perform actions in this account.
Use case
Allow an AWS service like EC2, Lambda, or others to perform actions in this account.
Redshift
Choose a use case for the specified service.
Redshift - Customizable
Redshift
Allows Redshift clusters to call AWS services on your behalf.
Redshift - Scheduler
Allow Redshift Scheduler to call Redshift on your behalf.
Cancel
Next
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

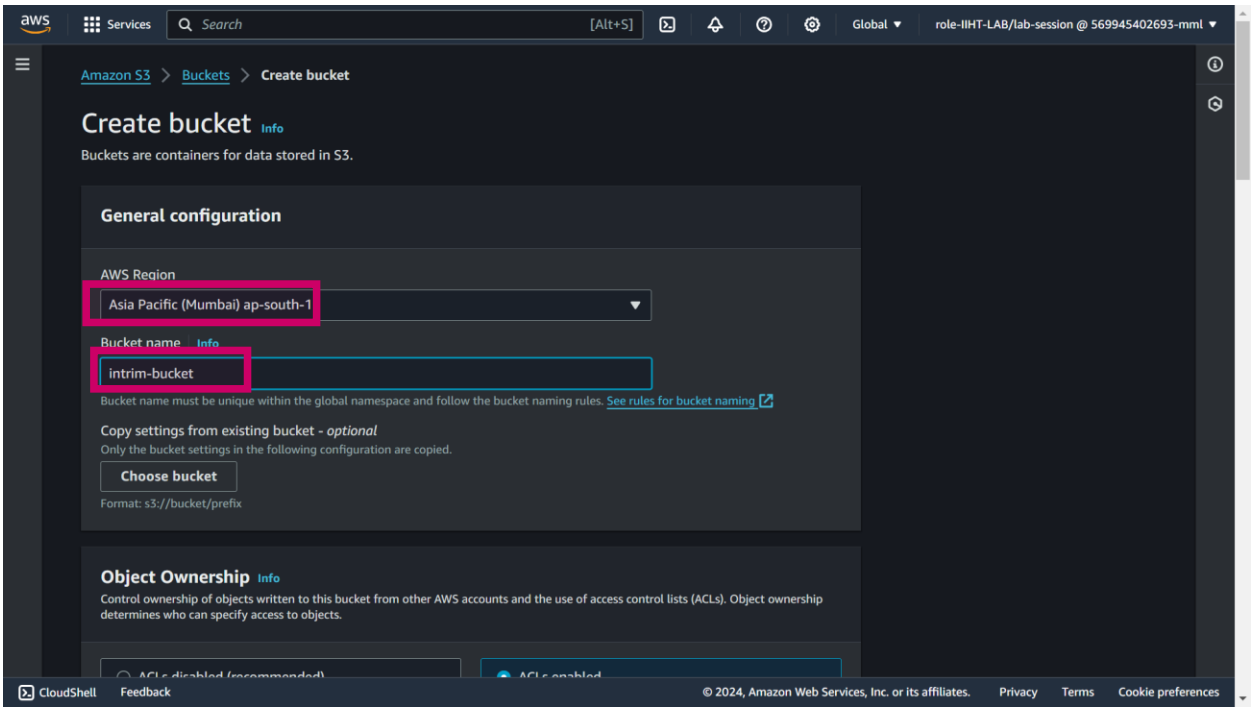

Services
Search
[Alt+S]
Global
role-IIHT-LAB/lab-session @ 569945402693-mml
IAM > Roles > Create role
Step 1
Select trusted entity
Step 2
Add permissions
Step 3
Name, review, and create
Add permissions Info
Permissions policies (1/913) Info
Choose one or more policies to attach to your new role.
Filter by Type
s3
All types
9 matches
Policy name
Type
Description
AmazonDMSRedshiftS3Role
AWS managed
Provides access to manage S3 settings for...
AmazonS3FullAccess
AWS managed
Provides full access to all buckets via the ...
AmazonS3ObjectLambdaExecutionRolePolicy
AWS managed
Provides AWS Lambda functions permissi...
AmazonS3OutpostsFullAccess
AWS managed
Provides full access to Amazon S3 on Out...
AmazonS3OutpostsReadOnlyAccess
AWS managed
Provides read only access to Amazon S3 o...
AmazonS3ReadOnlyAccess
AWS managed
Provides read only access to all buckets vi...
AWSBackupServiceRolePolicyForS3Backup
AWS managed
Policy containing permissions necessary f...
AWSBackupServiceRolePolicyForS3Restore
AWS managed
Policy containing permissions necessary f...
QuickSightAccessForS3StorageManagementAnalytic...
AWS managed
Policy used by QuickSight team to access ...
Set permissions boundary - optional
Cancel
Previous
Next
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

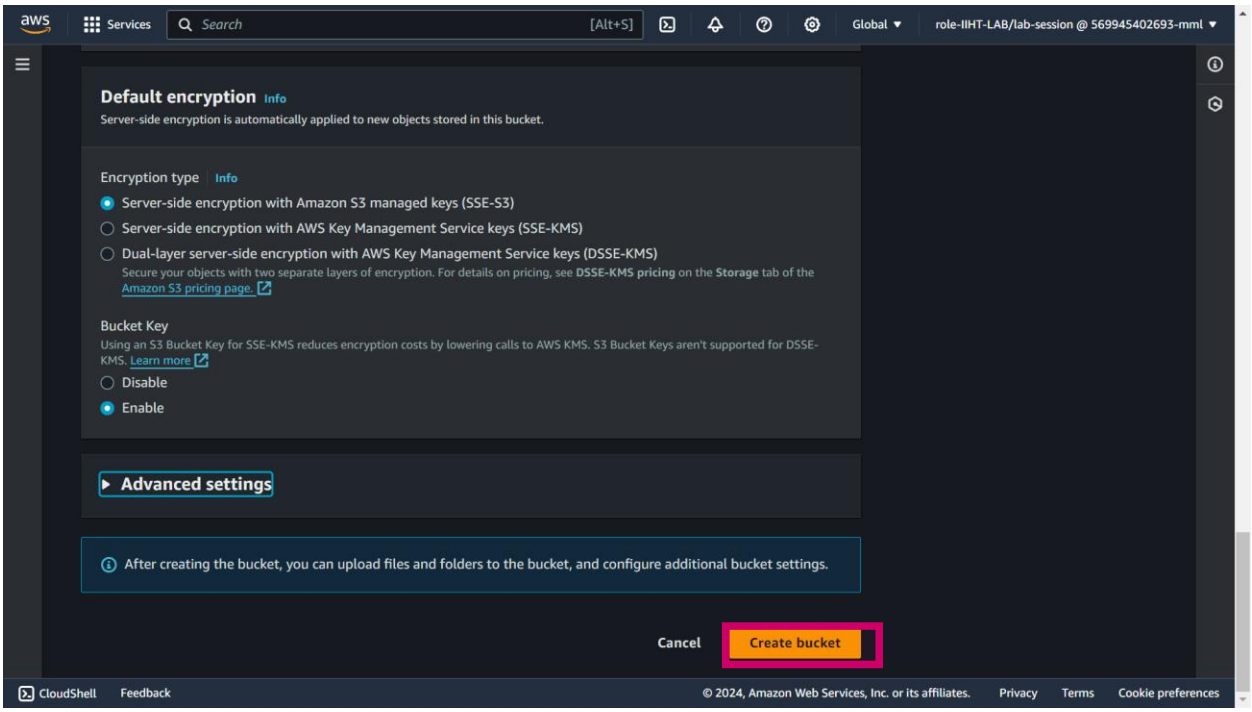

Name, review, and create
Role details
Role name
S3access
Allows Redshift clusters to call AWS services on your behalf.
Step 1: Select trusted entities
Edit
Trust policy
{
"Version": "2012-10-17",
"Statement": [
{
"Effect": "Allow",
"Action": [
"sts:AssumeRole"
],
"Principal": {
"Service": [
"redshift.amazonaws.com"
]
}
}
]
}
Step 2: Add permissions
Edit
Permissions policy summary
Policy name
Type
Attached as
AmazonS3FullAccess
AWS managed
Permissions policy
Step 3: Add tags
Add tags - optional
No tags associated with the resource.
Add new tag
Cancel
Create role

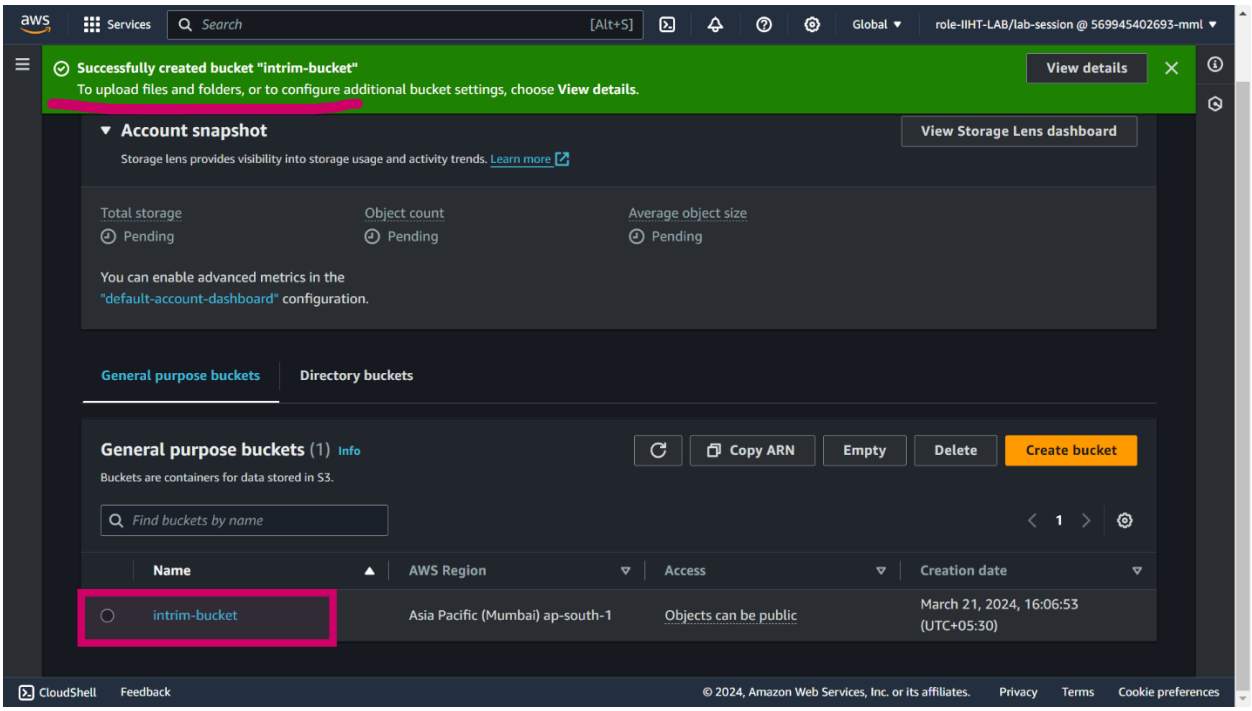


Step 5: Amazon S3 Bucket creation

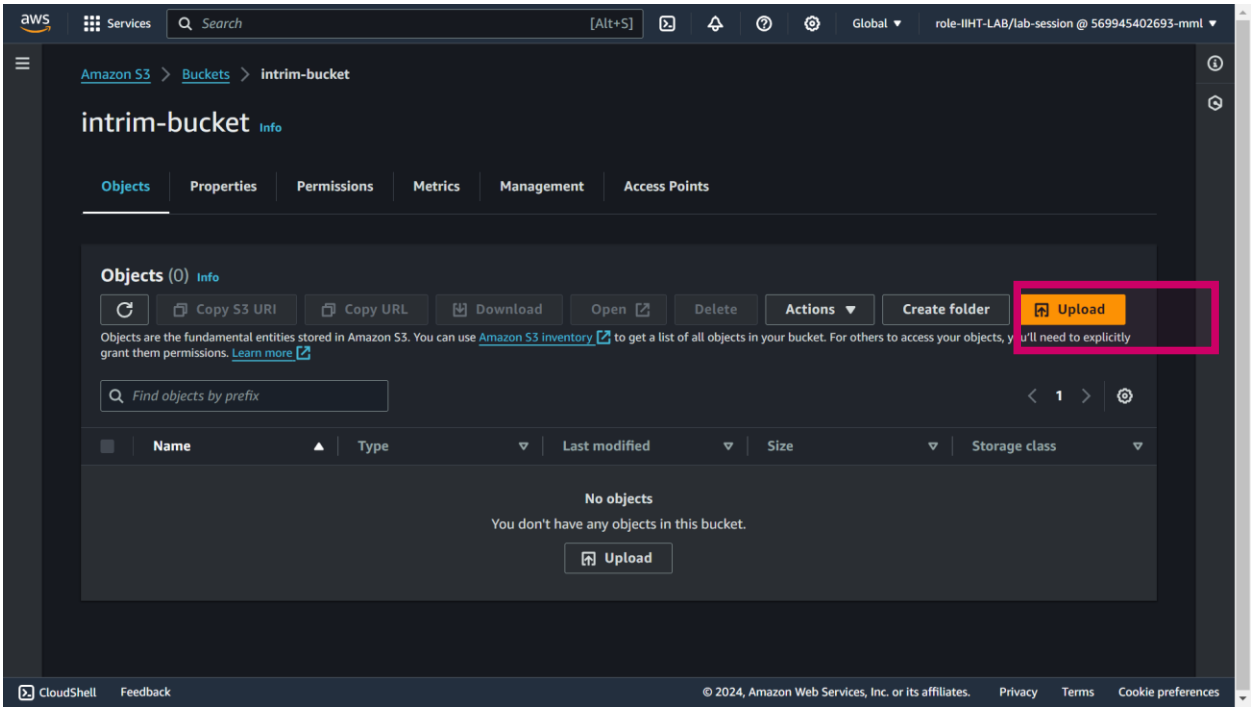


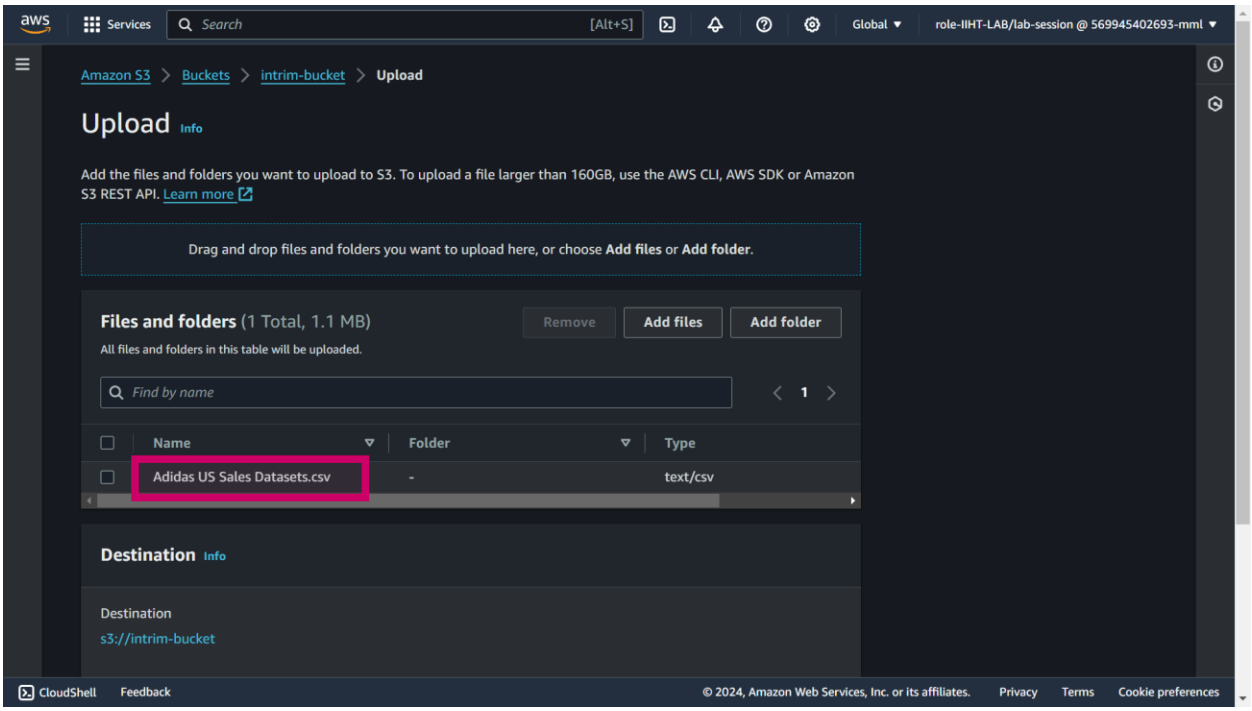

Services
Search
[Alt+S]
Global
role-IIHT-LAB/lab-session @ 569945402693-mml
Amazon S3 > Buckets > Create bucket
Create bucket Info
Buckets are containers for data stored in S3.
General configuration
AWS Region
Asia Pacific (Mumbai) ap-south-1
Bucket name Info
intrim-bucket
Bucket name must be unique within the global namespace and follow the bucket naming rules. See rules for bucket naming
Copy settings from existing bucket - optional
Only the bucket settings in the following configuration are copied.
Choose bucket
Format: s3://bucket/prefix
Object Ownership Info
Control ownership of objects written to this bucket from other AWS accounts and the use of access control lists (ACLs). Object ownership determines who can specify access to objects.
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

