

Machine Learning with AWS Hackathon

Team-6

A.Pranavya -21251A6602

M.Satwika- 21251A6659

P.Lakshmi Harini -21251A6626

S.Saisree- 21251A6660

P.Sathwika- 21251A0521

AWS(Amazon Web Services)

Amazon Web Services offers a broad set of global cloud-based products including compute, storage, databases, analytics, networking, mobile, developer tools, management tools, IoT, security, and enterprise applications: on-demand, available in seconds, with pay-as-you-go pricing. From data warehousing to deployment tools, directories to content delivery, over 200 AWS services are available. New services can be provisioned quickly, without the upfront fixed expense.

In 2006, Amazon Web Services (AWS) began offering IT infrastructure services to businesses as web services—now commonly known as cloud computing.

Cloud Computing

The term cloud refers to a network or the internet. It is a technology that uses remote servers on the internet to store, manage, and access data online without installing and maintaining them on-premises. The data can be anything such as files, images, documents, audio, video, and more.

Benefits of cloud computing

- Cost Savings.
- Security.
- Flexibility.
- Mobility.
- Insight.
- Increased Collaboration.
- Quality Control.
- Disaster Recovery.
- Loss Prevention.
- Automatic Software Updates.

Tools used on AWS

- Amazon VPC
- Amazon EC2
- Amazon Subnets
- Network security groups
- Network Access Control List (NACL)
- Network Address Translator (NAT)

Companies using AWS

AWS is used in wide range of sectors and companies. Amazon and Netflix are the companies which majorly dependent on AWS for its services. Few other companies which depend on AWS are:

- Airbnb
- Spotify
- Intuit
- Stripe
- Pfizer
- BMW
- The New York Times
- NASA
- Duolingo
- Samsung

ADVANTAGES OF AWS:

- User-friendly
- Flexible
- Secure
- Cost-effective
- Reliable
- Scalable and Elastic
- Highly Performant

LEVEL-1

Task 1 :

NETWORKING

To develop a VPC with web server in a public subnet and database in a private subnet. And then connect an internet gateway to the public subnet , access the database through the jump box, connect the internet to the private subnet through NAT and attach NACLs to both the subnets.

VPC(Virtual Private Cloud)

VPC (Virtual Private Cloud) is a fundamental networking feature provided by cloud service providers like Amazon Web Services (AWS) and others. It allows users to create isolated, private network environments within the cloud infrastructure.

VPCs are used to segment and isolate resources within a cloud provider's infrastructure, enabling better control over networking, security, and resource management.