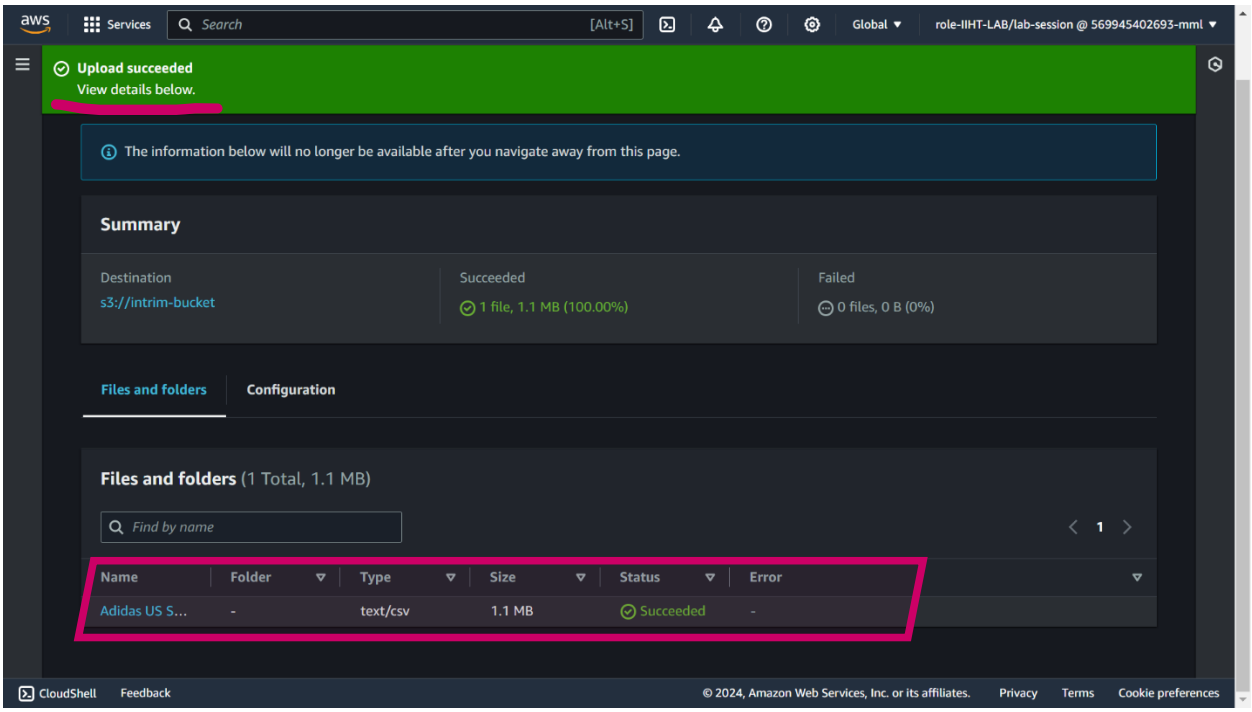

Services
Search
[Alt+S]
Global
role-IIHT-LAB/lab-session @ 569945402693-mml
Default encryption Info
Server-side encryption is automatically applied to new objects stored in this bucket.
Encryption type Info
Server-side encryption with Amazon S3 managed keys (SSE-S3)
Server-side encryption with AWS Key Management Service keys (SSE-KMS)
Dual-layer server-side encryption with AWS Key Management Service keys (DSSE-KMS)
Secure your objects with two separate layers of encryption. For details on pricing, see DSSE-KMS pricing on the Storage tab of the Amazon S3 pricing page.
Bucket Key
Using an S3 Bucket Key for SSE-KMS reduces encryption costs by lowering calls to AWS KMS. S3 Bucket Keys aren't supported for DSSE-KMS. Learn more
Disable
Enable
Advanced settings
After creating the bucket, you can upload files and folders to the bucket, and configure additional bucket settings.
Cancel
Create bucket
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences



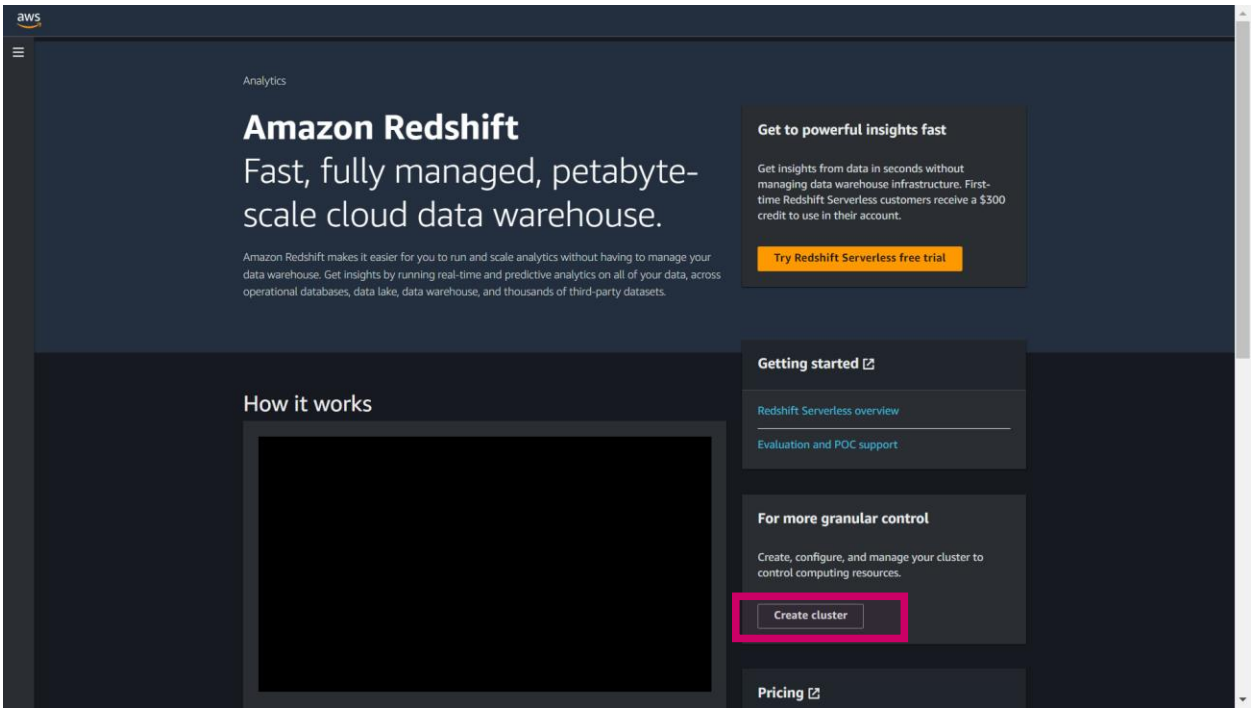
Step6: Uploading .csv file into S3 bucket