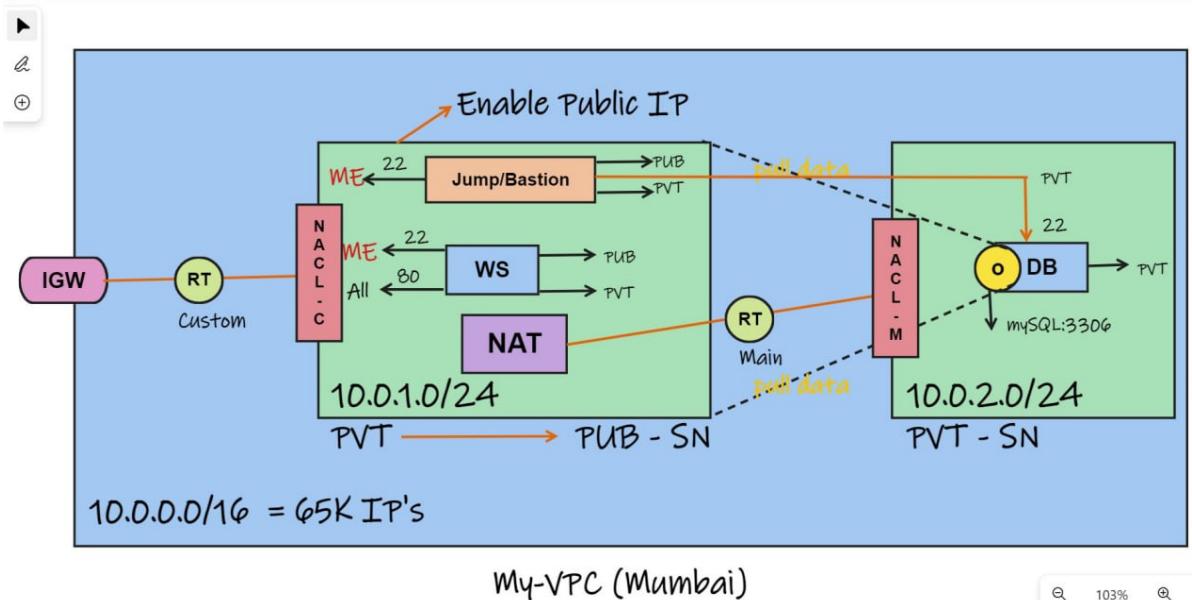
Advantages

- Isolation and security
- Customisation and control
- Scalability

and

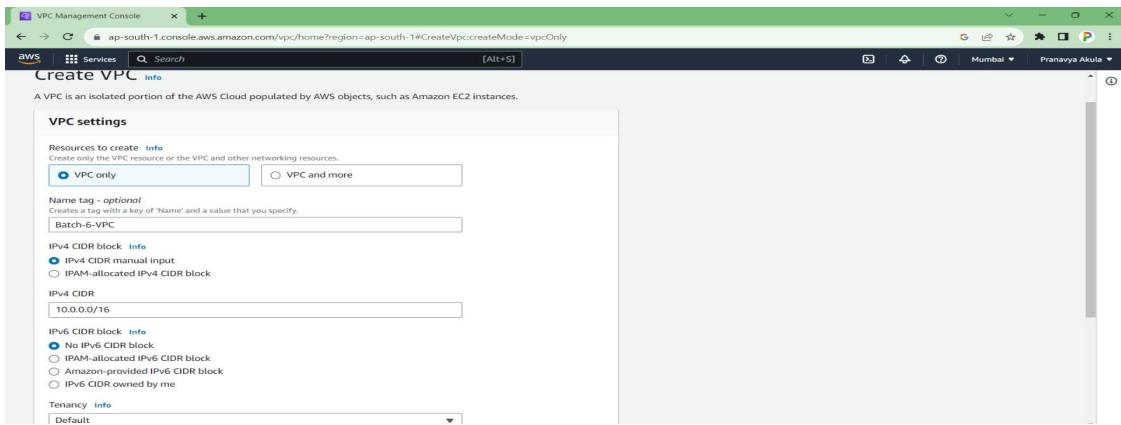
high

availability



CREATING VPC(VIRTUAL PRIVATE CLOUD):

- Open VPC Portal in aws console and click on create VPC.
- When we create the VPC we will also get the DHCP option set, main route table and main network ACL along with VPC.



Detail	Value
VPC ID	vpc-095c6a67d2a15ad5b
State	Available
DHCP option set	dopt-0c03473738be1ea01
Default VPC	No
IPv4 CIDR	10.0.0.0/16
Network Address Usage metrics	Disabled
Main route table	rtb-0e00cdf41b90ce348
IPv6 pool	-
Owner ID	153334110048
DNS resolution	Enabled
Main network ACL	acl-0a7fsb3c335f50f82
IPv6 CIDR (Network border group)	-

- Then select “subnets” from the available options and create two subsets with the following details and name one as “pub” and other as “pvt” with proper availability zones.
- The public subnet has the CIDR 10.0.1.0/24 whereas the private subnet has 10.0.2.0/24.

VPC Management Console

Create subnet [Info](#)

VPC

VPC ID
Create subnets in this VPC.
vpc-095c6a67d2a15ad5b (Batch-6-VPC)

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 2

Subnet name
Create a tag with a key of 'Name' and a value that you specify.
10.0.1.0/24-1a-Pub

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

CloudShell Feedback Language © 2023, Amazon Web Services India Private Limited or its affiliates. Privacy Terms Cookie preferences 25°C Cloudy 13:04 ENG IN 19-08-2023

VPC Management Console

Create subnet [Info](#)

IPv4 CIDR block [Info](#)
Q 10.0.1.0/24 X

Tags - optional

Key Value - optional
Name 10.0.1.0/24-1a-Pub Remove

Add new tag You can add 49 more tags. Remove

Subnet 2 of 2

Subnet name
Create a tag with a key of 'Name' and a value that you specify.
10.0.2.0/24-1b-Pvt

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-1b

IPv4 CIDR block [Info](#)
Q 10.0.2.0/24 X

Tags - optional

Key Value - optional

CloudShell Feedback Language © 2023, Amazon Web Services India Private Limited or its affiliates. Privacy Terms Cookie preferences 25°C 13:05

Both subnets are created.

The screenshot shows the AWS VPC Management Console. On the left, a sidebar lists various VPC-related services like EC2 Global View, Subnets, Route tables, and Internet gateways. The main pane displays a table of subnets. A success message at the top says, "You have successfully created 2 subnets: subnet-0fbaf21bc9196a958, subnet-0c60e3e81a325c108". The table has columns for Name, Subnet ID, State, VPC, and IPv4 CIDR. Two subnets are listed: "10.0.1.0/24-1a-Pub" and "10.0.2.0/24-1b-Pvt", both in the "Available" state.

- By default both the subnets are private. To make a private subnet as public it has to be connected to an external cloud through an Internet Gateway. So, select and action and then edit subnet settings
- Click on edit subnet associations-->Enable Auto Assign public .

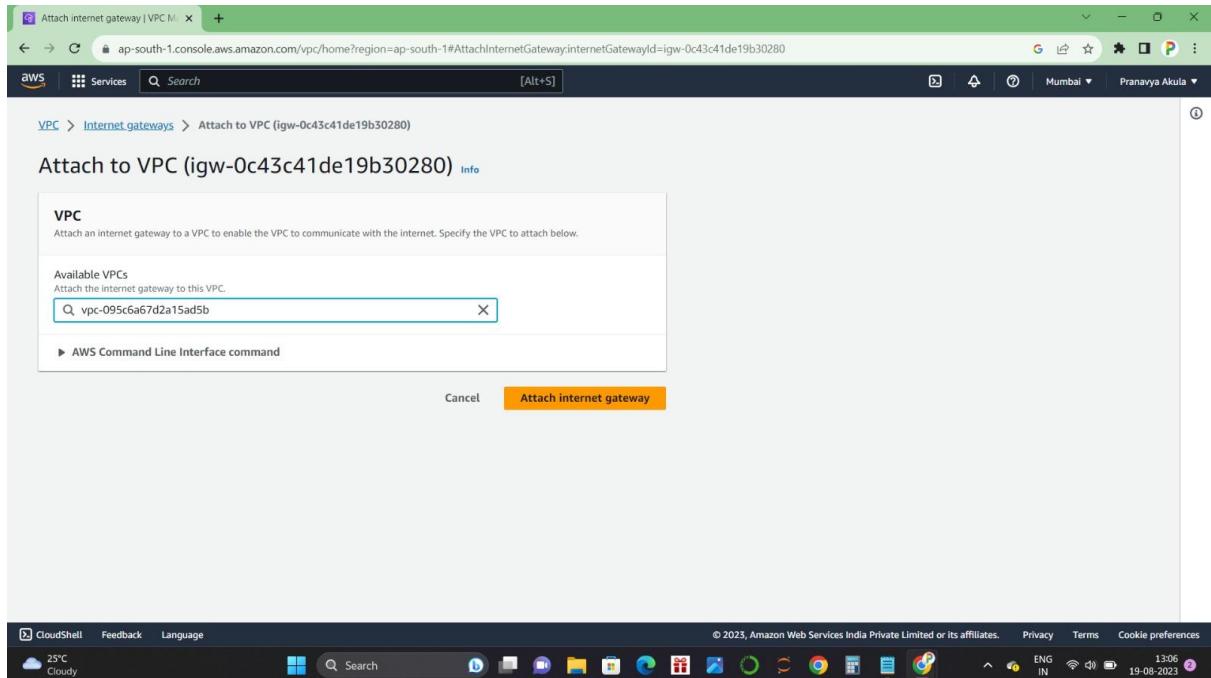
This screenshot shows the detailed configuration page for the subnet "subnet-0fbaf21bc9196a958". The left sidebar shows the navigation path: VPC > Subnets > subnet-0fbaf21bc9196a958. The main area displays the "Details" section with various configuration parameters. On the right, a context menu under the "Actions" button is open, with "Edit subnet settings" highlighted. Other options in the menu include Create flow log, Edit IPv6 CIDRs, Edit network ACL association, Edit route table association, Edit CIDR reservations, Share subnet, Manage tags, and Delete.

The screenshot shows the AWS VPC Management Console with the URL ap-south-1.console.aws.amazon.com/vpc/home?region=ap-south-1>EditSubnetSettings subnetId=subnet-0fbaf21bc9196a958. The page is titled 'Edit subnet settings' and displays the configuration for a public subnet with Subnet ID subnet-0fbaf21bc9196a958 and Name 10.0.1.0/24-1a-Pub. It includes sections for 'Auto-assign IP settings' (with 'Enable auto-assign public IPv4 address' checked), 'Resource-based name (RBN) settings' (with 'IP name' selected), and a toolbar at the bottom.

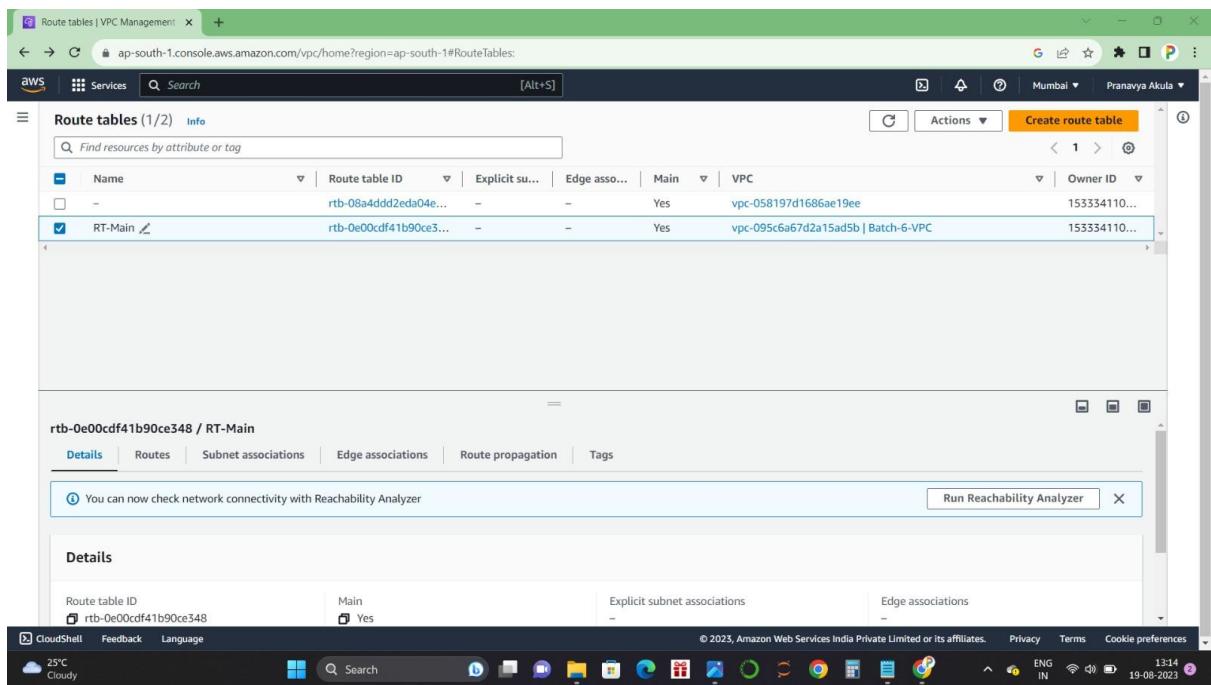
- We need to connect Internet Gateway to public subnet to establish a connection between a Virtual Private Cloud (VPC) and the wider internet, enabling resources within the VPC to communicate with external networks and services on the internet.
- For that we need to create an Internet Gateway

The screenshot shows the AWS VPC Management Console with the URL ap-south-1.console.aws.amazon.com/vpc/home?region=ap-south-1>CreateInternetGateway. The page is titled 'Create internet gateway' and provides instructions for creating a new internet gateway. It includes fields for 'Name tag' (containing 'IGW') and 'Tags - optional' (containing a single tag 'Name: IGW'). A 'Create internet gateway' button is at the bottom.