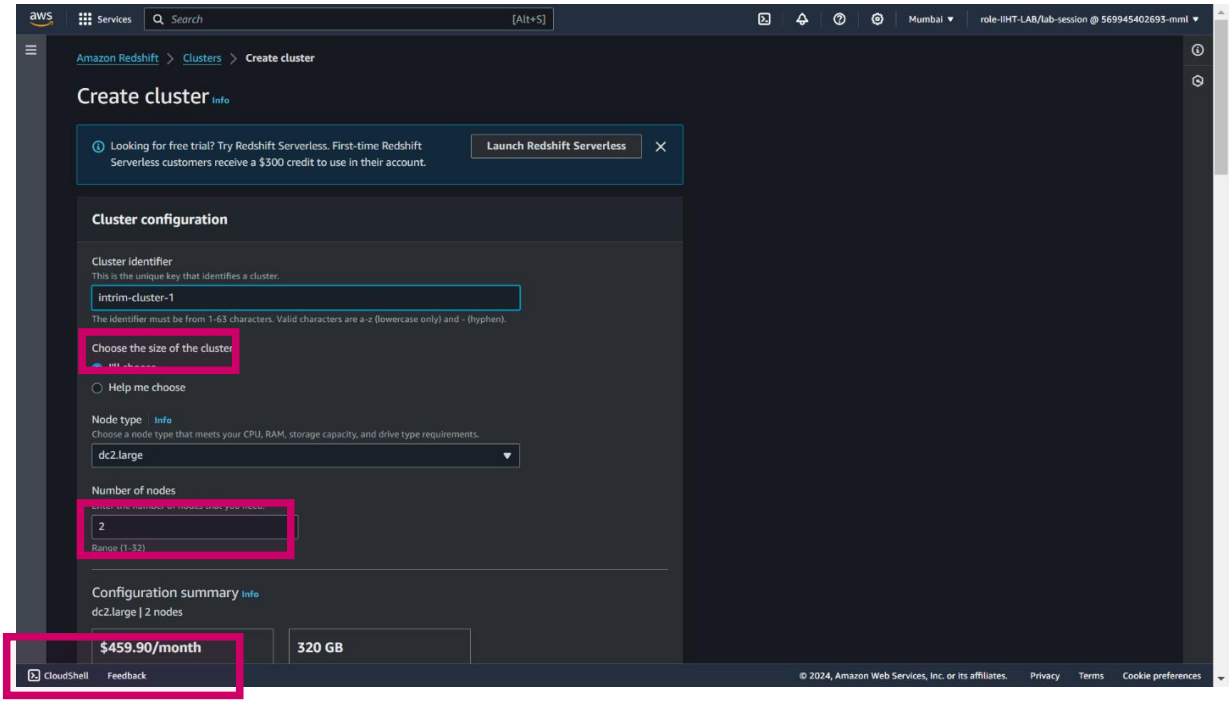




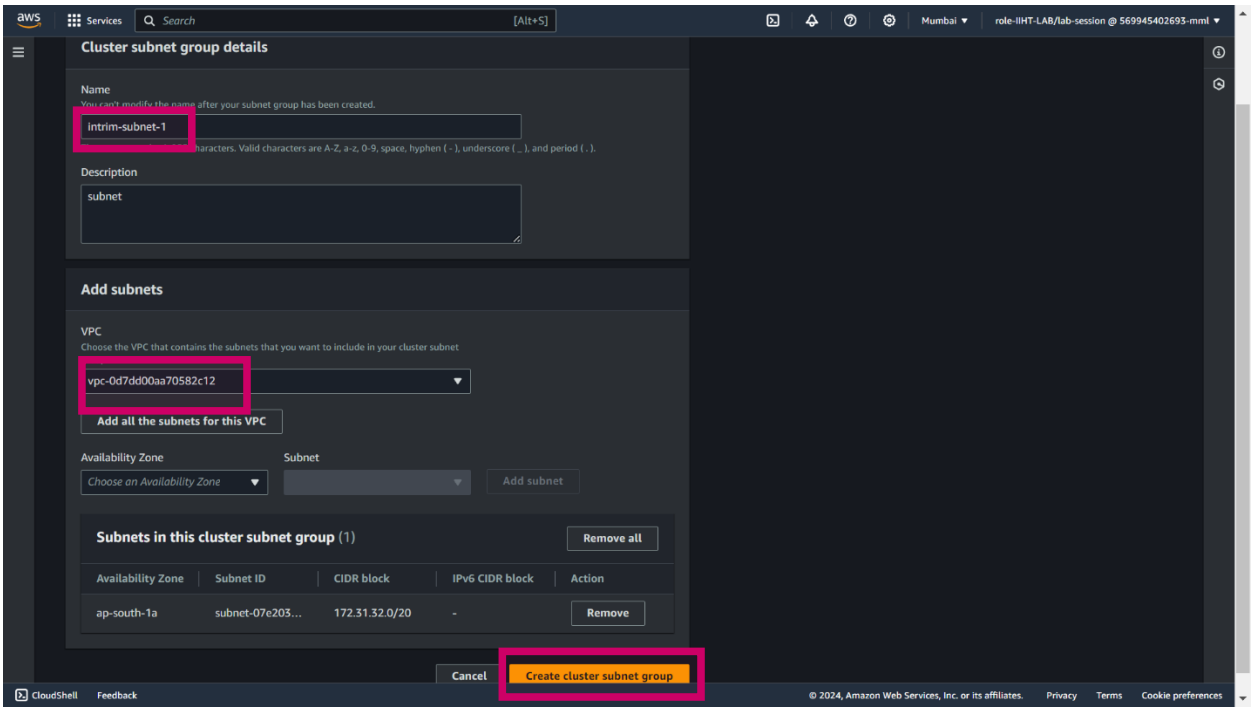


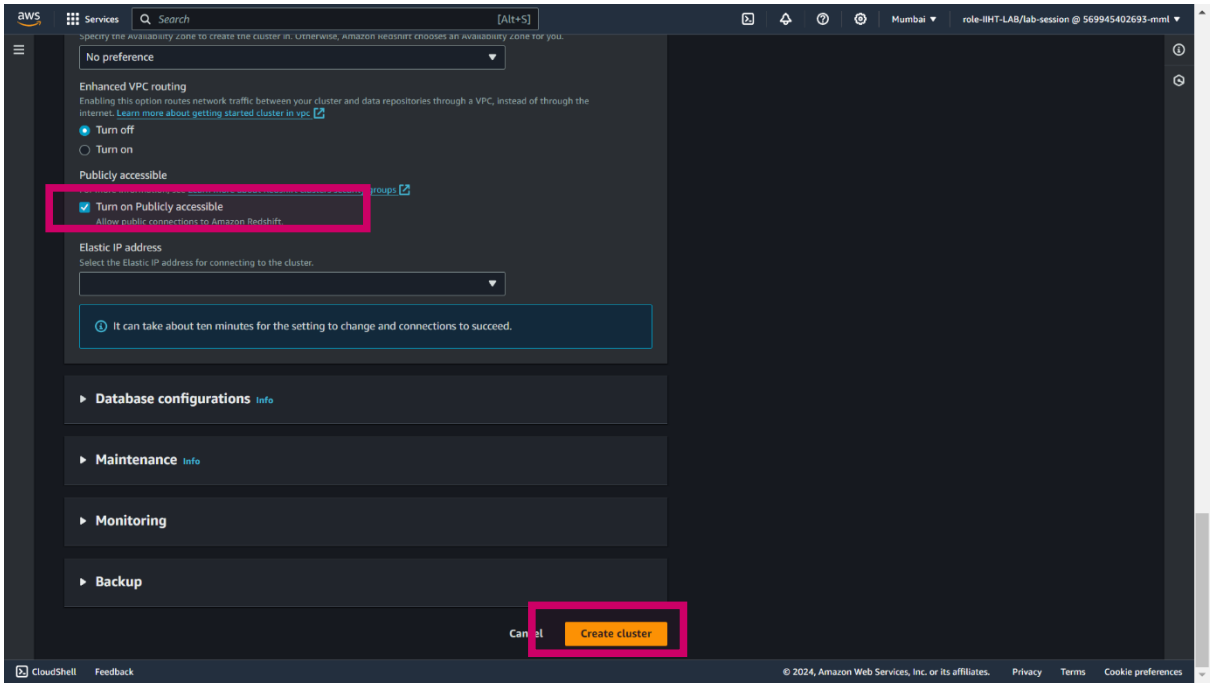
Step 7: Creating Cluster

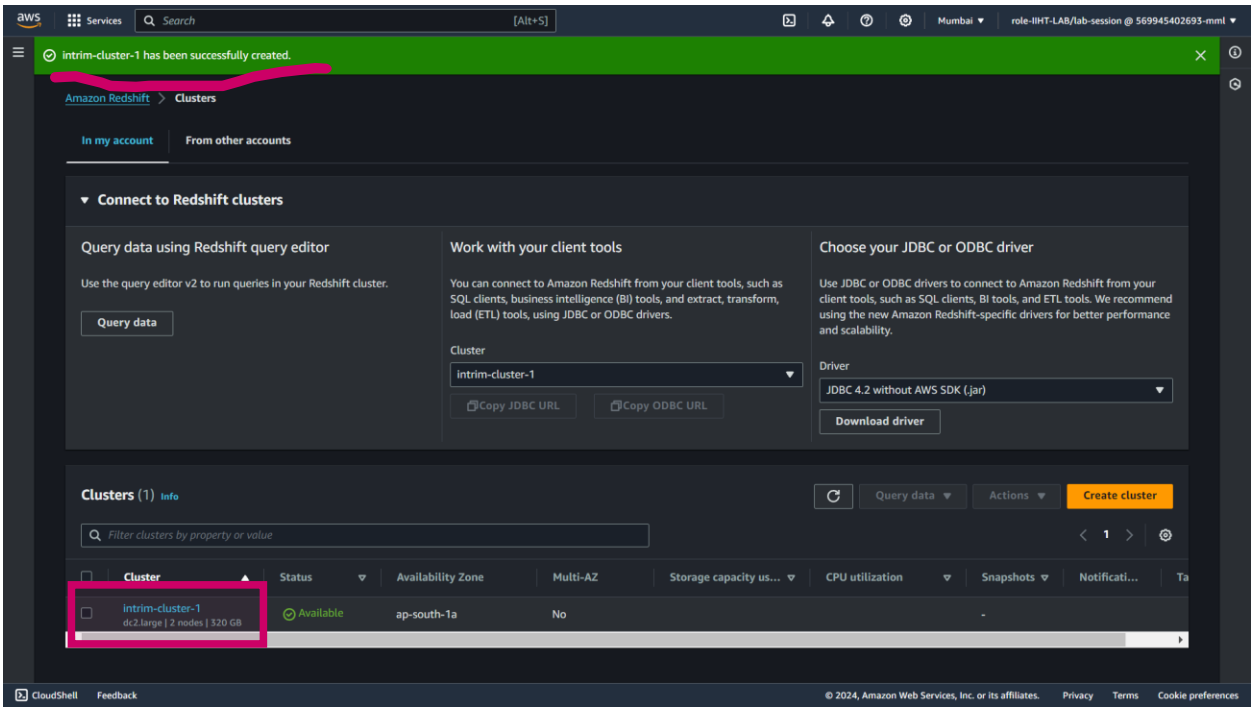




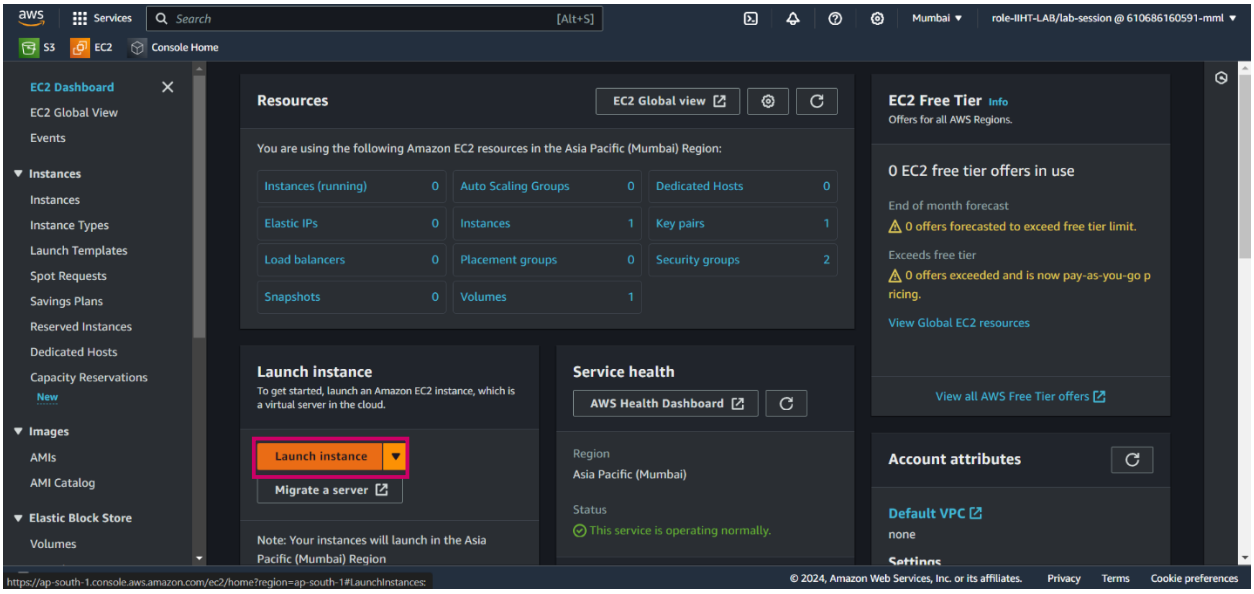
Step 8: Creating Cluster Subnet Group

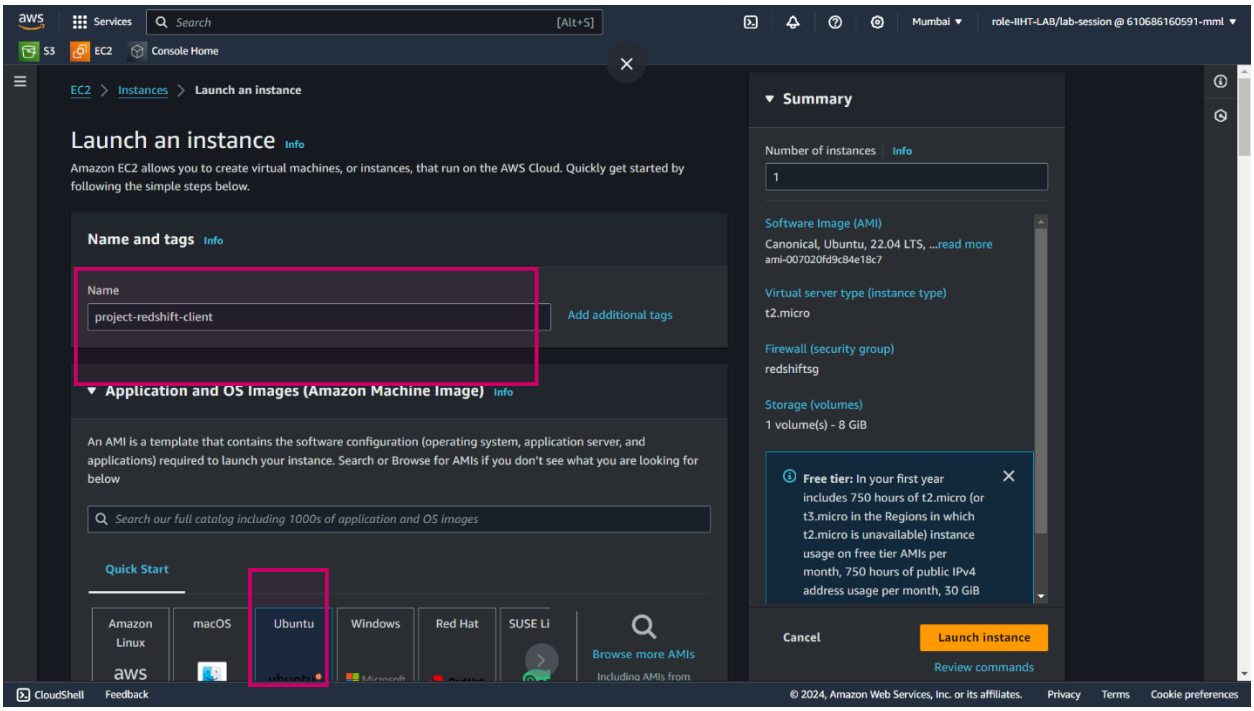

Cluster subnet group details
Name
intrim-subnet-1
Description
subnet
Add subnets
VPC
vpc-0d7dd00aa70582c12
Add all the subnets for this VPC
Availability Zone
Choose an Availability Zone
Subnet
Add subnet
Subnets in this cluster subnet group (1)
Remove all
Availability Zone
Subnet ID
CIDR block
IPv6 CIDR block
Action
ap-south-1a
subnet-07e203...
172.31.32.0/20
Remove
Cancel
Create cluster subnet group

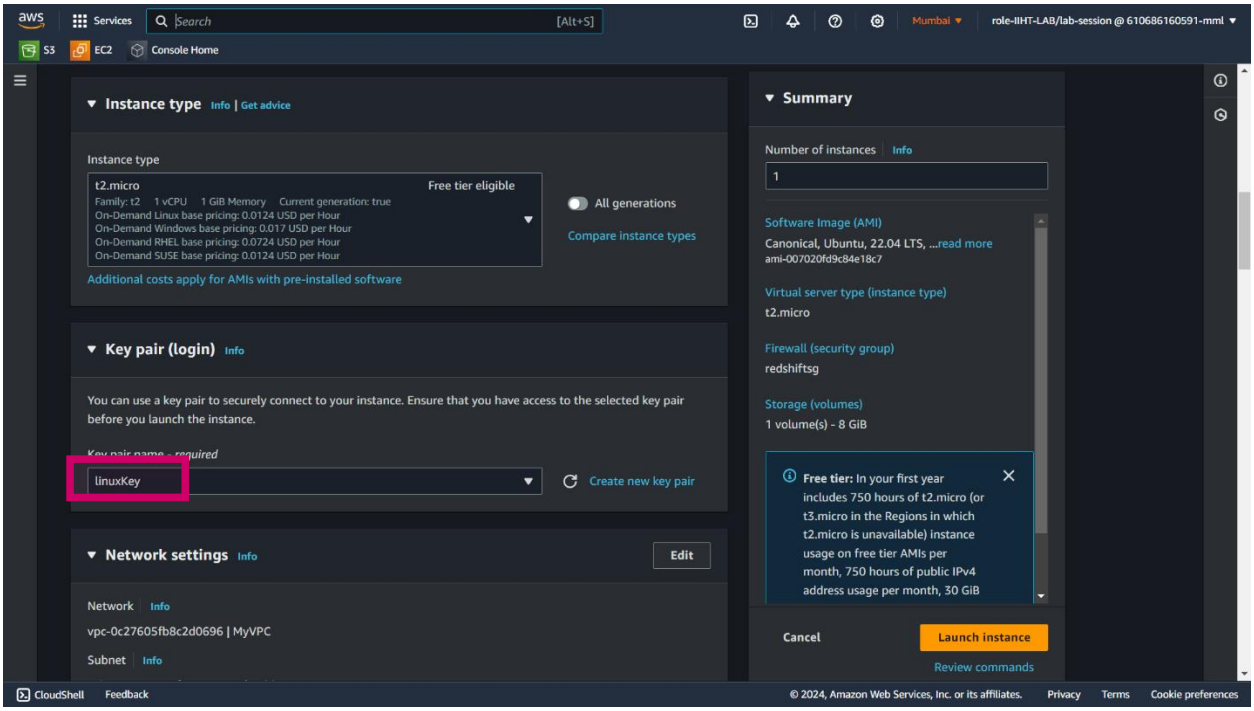

No preference
Enhanced VPC routing
Turn off
Turn on
Publicly accessible
Turn on Publicly accessible
Elastic IP address
It can take about ten minutes for the setting to change and connections to succeed.
Database configurations
Maintenance
Monitoring
Backup
Create cluster