- Attach created VPC to Internet Gateway.
- Attach this internet gateway to VPC



- We have to create the route tables. Route table is necessary to determine how network traffic is directed within the VPC and between the VPC and external networks.
- For this click on route tables and click on create route table.



The screenshot shows the 'Create route table' wizard in the AWS VPC Management Console. The 'Route table settings' section is active, displaying fields for 'Name - optional' (RT-Custom) and 'VPC' (vpc-095c6a67d2a15ad5b (Batch-6-VPC)). The 'Tags' section shows a single tag 'Name: RT-Custom'. At the bottom right is the 'Create route table' button.

- After creating Route table edit subnet associations.
- Select public subnet and save associations.
- Then edit routes set destination 0.0.0.0/0 and Target as igw and save changes

The screenshot shows the 'Edit subnet associations' page for route table rtb-05eb0250e7bc4a286. It lists available subnets (10.0.1.0/24-1a-Pub and 10.0.2.0/24-1b-Pvt) and selected subnets (10.0.1.0/24-1a-Pub). The 'Save associations' button is visible at the bottom right.

Name	Subnet ID	IPv4 CIDR	IPv6 CIDR	Route table ID
10.0.1.0/24-1a-Pub	subnet-0fbaf21bc9196a958	10.0.1.0/24	-	Main (rtb-0e00ccdf41b90ce348 / RT-Ma..)
10.0.2.0/24-1b-Pvt	subnet-0c60e3e81a325c108	10.0.2.0/24	-	Main (rtb-0e00ccdf41b90ce348 / RT-Ma..)

VPC Management Console

Services Search [Alt+S]

Mumbai Pranavya Akula

VPC > Route tables > rtb-05eb0250e7bc4a286 > Edit routes

Edit routes

Destination	Target	Status	Propagated
10.0.0.0/16	local	Active	No
0.0.0.0/0	igw-0c43c41de19b30280	-	No

Add route Cancel Preview Save changes

- GO to EC2 and launch the instances as web-server
- Create a new key pair

VPC Management Console

Launch an instance | EC2 Manager

CloudShell Feedback Language

© 2023, Amazon Web Services India Private Limited or its affiliates. Privacy Terms Cookie preferences

Cloud Cloudy ENG IN 13:16 19-08-2023

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: web-server 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: linux

AMI from catalog Quick Start

Amazon Machine Image (AMI)

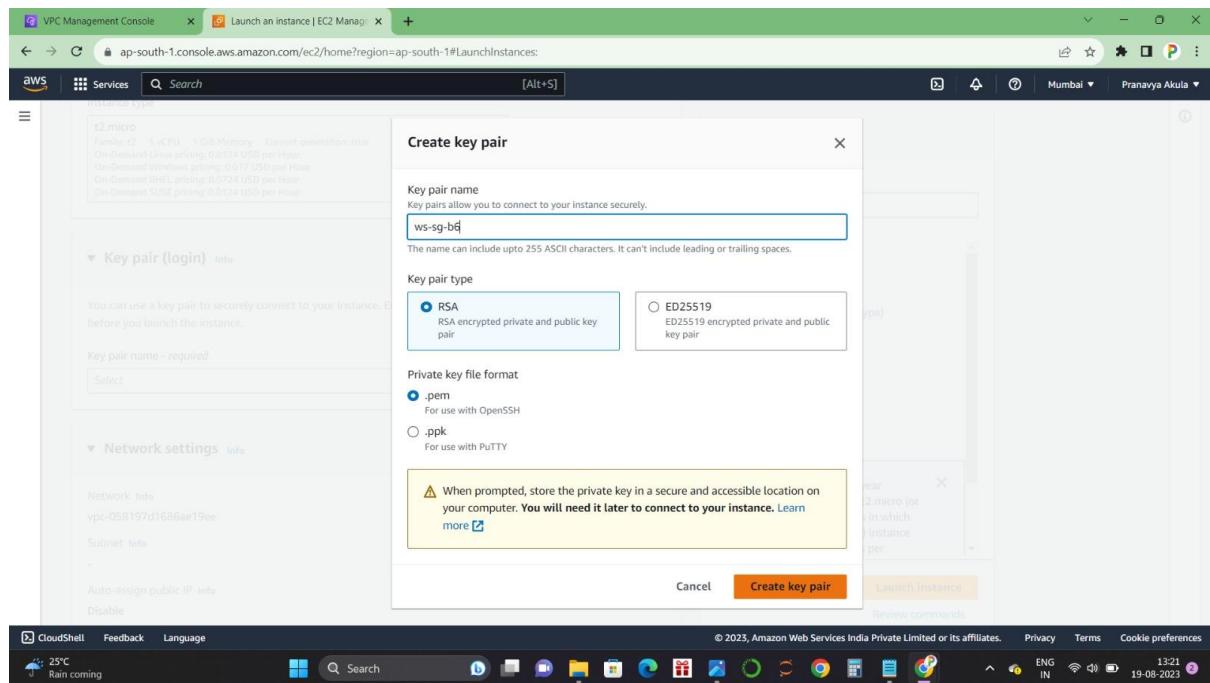
al2023-ami-2023.1.20230809.0-kernel-6.1-x86_64 ami-0da59f1af71ea4ad2

Verified provider Free tier eligible

Browse more AMIs

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

Cancel Launch instance Review commands



- Edit Network settings select VPC and select public subnet. Add security group name and description.
- Edit and add security group rules for web server.
- In advanced details add Bin bash commands
- Web Server instance is launched!

VPC Management Console | Launch an instance | EC2 Manager

ap-south-1.console.aws.amazon.com/ec2/home?region=ap-south-1#LaunchInstances:

Services Search [Alt+S]

Pranaya Akula Mumbai

Network settings

VPC - required Info: VPC: vpc-095c6a67d2a15ad5b (Batch-6-VPC) 10.0.0.0/16

Subnet Info: Subnet: subnet-0fbaf21bc9196a958 VPC: vpc-095c6a67d2a15ad5b Owner: 15334110048 Availability Zone: ap-south-1a IP addresses available: 251 CIDR: 10.0.1.0/24

Auto-assign public IP Info: Enable

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

Security group name - required: ws-sg-b6

This security group will be added to all network interfaces. The name can't be edited after the security group is created. Max length is 255 characters. Valid characters: a-z, A-Z, 0-9, spaces, and _-~/@[]+=&{|\$^

Description - required Info: ws-sg-b6

Inbound Security Group Rules

CloudShell Feedback Language 25°C Near record Search ENG IN 13:23 19-08-2023

Number of instances Info: 1

Software Image (AMI): Amazon Linux 2023 AMI ami-0da59f1af71ea4ad2

Virtual server type (instance type): t2.micro

Firewall (security group): New security group

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 year

Launch instance Review commands

© 2023, Amazon Web Services India Private Limited or its affiliates. Privacy Terms Cookie preferences

VPC Management Console | Launch an instance | EC2 Manager

ap-south-1.console.aws.amazon.com/ec2/home?region=ap-south-1#LaunchInstances:

Services Search [Alt+S]

Pranaya Akula Mumbai

Security group rule 1 (TCP, 22, 43.225.26.106/52, admin)

Type Info: ssh Protocol Info: TCP Port range Info: 22

Source type Info: My IP Name Info: 43.225.26.106/32 Description - optional Info: admin

Security group rule 2 (TCP, 80, Multiple sources, all)

Type Info: HTTP Protocol Info: TCP Port range Info: 80

Source type Info: Anywhere Source Info: 0.0.0.0/0 Description - optional Info: all

⚠️ Rules with source of 0.0.0.0/0 allow all IP addresses to access your instance. We recommend setting security group rules to allow access from known IP addresses only.

Add security group rule Advanced network configuration

CloudShell Feedback Language 25°C Near record Search ENG IN 13:24 19-08-2023

Number of instances Info: 1

Software Image (AMI): Amazon Linux 2023 AMI ami-0da59f1af71ea4ad2

Virtual server type (instance type): t2.micro

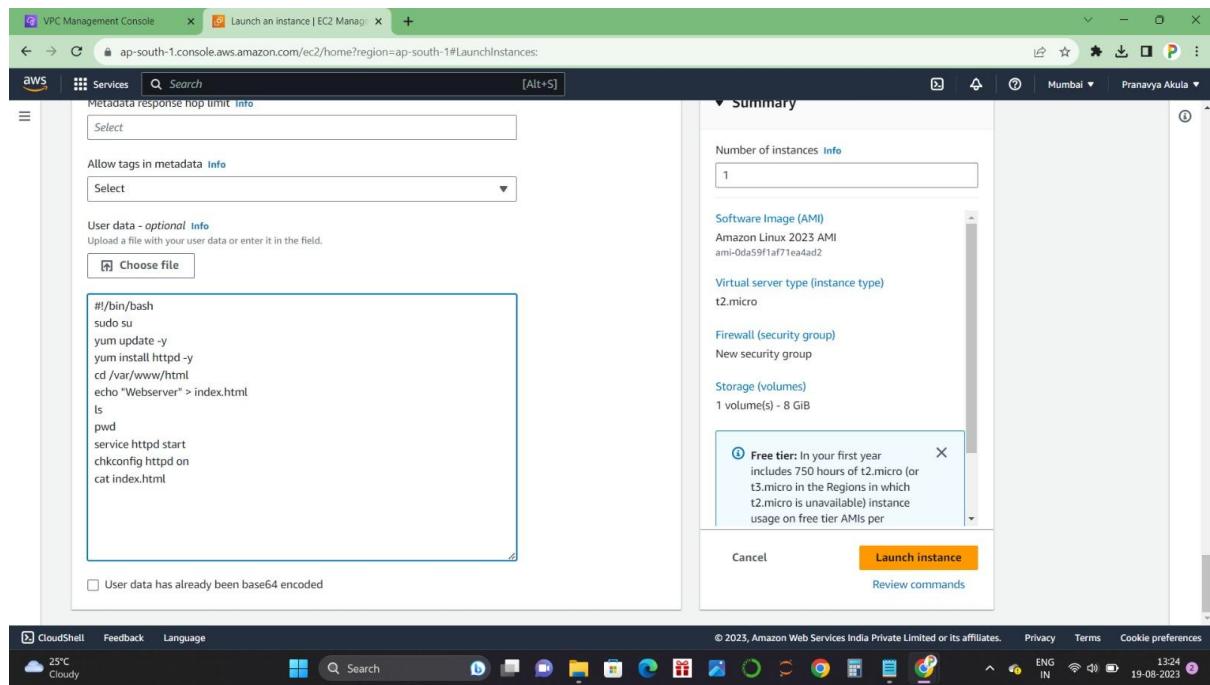
Firewall (security group): New security group