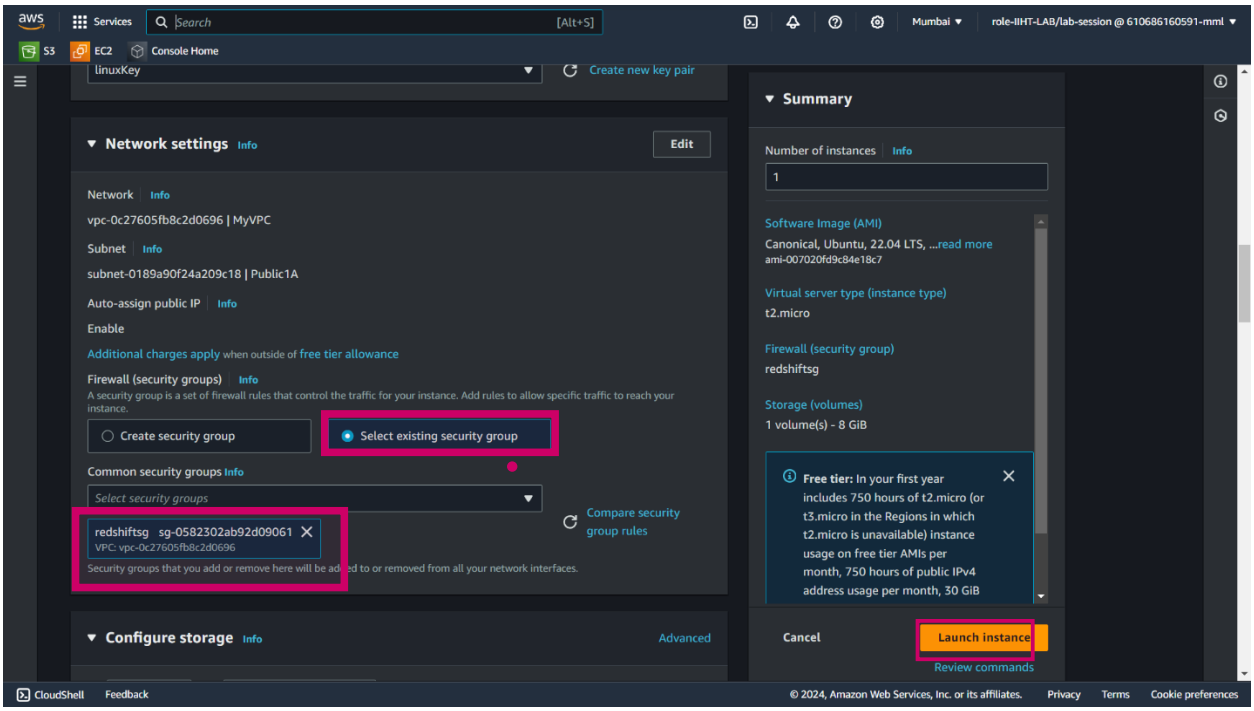


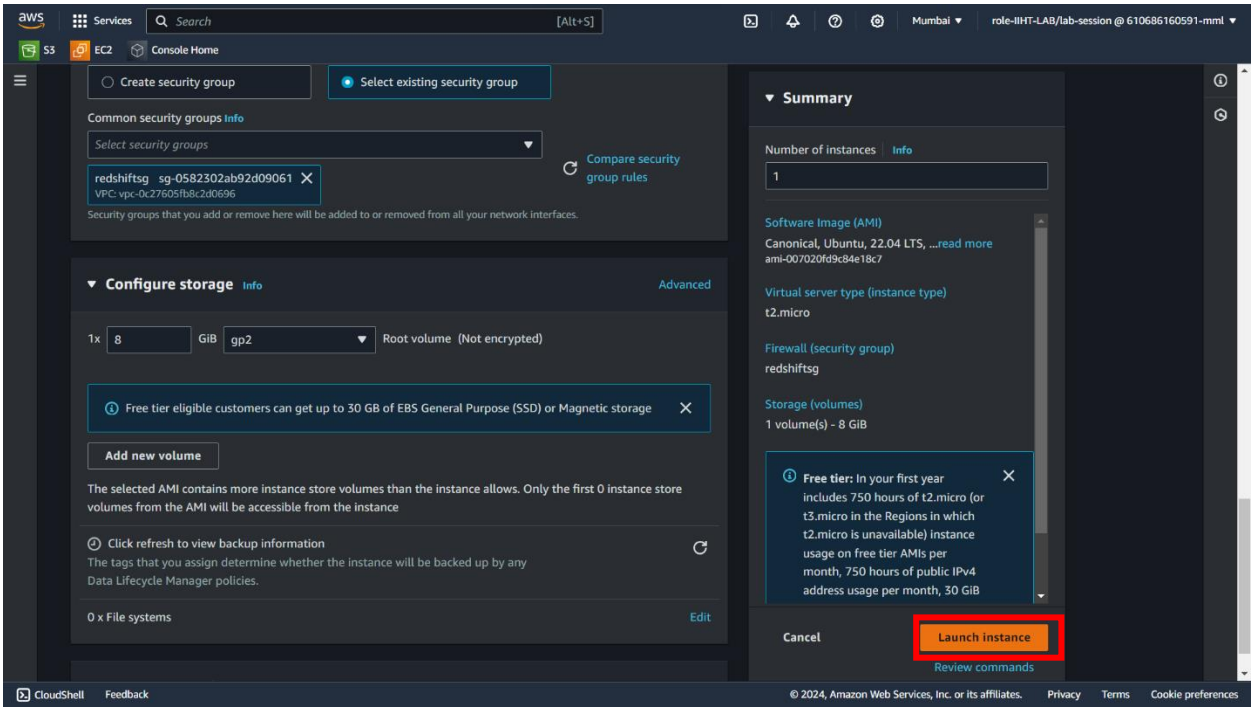
Step 9 : Creating an EC2 instance

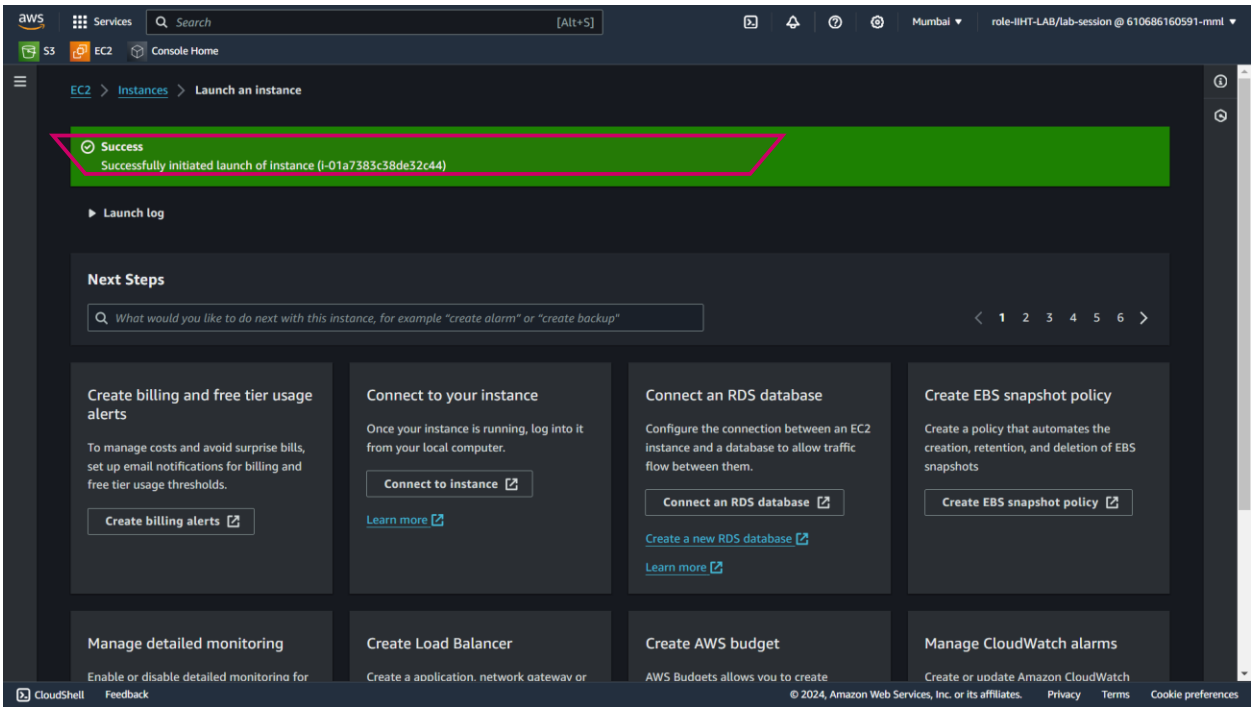


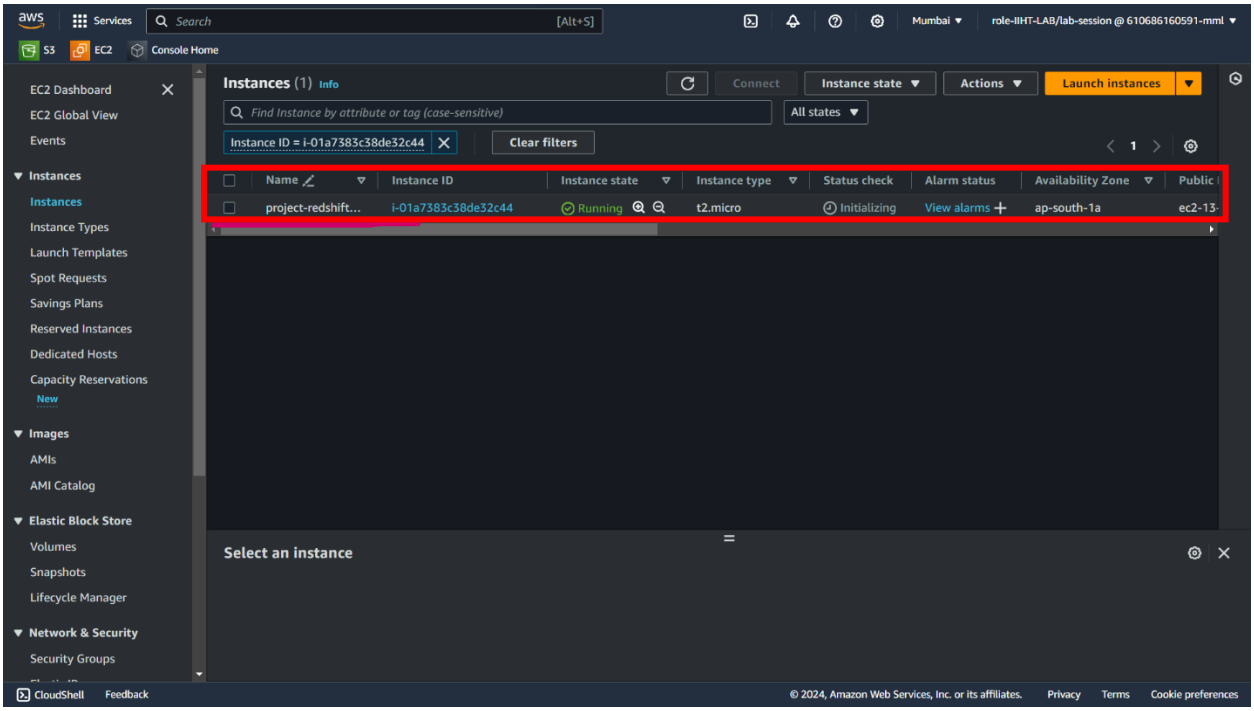

Services
Search
[Alt+S]
Mumbai
role-IIHT-LAB/lab-session @ 610686160591-mml
S3
EC2
Console Home
EC2 > Instances > Launch an instance
Launch an instance Info
Amazon EC2 allows you to create virtual machines, or instances, that run on the AWS Cloud. Quickly get started by following the simple steps below.
Name and tags Info
Name
project-redshift-client
Add additional tags
Application and OS Images (Amazon Machine Image) Info
An AMI is a template that contains the software configuration (operating system, application server, and applications) required to launch your instance. Search or Browse for AMIs if you don't see what you are looking for below
Search our full catalog including 1000s of application and OS images
Quick Start
Amazon Linux
macOS
Ubuntu
Windows
Red Hat
SUSE Li
Browse more AMIs
Including AMIs from
Summary
Number of instances Info
1
Software Image (AMI)
Canonical, Ubuntu, 22.04 LTS, ...read more
ami-007020fd9c84e18c7
Virtual server type (instance type)
t2.micro
Firewall (security group)
redshiftsg
Storage (volumes)
1 volume(s) - 8 GiB
Free tier: In your first year includes 750 hours of t2.micro (or t3.micro in the Regions in which t2.micro is unavailable) instance usage on free tier AMIs per month, 750 hours of public IPv4 address usage per month, 30 GiB
Cancel
Launch instance
Review commands
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

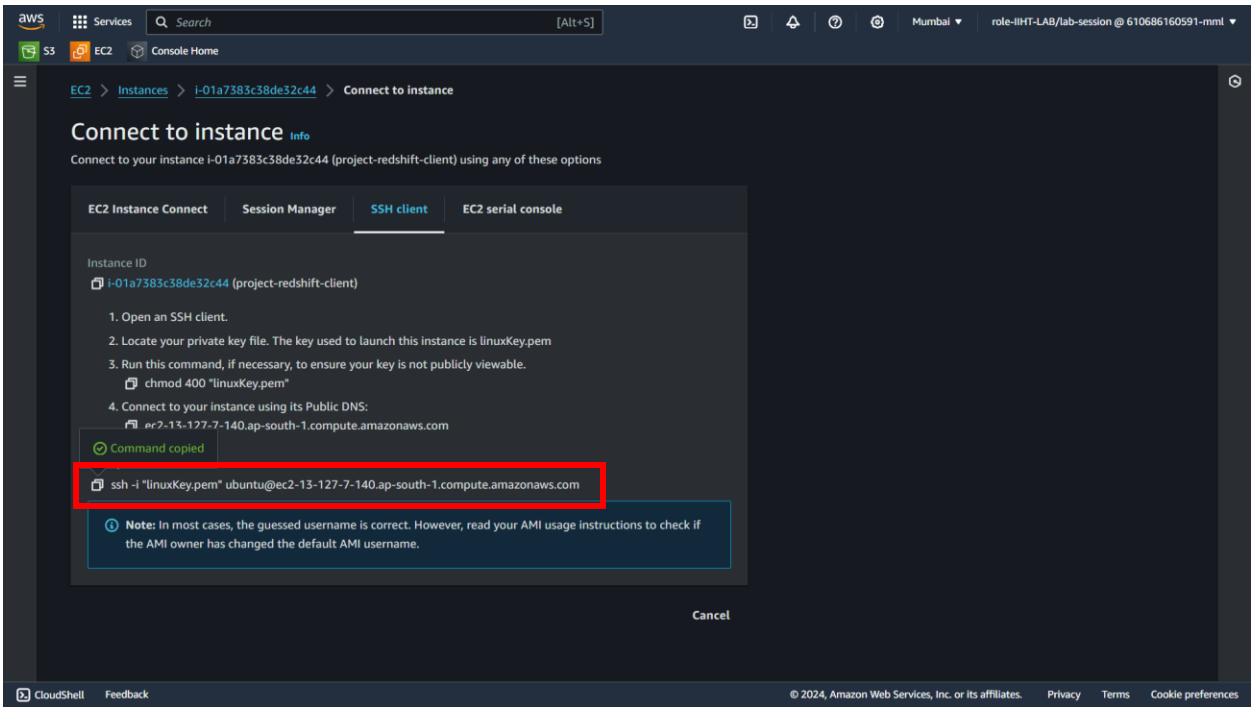

Services
Search
[Alt+S]
Mumbai
role-IIHT-LAB/lab-session @ 610686160591-mml
S3
EC2
Console Home
Instance type Info | Get advice
Instance type
t2.micro
Free tier eligible
Family: t2 1 vCPU 1 GiB Memory Current generation: true
On-Demand Linux base pricing: 0.0124 USD per Hour
On-Demand Windows base pricing: 0.017 USD per Hour
On-Demand RHEL base pricing: 0.0724 USD per Hour
On-Demand SUSE base pricing: 0.0124 USD per Hour
All generations
Compare instance types
Additional costs apply for AMIs with pre-installed software
Key pair (login) Info
You can use a key pair to securely connect to your instance. Ensure that you have access to the selected key pair before you launch the instance.
Key pair name - required
linuxKey
Create new key pair
Network settings Info
Edit
Network Info
vpc-0c27605fb8c2d0696 | MyVPC
Subnet Info
Summary
Number of instances Info
1
Software Image (AMI)
Canonical, Ubuntu, 22.04 LTS, ...read more
ami-007020fd9c84e18c7
Virtual server type (instance type)
t2.micro
Firewall (security group)
redshiftsg
Storage (volumes)
1 volume(s) - 8 GiB
Free tier: In your first year includes 750 hours of t2.micro (or t3.micro in the Regions in which t2.micro is unavailable) instance usage on free tier AMIs per month, 750 hours of public IPv4 address usage per month, 30 GiB
Cancel
Launch instance
Review commands
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


Services
Search
[Alt+S]
Mumbai
role-IIHT-LAB/lab-session @ 610686160591-mml
S3
EC2
Console Home
linuxKey
Create new key pair
Network settings Info
Edit
Network Info
vpc-0c27605fb8c2d0696 | MyVPC
Subnet Info
subnet-0189a90f24a209c18 | Public1A
Auto-assign public IP Info
Enable
Additional charges apply when outside of free tier allowance
Firewall (security groups) Info
A security group is a set of firewall rules that control the traffic for your instance. Add rules to allow specific traffic to reach your instance.
Create security group
Select existing security group
Common security groups Info
Select security groups
redshiftsg sg-0582302ab92d09061
VPC: vpc-0c27605fb8c2d0696
Security groups that you add or remove here will be added to or removed from all your network interfaces.
Compare security group rules
Configure storage Info
Advanced
Summary
Number of instances Info
1
Software Image (AMI)
Canonical, Ubuntu, 22.04 LTS, ...read more
ami-007020fd9c84e18c7
Virtual server type (instance type)
t2.micro
Firewall (security group)
redshiftsg
Storage (volumes)
1 volume(s) - 8 GiB
Free tier: In your first year includes 750 hours of t2.micro (or t3.micro in the Regions in which t2.micro is unavailable) instance usage on free tier AMIs per month, 750 hours of public IPv4 address usage per month, 30 GiB
Cancel
Launch instance
Review commands
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


Services
Search
[Alt+S]
Mumbai
role-IIHT-LAB/lab-session @ 610686160591-mml
S3
EC2
Console Home
Create security group
Select existing security group
Common security groups Info
Select security groups
redshiftsg sg-0582302ab92d09061
VPC: vpc-0c27605fb8c2d0696
Security groups that you add or remove here will be added to or removed from all your network interfaces.
Compare security group rules
Configure storage Info
Advanced
1x 8 GiB gp2 Root volume (Not encrypted)
Free tier eligible customers can get up to 30 GB of EBS General Purpose (SSD) or Magnetic storage
Add new volume
The selected AMI contains more instance store volumes than the instance allows. Only the first 0 instance store volumes from the AMI will be accessible from the instance
Click refresh to view backup information
The tags that you assign determine whether the instance will be backed up by any Data Lifecycle Manager policies.
0 x File systems
Edit
Summary
Number of instances Info
1
Software Image (AMI)
Canonical, Ubuntu, 22.04 LTS, ...read more
ami-007020fd9c84e18c7
Virtual server type (instance type)
t2.micro
Firewall (security group)
redshiftsg
Storage (volumes)
1 volume(s) - 8 GiB
Free tier: In your first year includes 750 hours of t2.micro (or t3.micro in the Regions in which t2.micro is unavailable) instance usage on free tier AMIs per month, 750 hours of public IPv4 address usage per month, 30 GiB
Cancel
Launch instance
Review commands
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


Services
Search
[Alt+S]
Mumbai
role-IIHT-LAB/lab-session @ 610686160591-mml
S3
EC2
Console Home
EC2 > Instances > Launch an instance
Success
Successfully initiated launch of instance (i-01a7383c38de32c44)
Launch log
Next Steps
What would you like to do next with this instance, for example "create alarm" or "create backup"
Create billing and free tier usage alerts
To manage costs and avoid surprise bills, set up email notifications for billing and free tier usage thresholds.
Create billing alerts
Connect to your instance
Once your instance is running, log into it from your local computer.
Connect to instance
Learn more
Connect an RDS database
Configure the connection between an EC2 instance and a database to allow traffic flow between them.
Connect an RDS database
Create a new RDS database
Learn more
Create EBS snapshot policy
Create a policy that automates the creation, retention, and deletion of EBS snapshots
Create EBS snapshot policy
Manage detailed monitoring
Create Load Balancer
Create AWS budget
Manage CloudWatch alarms
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


Services
Search
[Alt+S]
Mumbai
role-IIHT-LAB/lab-session @ 610686160591-mml
S3
EC2
Console Home
EC2 Dashboard
EC2 Global View
Events
Instances
Instances
Instance Types
Launch Templates
Spot Requests
Savings Plans
Reserved Instances
Dedicated Hosts
Capacity Reservations
New
Images
AMIs
AMI Catalog
Elastic Block Store
Volumes
Snapshots
Lifecycle Manager
Network & Security
Security Groups
Instances (1) Info
Connect
Instance state
Actions
Launch instances
Find Instance by attribute or tag (case-sensitive)
All states
Instance ID = i-01a7383c38de32c44
Clear filters
Name
Instance ID
Instance state
Instance type
Status check
Alarm status
Availability Zone
project-redshift...
i-01a7383c38de32c44
Running
t2.micro
Initializing
View alarms
ap-south-1a
Select an instance
CloudShell
Feedback
© 2024, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences



Step 10 : Open terminal and run the commands to create a table in Ec2 instance.

These are the commands to run the table in Ec2 instance.

- “sudo apt-get update
- sudo apt-get install -y postgresql-client && psql --version
- psql -h redshift-cluster-1.cjlluv3z3ar1.ap-south-1.redshift.amazonaws.com -p 5439 -U awsuser -d dev -W”

Syntax : pql –h <hostname> -p <port> -U <user> -d <database> -W

```
PS C:\Users\2320846> cd downloads
PS C:\Users\2320846\downloads> ssh -i "linuxKey.pem" ubuntu@ec2-13-127-7-140.ap-south-1.compute.amazonaws.com
The authenticity of host 'ec2-13-127-7-140.ap-south-1.compute.amazonaws.com (13.127.7.140)' can't be established.
ED25519 key fingerprint is SHA256:iAINGIoqfOLVshcfpGnAja3pPJ43zbJHmWFFV2FVPo0.
This key is not known by any other names
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'ec2-13-127-7-140.ap-south-1.compute.amazonaws.com' (ED25519) to the list of known hosts.
Welcome to Ubuntu 22.04.4 LTS (GNU/Linux 6.5.0-1014-aws x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:     https://landscape.canonical.com
 * Support:        https://ubuntu.com/pro

  System information as of Thu Mar 21 14:42:13 UTC 2024

  System load:  0.0               Processes:             98
  Usage of /:   20.4% of 7.57GB   Users logged in:       0
  Memory usage: 19%               IPv4 address for eth0: 10.0.1.38
  Swap usage:   0%

Expanded Security Maintenance for Applications is not enabled.

0 updates can be applied immediately.

Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status


The list of available updates is more than a week old.
To check for new updates run: sudo apt update


The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

ubuntu@ip-10-0-1-38:~$
```

- Connecting to the Redshift cluster from ec2 through postgreSQL client

```
ubuntu@ip-10-0-1-38:~$ psql -h redshift-cluster-1.c4hapfhrjfor.ap-south-1.redshift.amazonaws.com -p 5439 -U awsuser -d dev -W
Password:
psql (14.11 (Ubuntu 14.11-0ubuntu0.22.04.1), server 8.0.2)
SSL connection (protocol: TLSv1.2, cipher: ECDHE-RSA-AES256-GCM-SHA384, bits: 256, compression: off)
Type "help" for help.
```

- create database for our use

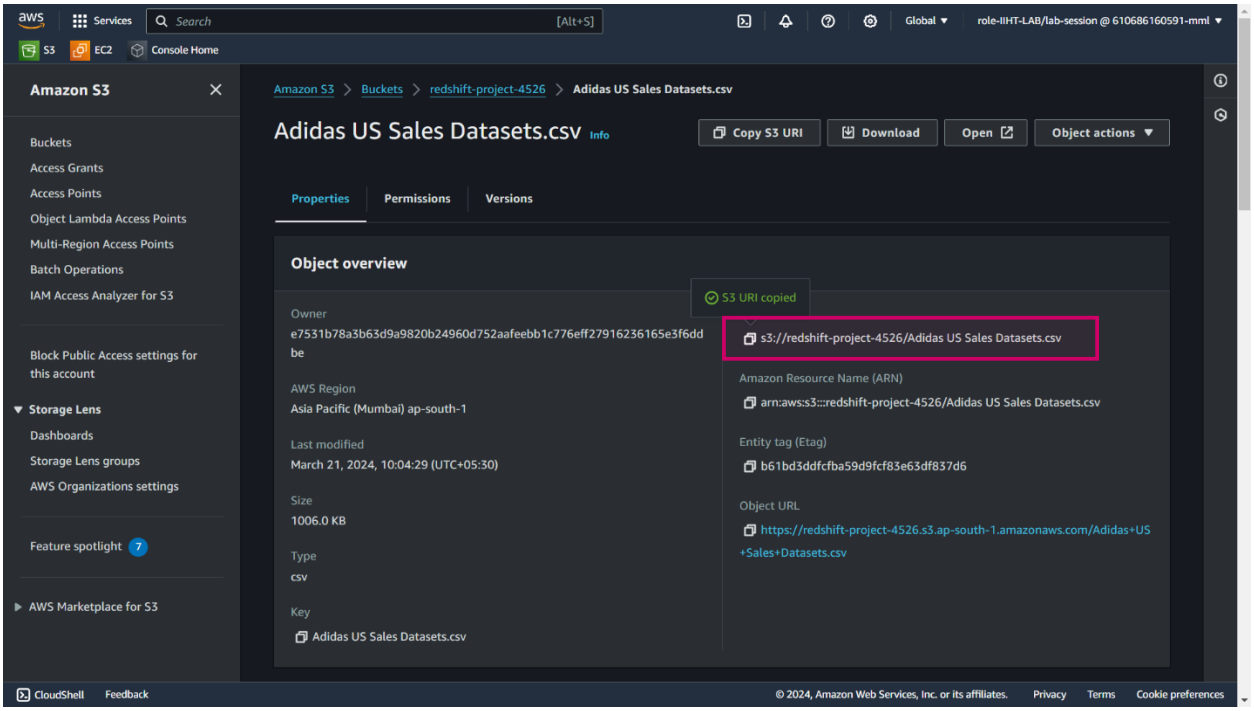
```
dev=# create database project ;
CREATE DATABASE
dev=#
```

- Again we connect to Redshift and login to the created database
- Create a table

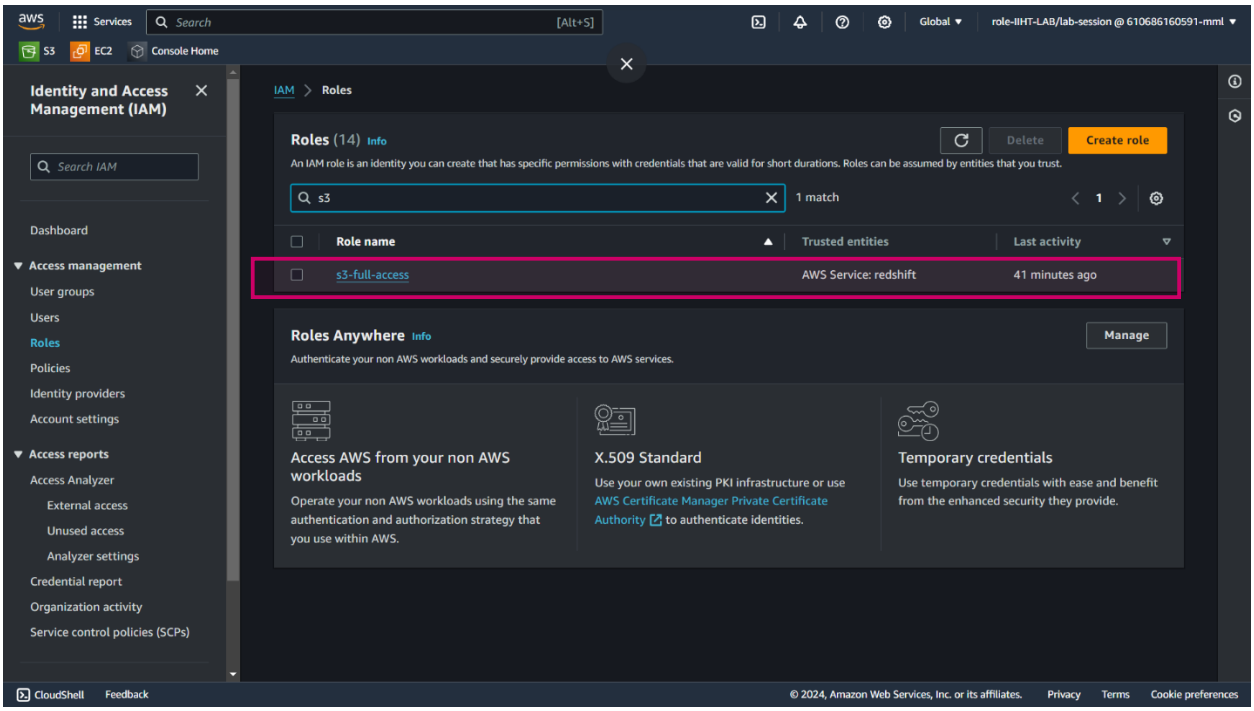
```
ubuntu@ip-10-0-1-38:~$ psql -h redshift-cluster-1.c4hapfhrjfor.ap-south-1.redshift.amazonaws.com -p 5439 -U awsuser -d project -W
Password:
psql (14.11 (Ubuntu 14.11-0ubuntu0.22.04.1), server 8.0.2)
SSL connection (protocol: TLSv1.2, cipher: ECDHE-RSA-AES256-GCM-SHA384, bits: 256, compression: off)
Type "help" for help.

project=# \d
Did not find any relations.
project=#
CREATE TABLE public.sales_data (
    retailer character varying(255) ENCODE lzo,
    retailer_id integer ENCODE az64,
    Invoice_date character varying(255) ENCODE lzo,
    region
        character varying(255) ENCODE lzo,
        state character varying(255) ENCODE lzo,
        city character varying(255) ENCODE lzo,
        product character varying(255) ENCODE lzo,
        price_per_unit numeric(10, 2) ENCODE az64,
        units_sold integer ENCODE az64,
        total_sales numeric(10, 2) ENCODE az64,
        operating_profit numeric(10, 2) ENCODE az64,
        operating_margin numeric(5, 2) ENCODE az64,
        sales_method character varying(255) ENCODE lzo
) DISTSTYLE AUTO;
CREATE TABLE
project=#
```

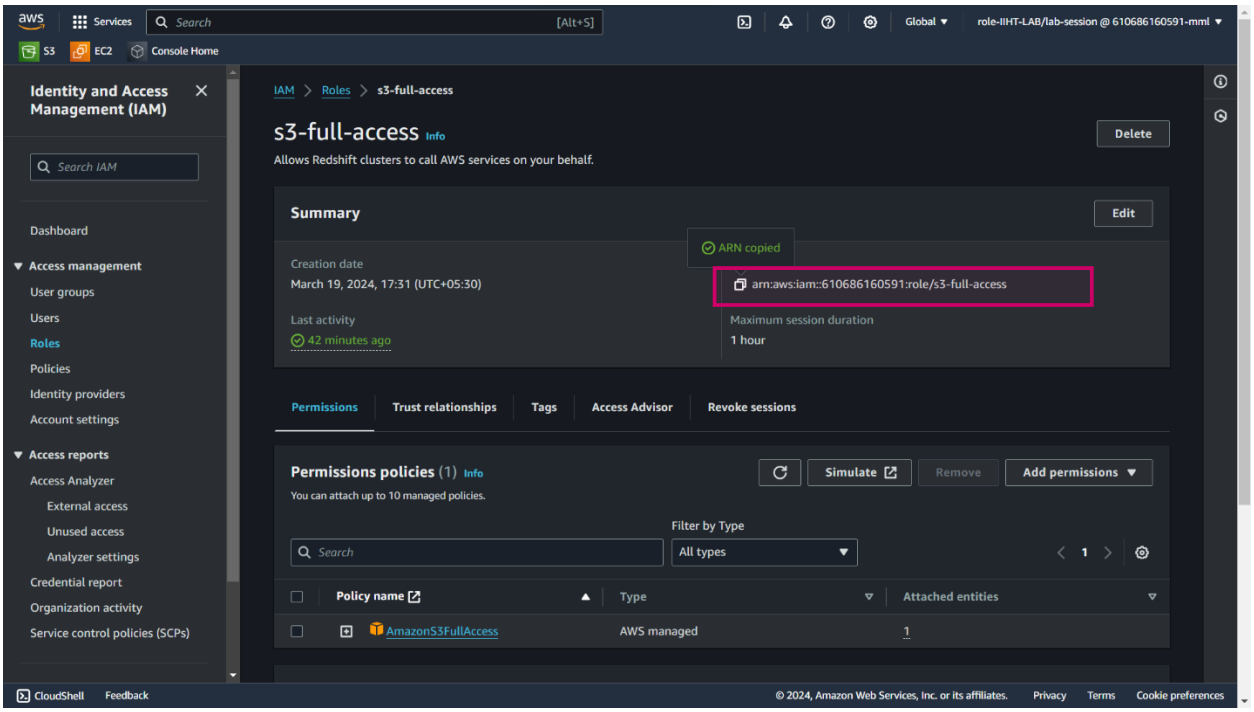
Step 11: Load the data from S3 to Redshift

To load the data we need S3 object URI



We need the IAM S3 Role ARN that we created earlier





Then use the following command to load the data

```
project=#
COPY project.public.sales_data FROM 's3://redshift-project-4526/Adidas US Sales Datasets.csv' IAM_ROLE 'arn:aws:iam::610686160591:rol
e/s3-full-access' FORMAT AS CSV DELIMITER ',' QUOTE '"' IGNOREHEADER 1 REGION AS 'ap-south-1' ;

INFO:  Load into table 'sales_data' completed, 9648 record(s) loaded successfully.
COPY
project=# select * from sales_data limit 10 ;
```

COPY <tablename> FROM <S3 bucket URI> IAM_ROLE <IAM role A> FORMAT AS CSV DELIMITER<file delimiter> REGION AS <cluster region>

retailer	retailer_id	invoice_date	region	state	city	product	price_per_unit	units_sold	total_sales	operating_profit	operating_margin	sales_method
Foot Locker	1185732	01-01-2020	Northeast	New York	New York	Men's Street Footwear	50.00	1200	600000.00	300000.00	50.00	In-store
Foot Locker	1185732	02-01-2020	Northeast	New York	New York	Men's Athletic Footwear	50.00	1000	500000.00	150000.00	30.00	In-store
Foot Locker	1185732	03-01-2020	Northeast	New York	New York	Women's Street Footwear	40.00	1000	400000.00	140000.00	35.00	In-store
Foot Locker	1185732	04-01-2020	Northeast	New York	New York	Women's Athletic Footwear	45.00	850	382500.00	133875.00	35.00	In-store
Foot Locker	1185732	05-01-2020	Northeast	New York	New York	Men's Apparel	60.00	900	540000.00	162000.00	30.00	In-store
Foot Locker	1185732	06-01-2020	Northeast	New York	New York	Women's Apparel	50.00	1000	500000.00	125000.00	25.00	In-store
Foot Locker	1185732	07-01-2020	Northeast	New York	New York	Men's Street Footwear	50.00	1250	625000.00	312500.00	50.00	In-store
Foot Locker	1185732	08-01-2020	Northeast	New York	New York	Men's Athletic Footwear	50.00	900	450000.00	135000.00	30.00	Outlet
Foot Locker	1185732	21-01-2020	Northeast	New York	New York	Women's Street Footwear	40.00	950	380000.00	133000.00	35.00	Outlet
Foot Locker	1185732	22-01-2020	Northeast	New York	New York	Women's Athletic Footwear	45.00	825	371250.00	129938.00	35.00	Outlet
Foot Locker	1185732	23-01-2020	Northeast	New York	New York	Men's Apparel	60.00	900	540000.00	162000.00	30.00	Outlet
Foot Locker	1185732	24-01-2020	Northeast	New York	New York	Women's Apparel	50.00	1000	500000.00	125000.00	25.00	Outlet
Foot Locker	1185732	25-01-2020	Northeast	New York	New York	Men's Street Footwear	50.00	1220	610000.00	305000.00	50.00	Outlet
Foot Locker	1185732	26-01-2020	Northeast	New York	New York	Men's Athletic Footwear	50.00	925	462500.00	138750.00	30.00	Outlet
Foot Locker	1185732	27-01-2020	Northeast	New York	New York	Women's Street Footwear	40.00	950	380000.00	133000.00	35.00	Outlet
Foot Locker	1185732	28-01-2020	Northeast	New York	New York	Women's Athletic Footwear	45.00	800	360000.00	126000.00	35.00	Outlet
Foot Locker	1185732	29-01-2020	Northeast	New York	New York	Men's Apparel	60.00	850	510000.00	153000.00	30.00	Outlet
Foot Locker	1185732	30-01-2020	Northeast	New York	New York	Women's Apparel	50.00	950	475000.00	118750.00	25.00	Outlet
Foot Locker	1185732	31-01-2020	Northeast	New York	New York	Men's Street Footwear	50.00	1200	600000.00	300000.00	50.00	Outlet
Foot Locker	1185732	01-02-2020	Northeast	New York	New York	Men's Athletic Footwear	50.00	900	450000.00	135000.00	30.00	Outlet

(20 rows)

Database schema :

sales_data

	Field	Type	NL	CMP
A	retailer	character varying(50)	NULL	lzo
#	retailerid	integer	NULL	az64
A	invoice_date	character varying(50)	NULL	lzo
A	region	character varying(40)	NULL	lzo
A	state	character varying(40)	NULL	lzo
A	city	character varying(40)	NULL	lzo
A	product	character varying(100)	NULL	lzo
#	priceperunit	numeric(10,2)	NULL	az64
#	units_sold	integer	NULL	az64
#	total_sales	numeric(12,2)	NULL	az64
#	operatingprofit	numeric(12,2)	NULL	az64
#	operatingmargin	numeric(5,2)	NULL	az64
A	salesmethod	character varying(50)	NULL	lzo

Step 12: Data Cleaning

Looking at the data we found some data mismatch , so we have to clean the data first before performing any analysis on the data .

- First we have to change the invoice_date colum from character variable to date format

```
-- changing the invoice_date from character varying(50) to date format

-- step 1 : Adding new column
ALTER TABLE sales_data
ADD COLUMN order_date DATE;

-- Step 2: Update the new column with the converted values
UPDATE sales_data
SET order_date = TO_DATE(invoice_date, 'DD-MM-YYYY');

-- Step 3: Drop the old column
ALTER TABLE sales_data
DROP COLUMN INVOICE_DATE;

SELECT * FROM SALES_DATA;
```

Now the schema looking like this

sales_data

	Field	Type	NL	CMP
A	retailer	character varying(50)	NULL	lzo
#	retailerid	integer	NULL	az64
A	region	character varying(40)	NULL	lzo
A	state	character varying(40)	NULL	lzo
A	city	character varying(40)	NULL	lzo
A	product	character varying(100)	NULL	lzo
#	priceperunit	numeric(10,2)	NULL	az64
#	units_sold	integer	NULL	az64
#	total_sales	numeric(12,2)	NULL	az64
#	operatingprofit	numeric(12,2)	NULL	az64
#	operatingmargin	numeric(5,2)	NULL	az64
A	salesmethod	character varying(50)	NULL	lzo
📅	order_date	date	NULL	az64

- We observed that in the dataset the some column’s data was mismatching , these columns are

- Total_Sales column
- Operating profit column

```
select priceperunit , units_Sold , total_Sales , operatingprofit , operatingmargin from sales_data ;
```

priceperunit	units_sold	total_sales	operatingprofit	operatingmargin
50	1200	600000	300000	50
50	1000	500000	150000	30
40	1000	400000	140000	35
45	850	382500	133875	35
60	900	540000	162000	30
50	1000	500000	125000	25
50	1250	625000	312500	50

So we perform some more data cleaning for

Total_Sales :

```
-- Data cleaning for total sales column

-- ADDING A NEW COL
ALTER TABLE sales_data
ADD COLUMN TOTAL_COST INT;


 -- MANUALLY MULTIPY UNITS SOLD * PRICE PER UNIT
UPDATE sales_data
SET TOTAL_COST = PRICEPERUNIT * UNITS_SOLD;


--DROP OLD COLUMN
ALTER TABLE SALES_DATA
DROP COLUMN TOTAL_SALES;


-- RENAME COLUMN TO THE ORIGINAL NAME
ALTER TABLE SALES_DATA
RENAME TOTAL_COST TO TOTAL_SALES;
```

priceperunit	units_sold	total_sales	operatingprofit	operatingmargin
50	1200	60000	300000	50
50	1000	50000	150000	30
40	1000	40000	140000	35
45	850	38250	133875	35
60	900	54000	162000	30
50	1000	50000	125000	25

Operating profit :

```sql
-- Data cleaning for operating profit column


-- add NEW PROFIT col
ALTER TABLE sales_data
ADD COLUMN new_profit INT;

 -- MANUALLY CALCULATING THE PROFIT
UPDATE sales_data
SET new_profit =  ( operatingmargin * total_sales ) / 100 ;

--DROP OLD COLUMN OPERATING_MARGIN
ALTER TABLE SALES_DATA
DROP COLUMN operatingprofit;

-- RENAME THE NEW COL TO THE ORIGINAL COLUMN NAME
ALTER TABLE SALES_DATA
RENAME new_profit TO operating_profit;
```

priceperunit	units_sold	total_sales	operating_profit	operatingmargin
50	1200	60000	30000	50
50	1000	50000	15000	30
40	1000	40000	14000	35
45	850	38250	13387	35
60	900	54000	16200	30
50	1000	50000	12500	25