Storage (volumes): 1 volume(s) - 8 GiB

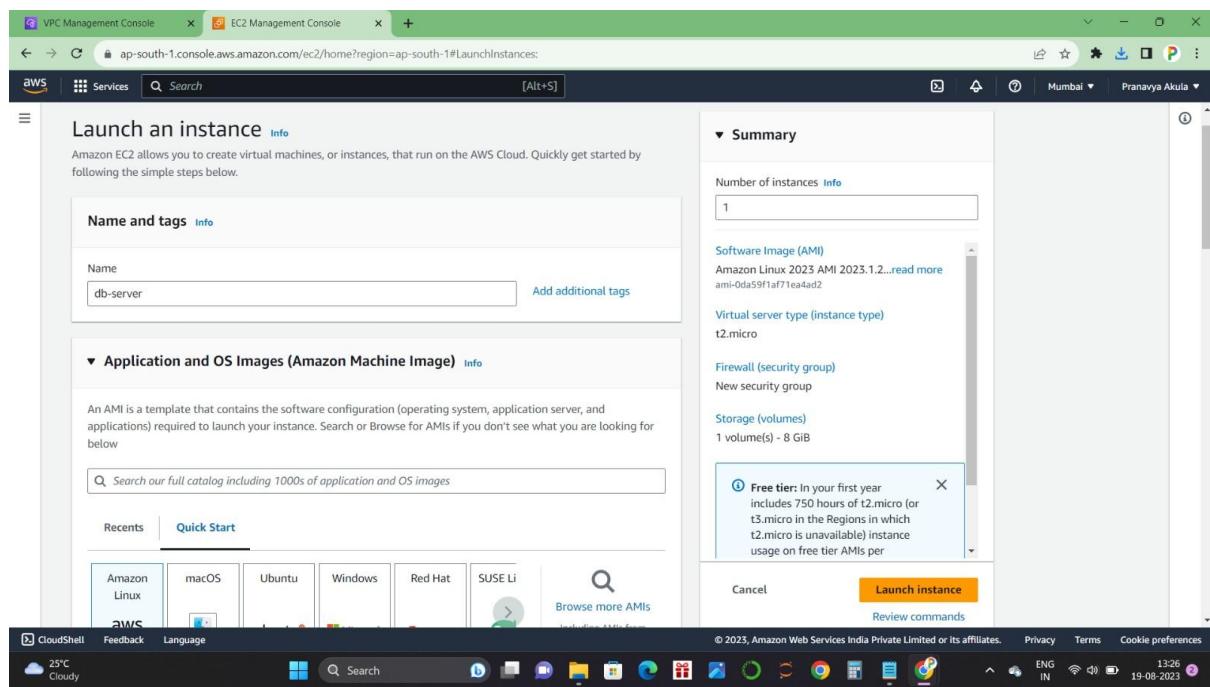
Free tier: In your first year includes 750 hours of t2.micro (or t3.micro in the Regions in which t2.micro is unavailable) instance usage on free tier AMIs per year