Now we will move to Redshift cluster to run the queries

Note: We can directly run queries from the instance but for better visuals we used Redshifts query editor v2

Step 13: Running SQL queries to analyze the sales trends

SQL queries to aggregate sales data by product, region, and time period.

```sql

  -- aggregate sales data by product
SELECT
    product,
    SUM(total_sales) AS total_sales
FROM sales_data
    GROUP BY product
    Order by total_sales DESC;


```

Result 1 (6)

product	total_sales
Men's Street Footwear	27680769
Women's Apparel	23870985
Men's Athletic Footwear	20577180
Women's Street Footwear	17201563
Men's Apparel	16520632
Women's Athletic Footwear	14315521

By Region

```
7   --aggregate sales data by region
8   SELECT
9       product, region  ,
10      SUM(total_sales) AS total_sales
11  FROM sales_data
12      GROUP BY region , product
13      Order by region;
14
```

Result 1 (30)

product	region	total_sales
Men's Apparel	Midwest	2223786
Women's Apparel	Midwest	3453008
Men's Street Footwear	Midwest	4707360
Men's Athletic Footwear	Midwest	2619289
Women's Street Footwear	Midwest	1997448
Women's Athletic Footwear	Midwest	1673543
Men's Street Footwear	Northeast	6841324
Men's Athletic Footwear	Northeast	3895862
Women's Street Footwear	Northeast	3152823
Women's Athletic Footwear	Northeast	2668013
Men's Apparel	Northeast	3475037
Women's Apparel	Northeast	5045208
Men's Apparel	South	2811194
Women's Apparel	South	4224937
Men's Street Footwear	South	4048261
Men's Athletic Footwear	South	3647045
Women's Street Footwear	South	3242822
Women's Athletic Footwear	South	2629097
Men's Apparel	Southeast	3183237
Women's Apparel	Southeast	4109786
Men's Street Footwear	Southeast	4693836
Men's Athletic Footwear	Southeast	3653645
Women's Street Footwear	Southeast	3059884
Women's Athletic Footwear	Southeast	2674048
Men's Apparel	West	4827378
Women's Apparel	West	7038046
Men's Street Footwear	West	7389988
Men's Athletic Footwear	West	6761339

By Time period

```sql
-- aggregate sales data by time period
SELECT
    product ,
    TO_CHAR(order_date::DATE,'MONTH' ) as month,
    sum(total_sales)  AS total_sales
FROM sales_data
    GROUP BY product , month  , REGION
    ORDER BY total_sales desc  ;
```

Virtual Servers in the Cloud 2

Result 1 (100)

product	month	total_sales
Women's Apparel	MAY	903900
Men's Street Footwear	JULY	894762
Men's Street Footwear	MAY	875565
Women's Apparel	JANUARY	870171
Men's Street Footwear	DECEMBER	865236
Women's Apparel	SEPTEMBER	845374
Men's Athletic Footwear	JANUARY	827784
Men's Athletic Footwear	MAY	798378
Men's Athletic Footwear	AUGUST	796036
Women's Street Footwear	SEPTEMBER	777403
Men's Street Footwear	AUGUST	775264
Men's Street Footwear	APRIL	771903
Women's Apparel	APRIL	757109
Men's Athletic Footwear	SEPTEMBER	753147
Men's Street Footwear	NOVEMBER	734340
Men's Street Footwear	JANUARY	733706
Men's Street Footwear	SEPTEMBER	732989
Women's Apparel	JUNE	727454
Men's Athletic Footwear	JULY	720404
Women's Apparel	JULY	718496
Men's Street Footwear	FEBRUARY	717585
Men's Athletic Footwear	FEBRUARY	717550
Women's Street Footwear	JANUARY	711685
Men's Street Footwear	JANUARY	693038
Women's Apparel	AUGUST	676628
Men's Street Footwear	JULY	674139
Women's Apparel	FEBRUARY	650157
Men's Street Footwear	JUNE	642840
Men's Street Footwear	MARCH	638749
Women's Apparel	DECEMBER	630685
Women's Street Footwear	MAY	627273
Women's Apparel	OCTOBER	623236
Men's Athletic Footwear	APRIL	620850
Men's Street Footwear	DECEMBER	620123
Men's Street Footwear	AUGUST	619224
Men's Street Footwear	SEPTEMBER	618857
Women's Street Footwear	FEBRUARY	617181
Women's Street Footwear	JULY	604788
Men's Apparel	MAY	601740
Women's Street Footwear	APRIL	595061
Men's Apparel	JULY	578730

product	month	total_sales
Men's Apparel	MAY	601740
Women's Street Footwear	APRIL	595061
Men's Apparel	JULY	578730
Men's Street Footwear	JUNE	566219
Men's Athletic Footwear	AUGUST	561319
Women's Athletic Footwear	MAY	559951
Men's Street Footwear	AUGUST	558716
Men's Street Footwear	MAY	557409
Men's Street Footwear	DECEMBER	552861
Men's Street Footwear	JULY	551740
Men's Athletic Footwear	JUNE	549573
Women's Athletic Footwear	JANUARY	543866
Men's Apparel	JANUARY	541701
Women's Street Footwear	JUNE	539924
Women's Street Footwear	AUGUST	539454
Women's Apparel	NOVEMBER	538436
Women's Apparel	DECEMBER	532298
Women's Apparel	JULY	525885
Men's Street Footwear	MARCH	524016
Women's Apparel	JULY	522292
Men's Apparel	SEPTEMBER	513620
Men's Street Footwear	OCTOBER	511351
Men's Street Footwear	OCTOBER	508353
Men's Street Footwear	APRIL	506990
Men's Street Footwear	JULY	503231
Women's Athletic Footwear	SEPTEMBER	501163
Men's Apparel	FEBRUARY	497137
Women's Apparel	NOVEMBER	495827
Women's Athletic Footwear	JULY	494946
Women's Apparel	SEPTEMBER	494112
Men's Athletic Footwear	DECEMBER	492676
Men's Athletic Footwear	MARCH	487809
Men's Athletic Footwear	SEPTEMBER	485515
Women's Apparel	APRIL	485057
Men's Street Footwear	AUGUST	484315
Women's Athletic Footwear	FEBRUARY	483639
Women's Street Footwear	OCTOBER	483456
Men's Athletic Footwear	OCTOBER	482089
Men's Athletic Footwear	AUGUST	480367
Men's Apparel	AUGUST	478597
Women's Street Footwear	MARCH	474928

Identify top-selling products by calculating total sales or units sold.

By total sales

```
28
29  -- Identify top-selling products by calculating total sales
30  select
31      product
32          from (
33                  select
34                      product ,
35                      sum(total_sales ) as total_sales
36                  from sales_data
37                  group by product
38                  order by total_sales DESC
39              )
40  ;
41
```

product
Men's Street Footwear
Women's Apparel
Men's Athletic Footwear
Women's Street Footwear
Men's Apparel
Women's Athletic Footwear

By units sold

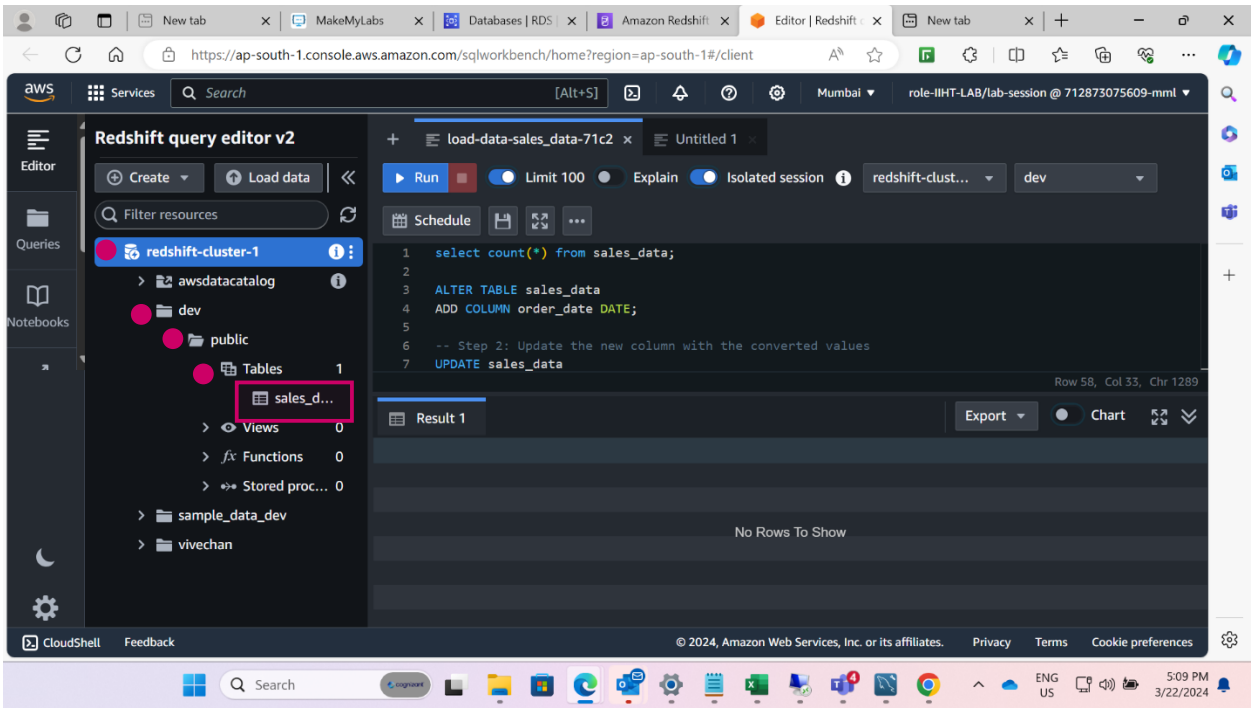
```
44  -- Identify top-selling products by calculating units sold.
45
46  select
47      product as top_selling_products
48          from (
49                  select
50                      product ,
51                      sum(units_sold) as units_sold
52                  from sales_data
53                  group by product
54                  order by units_sold  DESC
55              ) as tab
56  ;
```

top_selling_products
Men's Street Footwear
Men's Athletic Footwear
Women's Apparel
Women's Street Footwear
Women's Athletic Footwear
Men's Apparel

Determine highest revenue-generating regions by analyzing sales revenue per region.

```sql
-- Determine the highest revenue generating regions by analyzing  sales revenue per region

select
    region ,
    sum(total_sales ) as total_sales
from sales_data
    group by region
    order by total_sales  desc ;
```

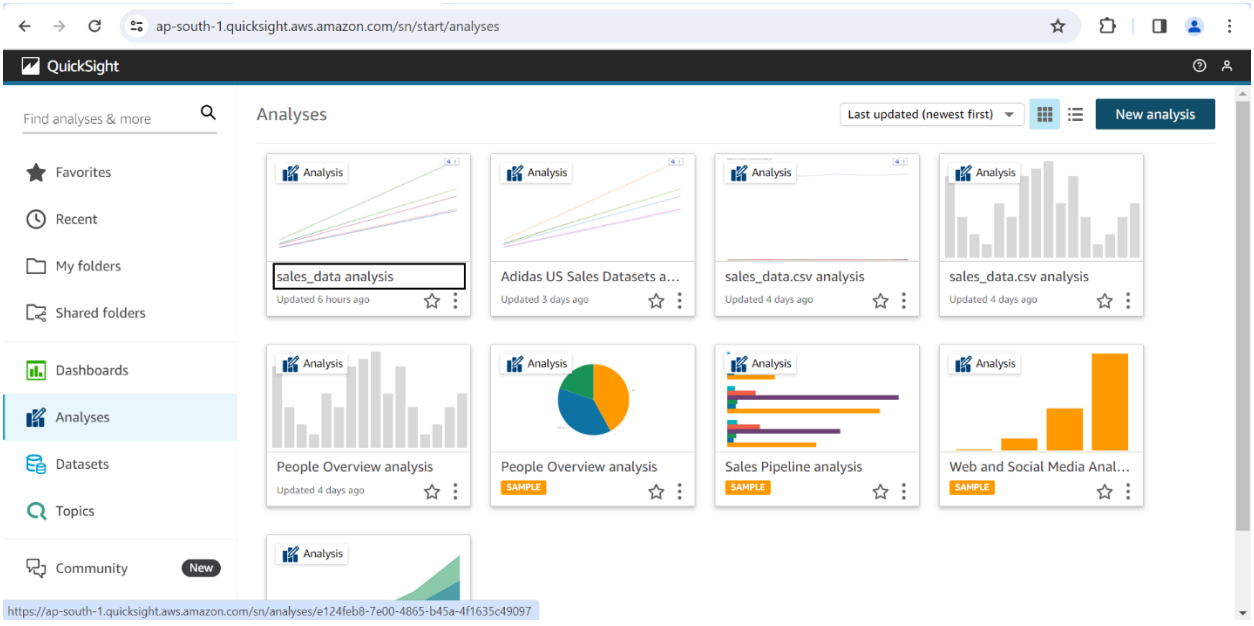
region	total_sales
West	36436157
Northeast	25078267
Southeast	21374436
South	20603356
Midwest	16674434

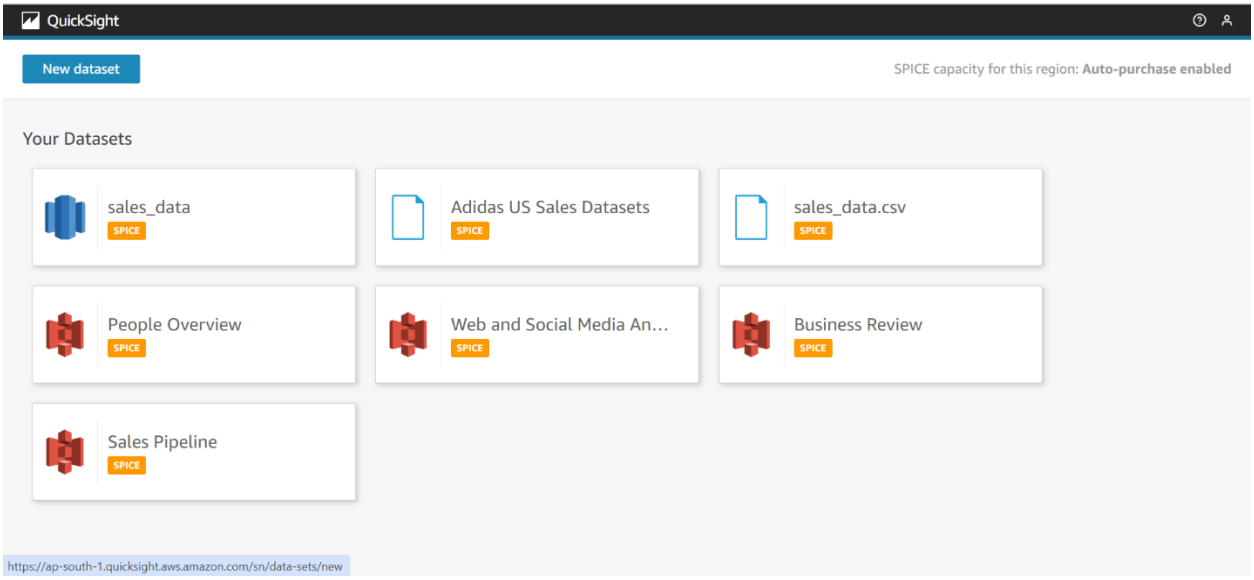


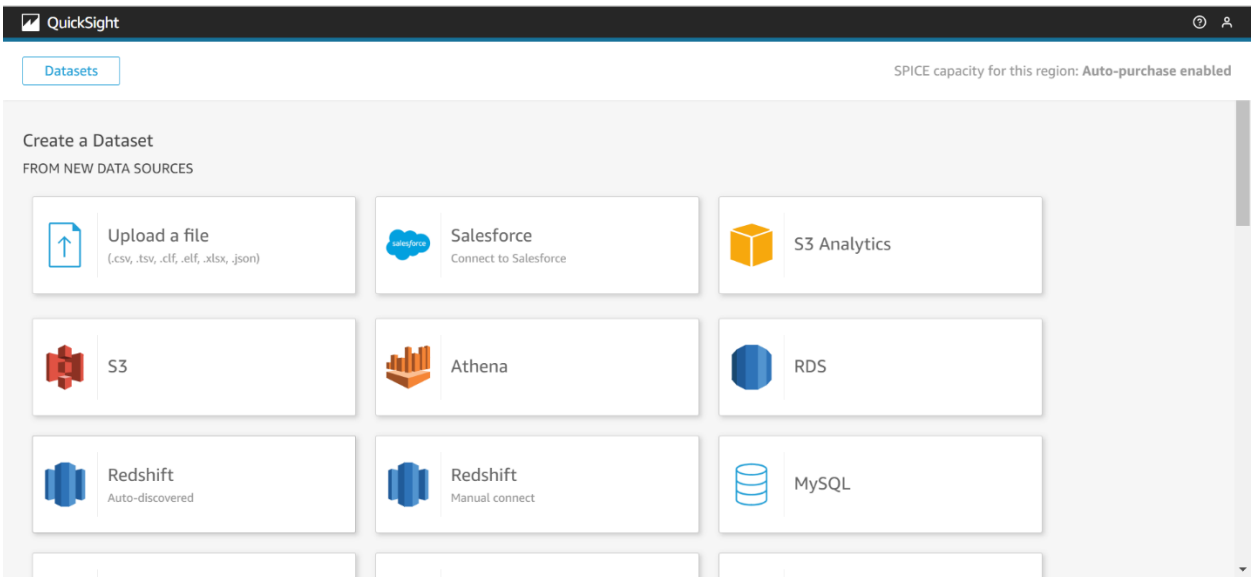
Visualization

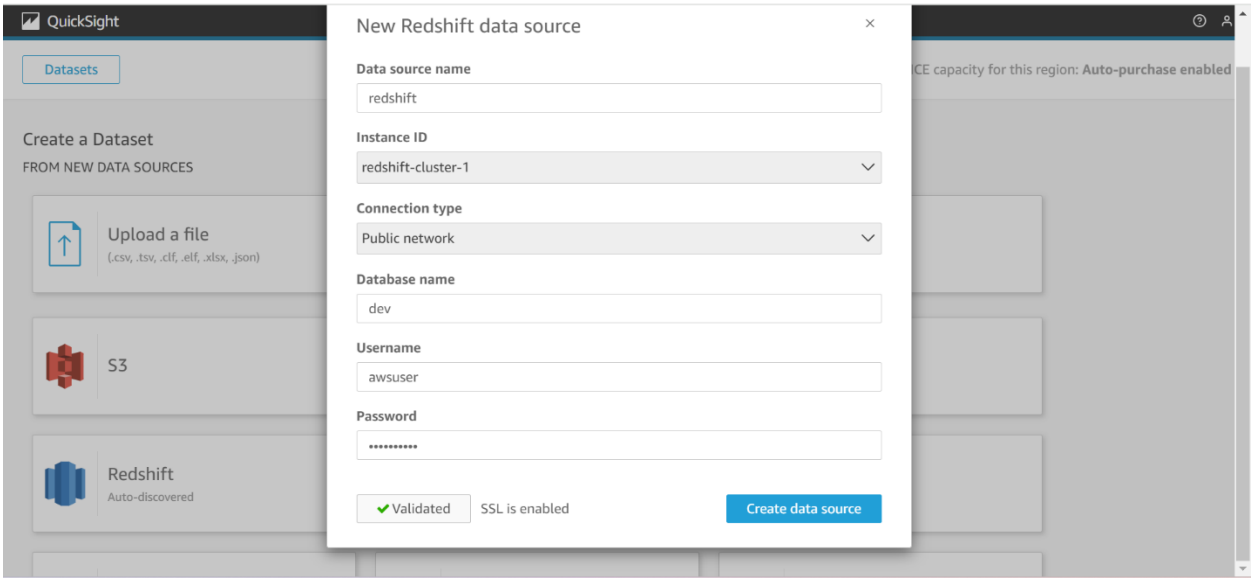
For dashboarding we will be using AWS QuickSight Service

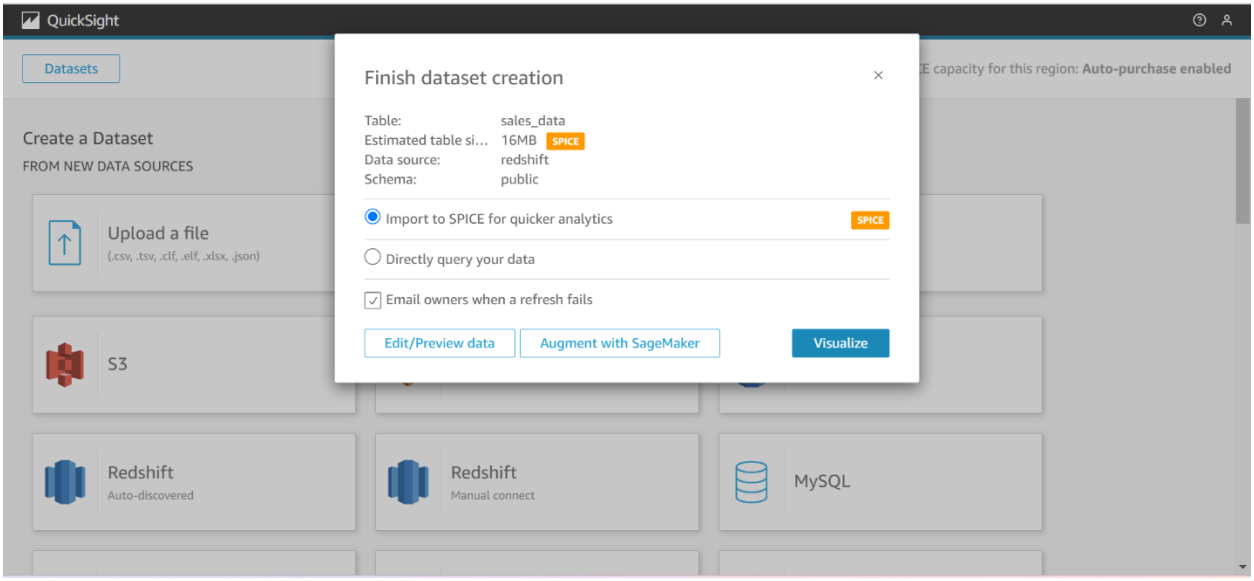
We first start importing data to QuickSight from Redshift Database

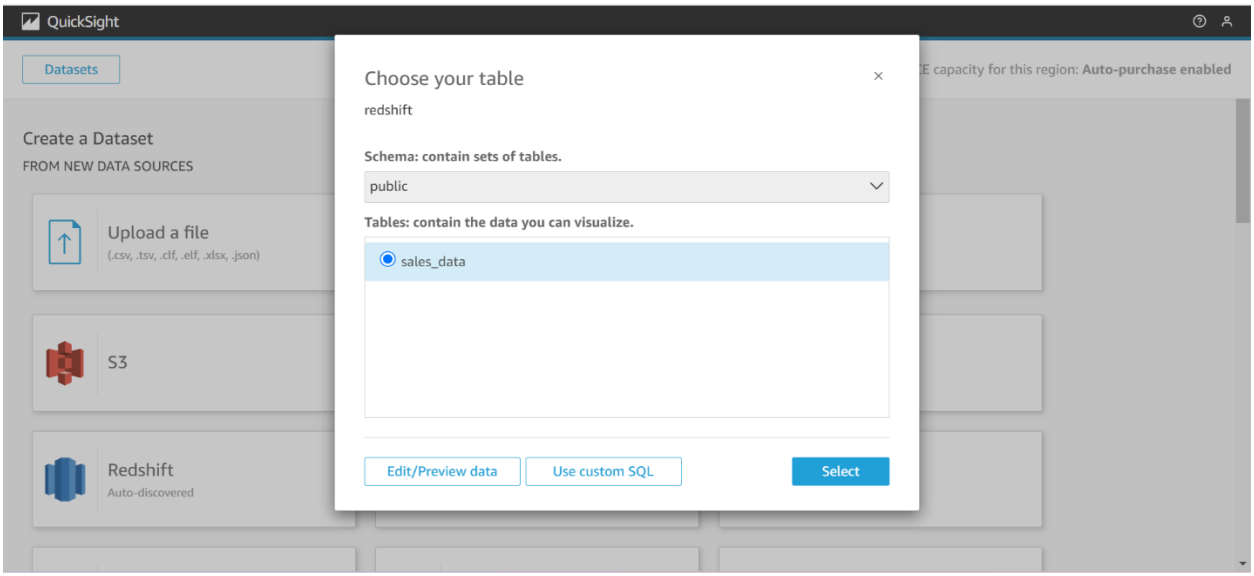


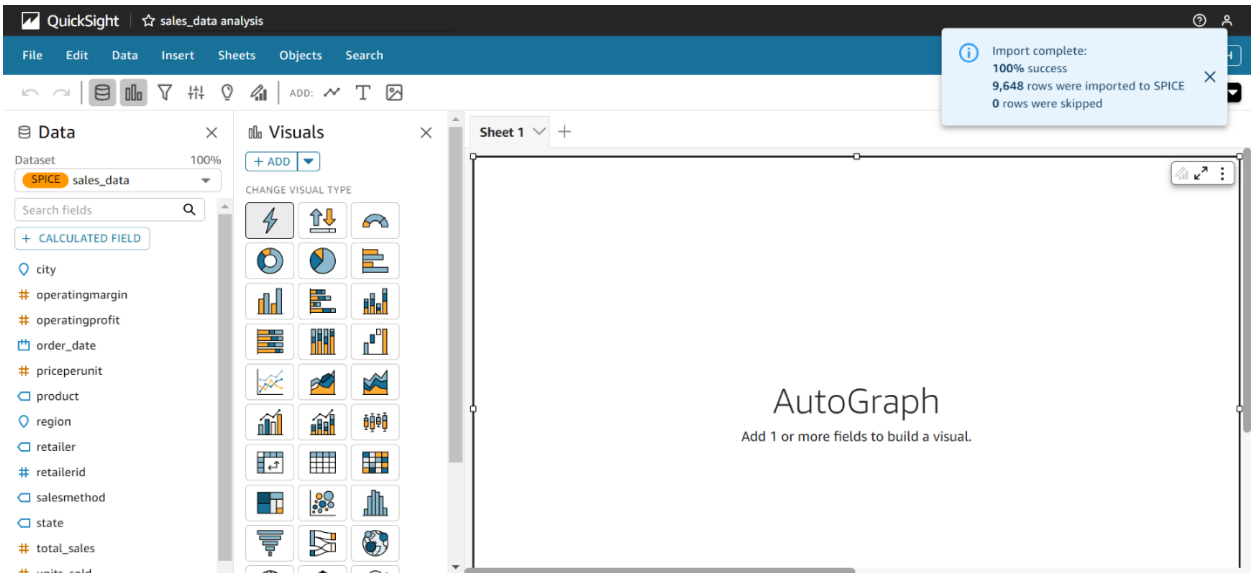


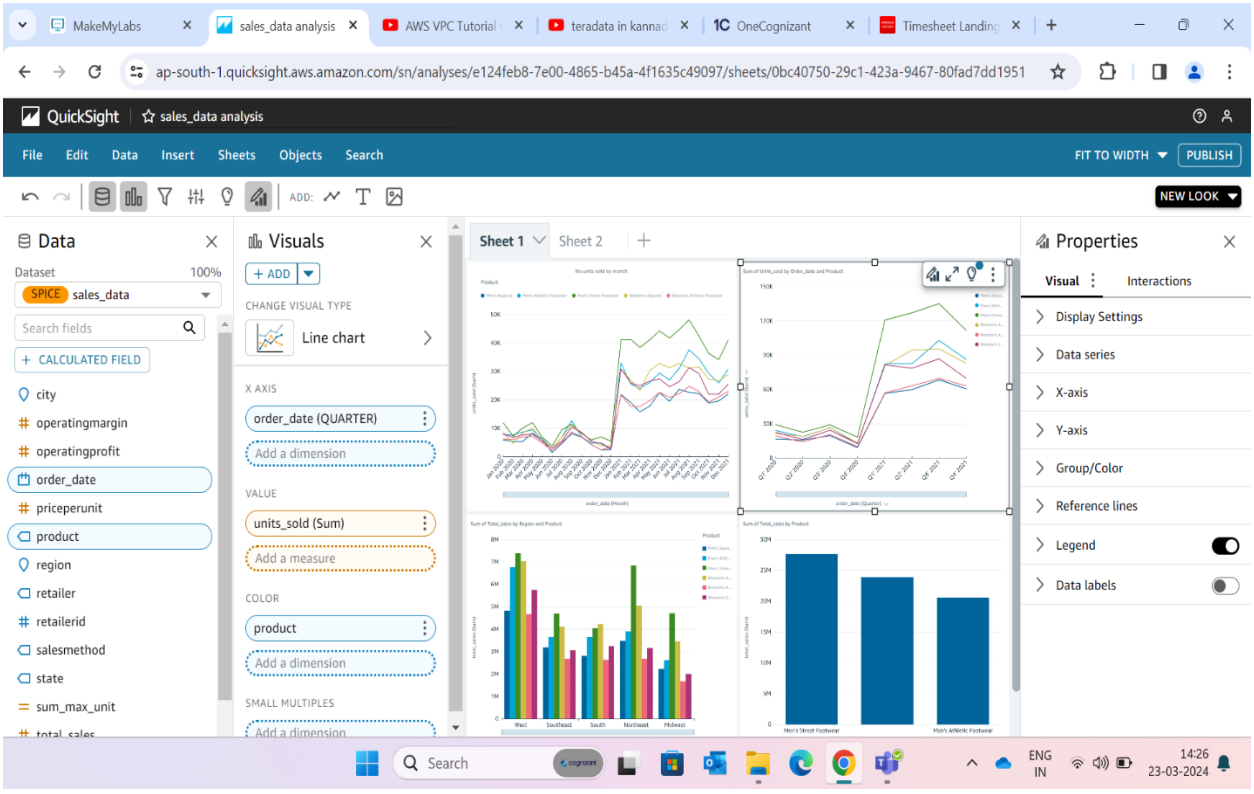

QuickSight
Datasets
SPICE capacity for this region: Auto-purchase enabled
Create a Dataset
FROM NEW DATA SOURCES
Upload a file
(.csv, .tsv, .clf, .elf, .xlsx, .json)
Salesforce
Connect to Salesforce
S3 Analytics
S3
Athena
RDS
Redshift
Auto-discovered
Redshift
Manual connect
MySQL

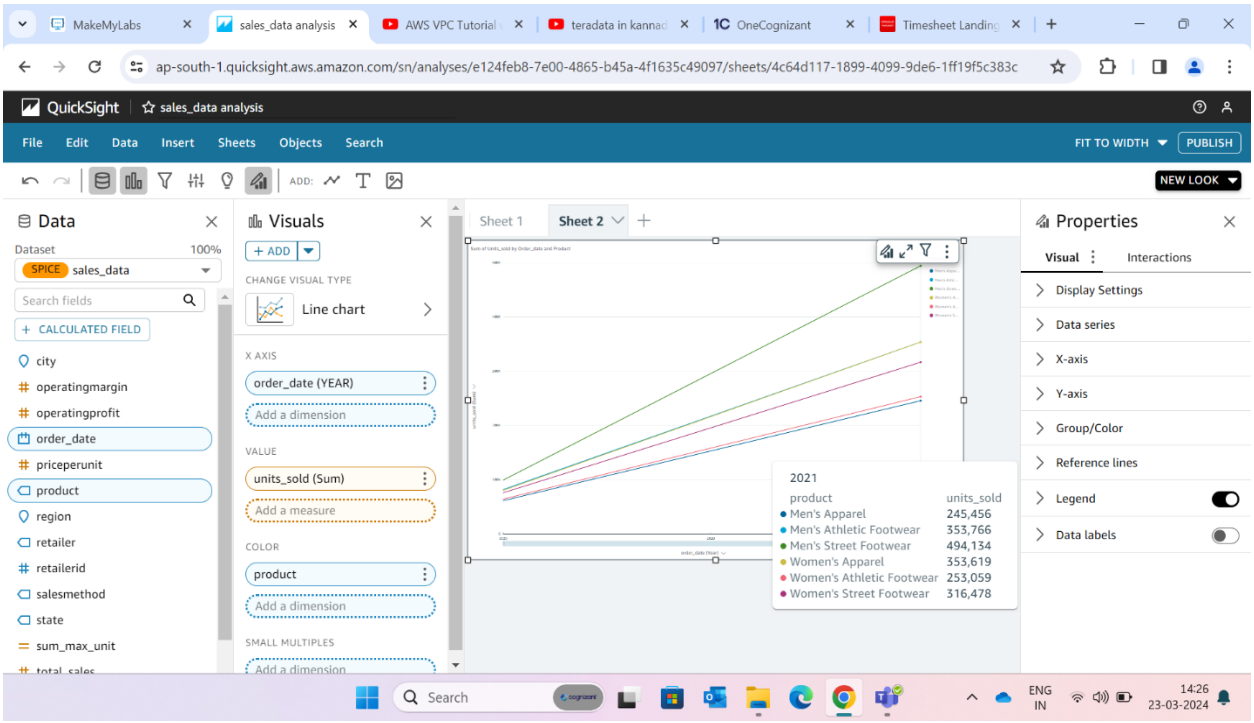

QuickSight
Datasets
Create a Dataset
FROM NEW DATA SOURCES
Upload a file
(.csv, .tsv, .clf, .elf, .xlsx, .json)
S3
Redshift
Auto-discovered
New Redshift data source
Data source name
redshift
Instance ID
redshift-cluster-1
Connection type
Public network
Database name
dev
Username
awsuser
Password
Validated
SSL is enabled
Create data source


QuickSight
Datasets
Create a Dataset
FROM NEW DATA SOURCES
Upload a file
(.csv, .tsv, .clf, .elf, .xlsx, .json)
S3
Redshift
Auto-discovered
Redshift
Manual connect
MySQL
Finish dataset creation
Table: sales_data
Estimated table si... 16MB SPICE
Data source: redshift
Schema: public
Import to SPICE for quicker analytics
Directly query your data
Email owners when a refresh fails
Edit/Preview data
Augment with SageMaker
Visualize