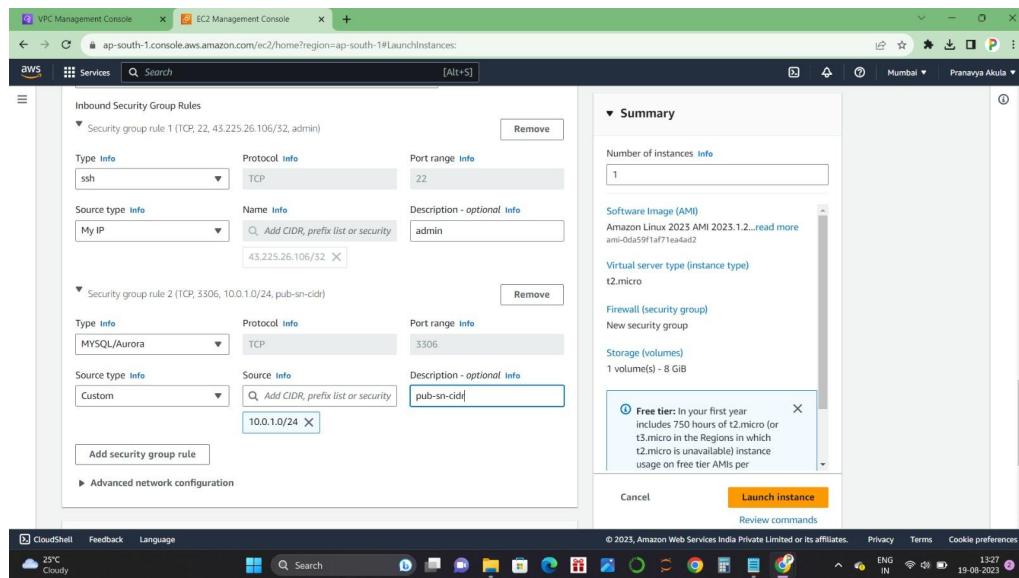
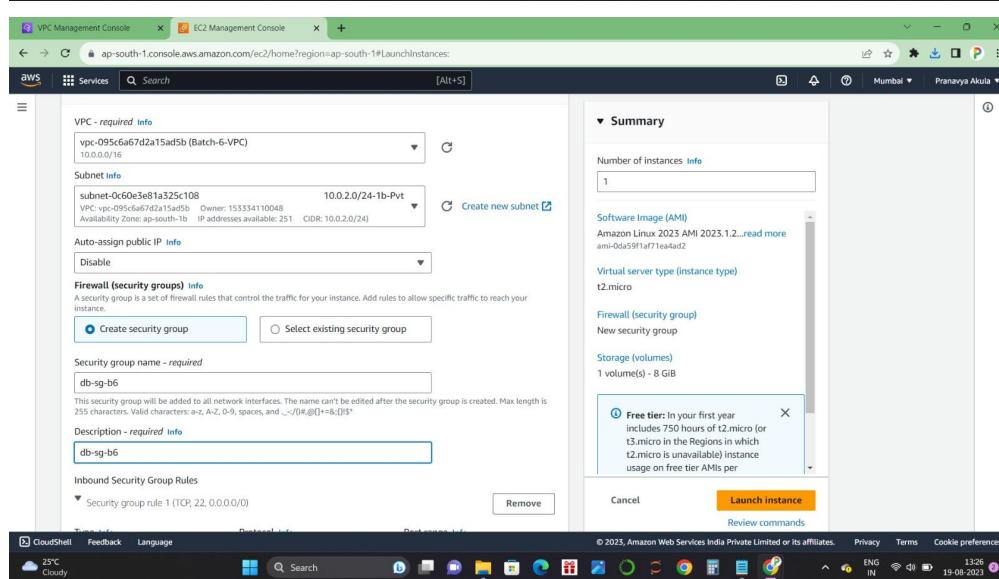
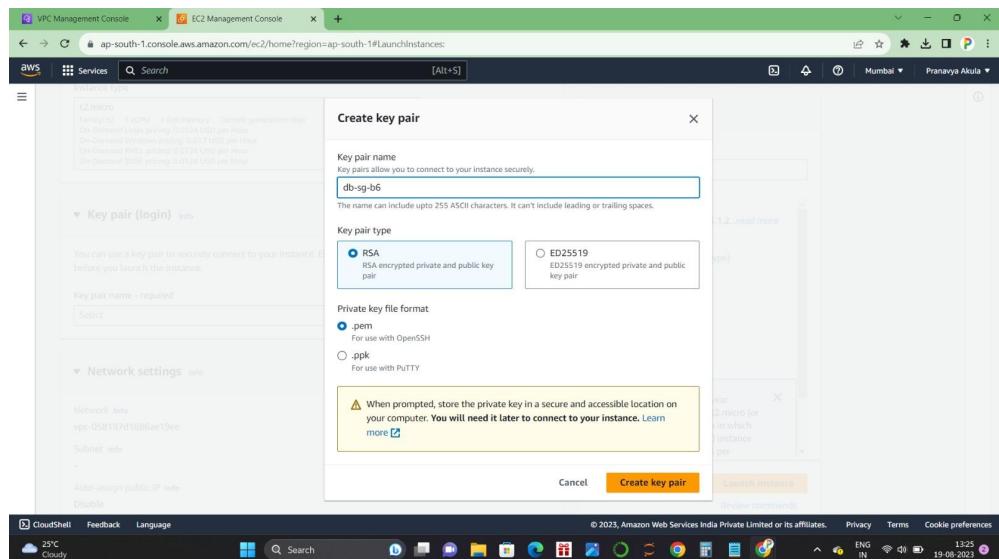
Launch instance Review commands

© 2023, Amazon Web Services India Private Limited or its affiliates. Privacy Terms Cookie preferences



Now launch another instance Database server,jump server.same as web-server





VPC Management Console X EC2 Management Console X 3.110.170.213 X +

ap-south-1.console.aws.amazon.com/ec2/home?region=ap-south-1#LaunchInstances:

AWS Services Search [Alt+S]

Pranavya Akula Mumbai

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: jump-box 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

Recent OS Images: Amazon Linux, macOS, Ubuntu, Windows, Red Hat, SUSE Linux Enterprise Server