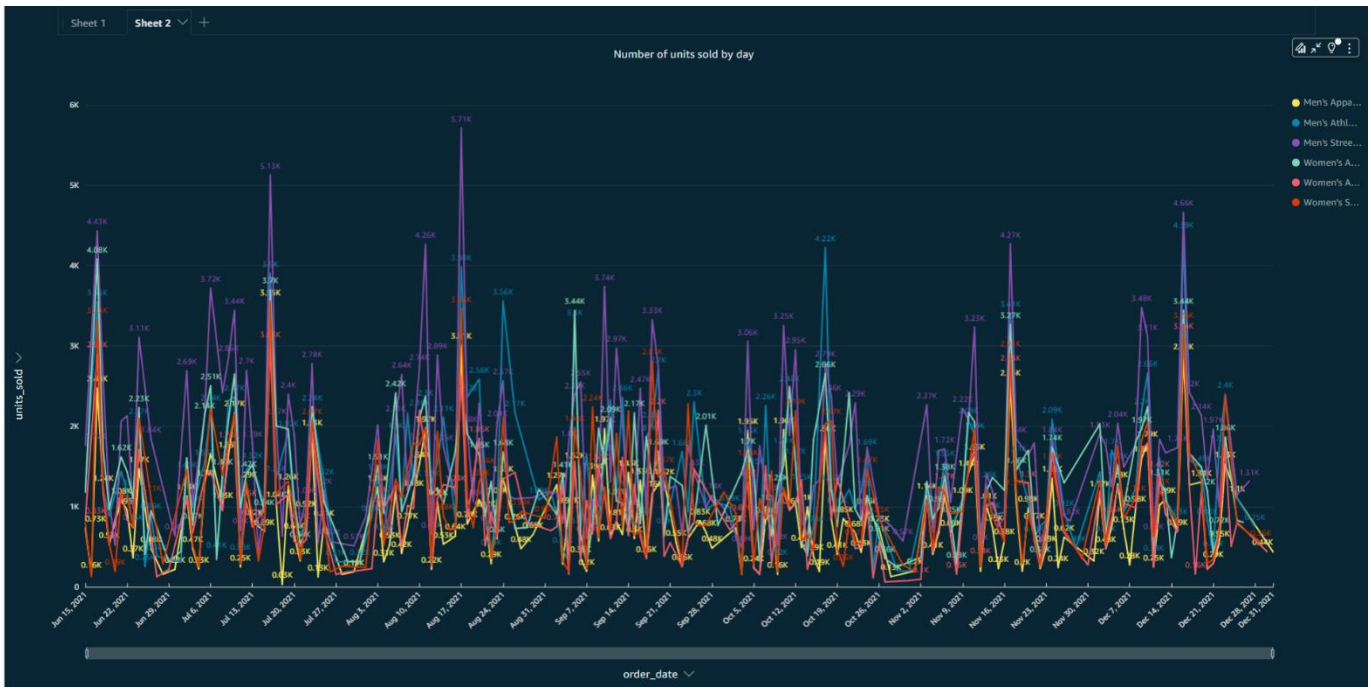

QuickSight
Datasets
Create a Dataset
FROM NEW DATA SOURCES
Upload a file
(.csv, .tsv, .clf, .elf, .xlsx, .json)
S3
Redshift
Auto-discovered
Choose your table
redshift
Schema: contain sets of tables.
public
Tables: contain the data you can visualize.
sales_data
Edit/Preview data
Use custom SQL
Select


QuickSight
sales_data analysis
File Edit Data Insert Sheets Objects Search
Import complete:
100% success
9,648 rows were imported to SPICE
0 rows were skipped
Data
Dataset 100%
SPICE sales_data
Search fields
CALCULATED FIELD
city
operatingmargin
operatingprofit
order_date
priceperunit
product
region
retailer
retailerid
salesmethod
state
total_sales
Visuals
ADD
CHANGE VISUAL TYPE
Sheet 1
AutoGraph
Add 1 or more fields to build a visual.


MakeMyLabs
sales_data analysis
AWS VPC Tutorial
teradata in kannad
OneCognizant
Timesheet Landing
ap-south-1.quicksight.aws.amazon.com/sn/analyses/e124feb8-7e00-4865-b45a-4f1635c49097/sheets/0bc40750-29c1-423a-9467-80fad7dd1951
QuickSight
sales_data analysis
File Edit Data Insert Sheets Objects Search
FIT TO WIDTH
PUBLISH
NEW LOOK
Data
Dataset 100%
SPICE sales_data
Search fields
CALCULATED FIELD
city
operatingmargin
operatingprofit
order_date
priceperunit
product
region
retailer
retailerid
salesmethod
state
sum_max_unit
Visuals
ADD
CHANGE VISUAL TYPE
Line chart
X AXIS
order_date (QUARTER)
Add a dimension
VALUE
units_sold (Sum)
Add a measure
COLOR
product
Add a dimension
SMALL MULTIPLES
Add a dimension
Sheet 1
Sheet 2
Properties
Visual
Interactions
Display Settings
Data series
X-axis
Y-axis
Group/Color
Reference lines
Legend
Data labels
Search
ENG
IN
14:26
23-03-2024

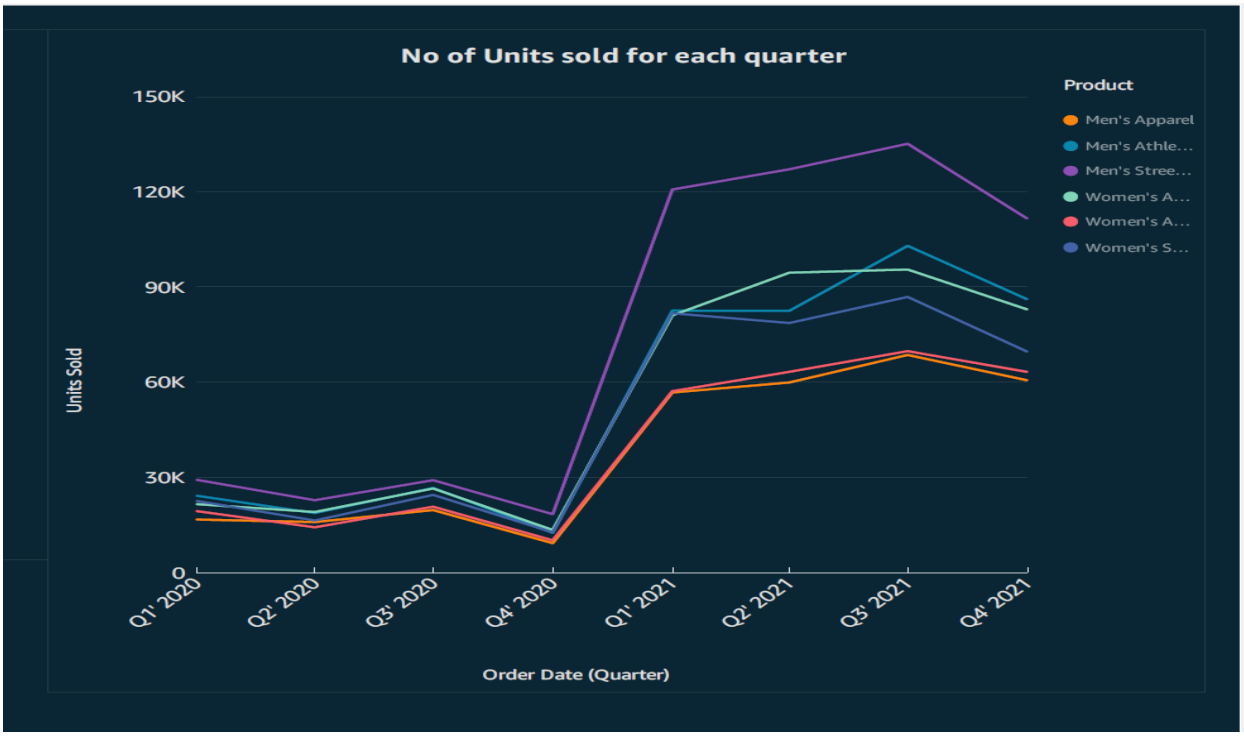


Plotting no of units sold for each day

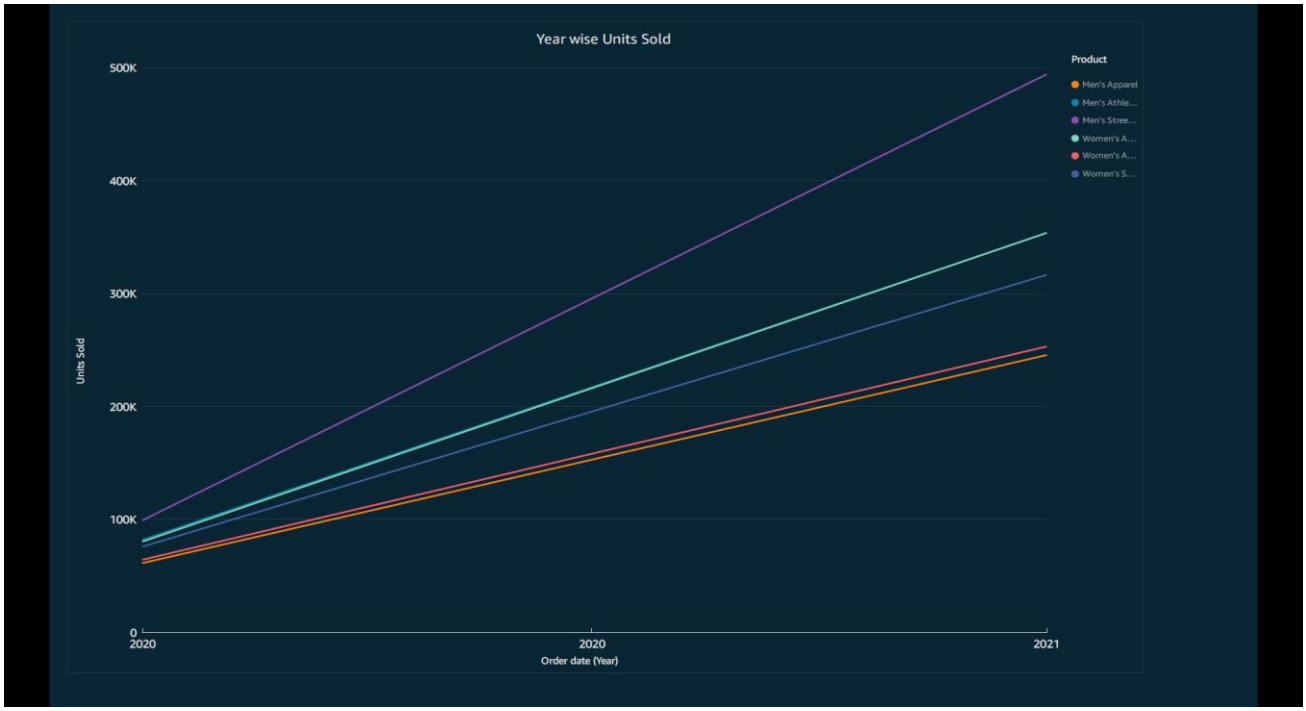


Plotting no of units sold for each month

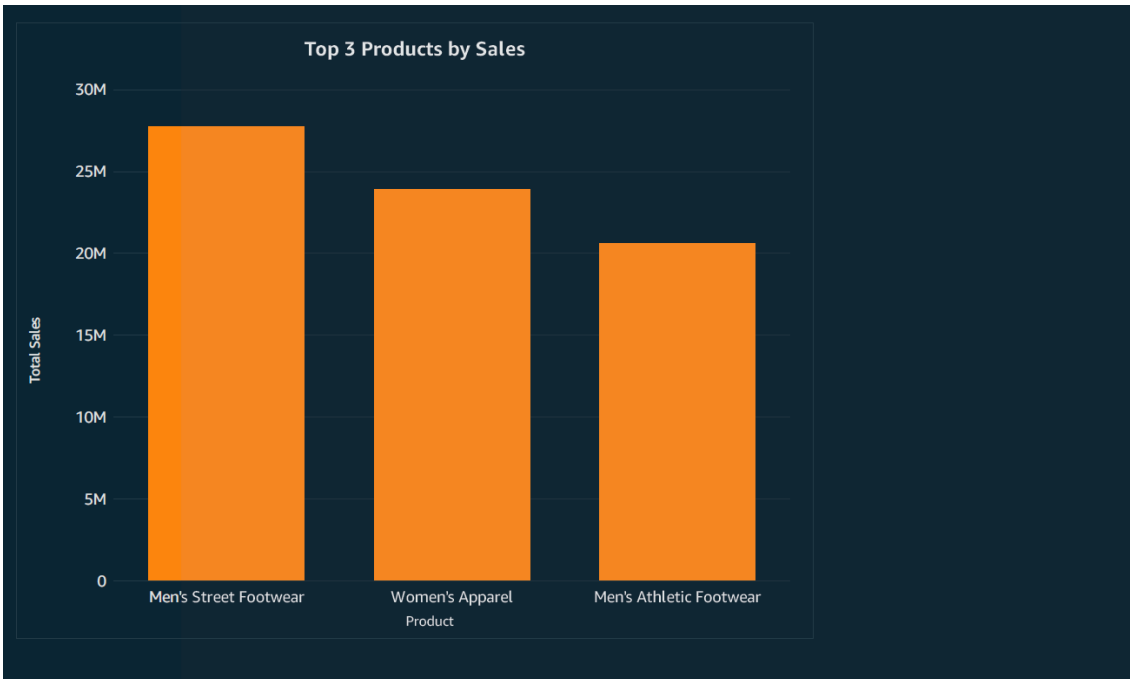
Quarterly wise units sold



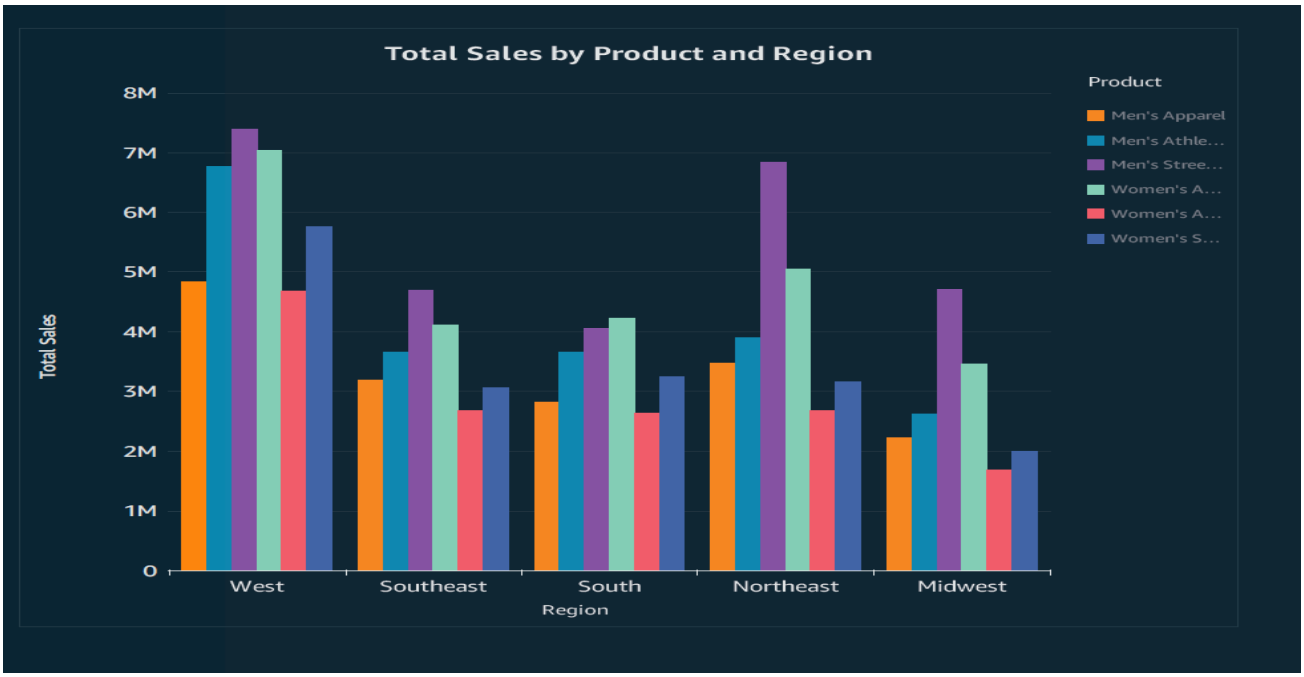
Year wise units sold



Top 3 products by sales



Region wise product sales



Conclusion

1. Data Warehousing with Amazon Redshift: By leveraging Amazon Redshift, we successfully centralized and managed the sales data. Redshift's scalability and performance allowed us to handle large volumes of information efficiently. We transformed raw data into a structured format, enabling complex queries and aggregations. The data warehouse served as the backbone for our analysis, providing a solid foundation for further insights.
2. Visual Insights with Amazon Quick Sight: Amazon Quick Sight played a pivotal role in turning data into actionable insights. We created interactive dashboards and visualizations that allowed stakeholders to explore sales trends, product performance, and customer behavior. Quick Sight's user-friendly interface empowered business users to make informed decisions without relying on technical expertise. Whether tracking revenue, monitoring inventory, or identifying growth opportunities, Quick Sight provided a responsive platform for data exploration.

   "Armed with Redshift and Quick Sight, the sales team can gain valuable insights. They can identify top-performing products, can analyze regional sales patterns, and optimize inventory management. By embracing AWS tools, they can elevate their sales analysis capabilities, driving smarter business decisions and enhancing overall performance".