Browse more AMIs Including AMIs from AWS Marketplace and AWS Lambda

Virtual server type (instance type) t2.micro

Firewall (security group) New security group

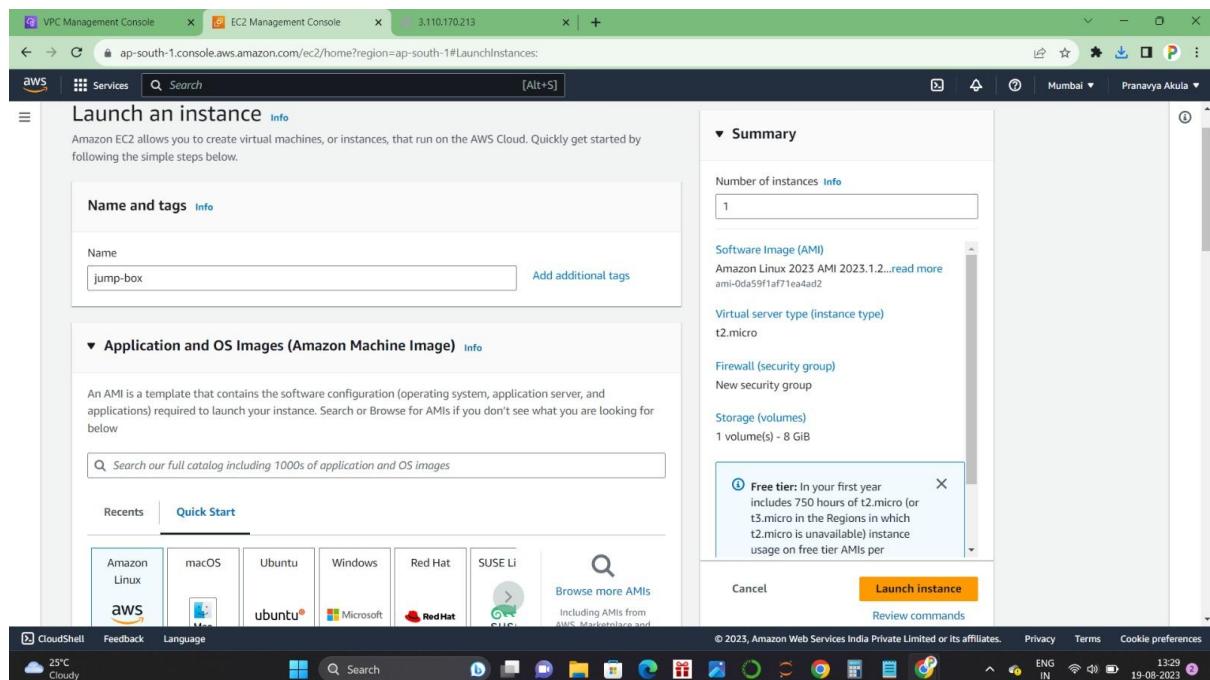
Storage (volumes) 1 volume(s) - 8 GiB

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 year

Launch instance Review commands

© 2023, Amazon Web Services India Private Limited or its affiliates. Privacy Terms Cookie preferences ENG IN 13:29 19-08-2023

CloudShell Feedback Language 25°C Cloudy



VPC Management Console X EC2 Management Console X 3.110.170.213 X +

ap-south-1.console.aws.amazon.com/ec2/home?region=ap-south-1#LaunchInstances:

AWS Services Search [Alt+S]

Pranavya Akula Mumbai

Instance type Info

Instance type: t2.micro

Family t2 - 1 vCPU - 1 GiB Memory - Current generation: true

On-Demand Linux pricing: 0.0124 USD per Hour

On-Demand Windows pricing: 0.0171 USD per Hour

On-Demand RHEL pricing: 0.0224 USD per Hour

On-Demand SUSE pricing: 0.0124 USD per Hour

Key pair (login) Info

You can use a key pair to securely connect to your instance before you launch the instance.

Key pair name - required Select

Network settings Info

Network: vpc-058197d1586ae19ee

Subnet: info

Create key pair

Key pair name: jump-b6

Key pairs allow you to connect to your instance securely.

Key pair type: RSA (selected) ED25519

RSA encrypted private and public key pair

ED25519 encrypted private and public key pair

Private key file format: .pem (selected) .ppk

.pem For use with OpenSSH

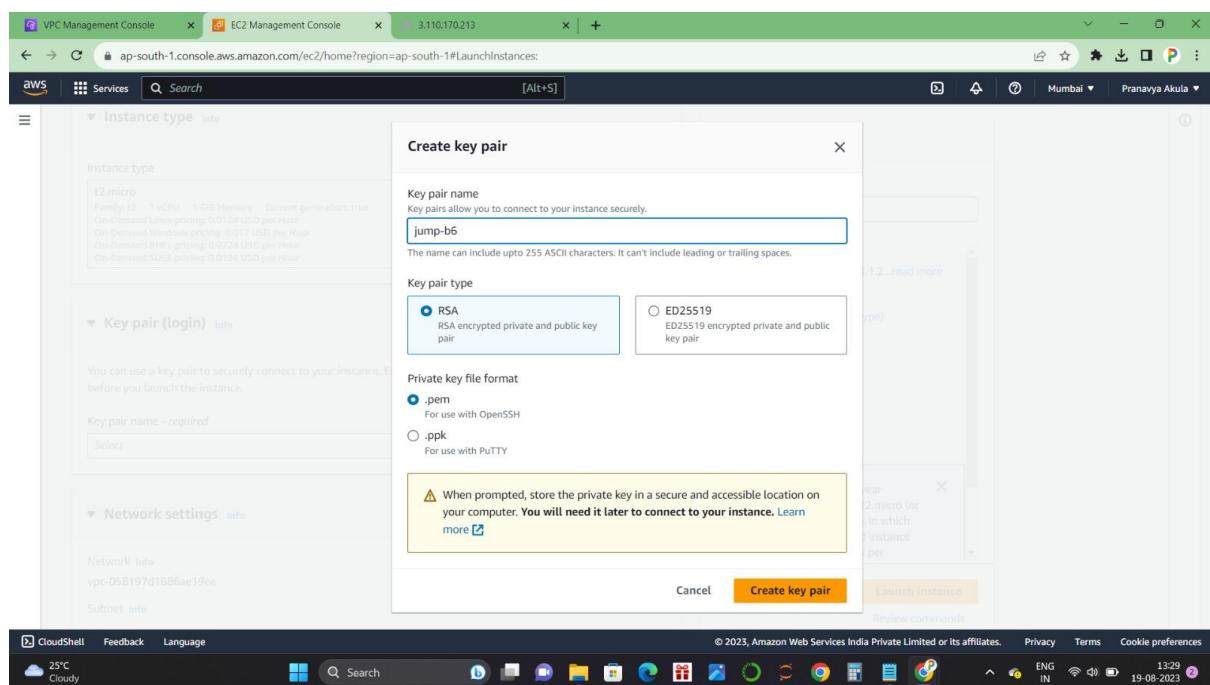
.ppk For use with PuTTY

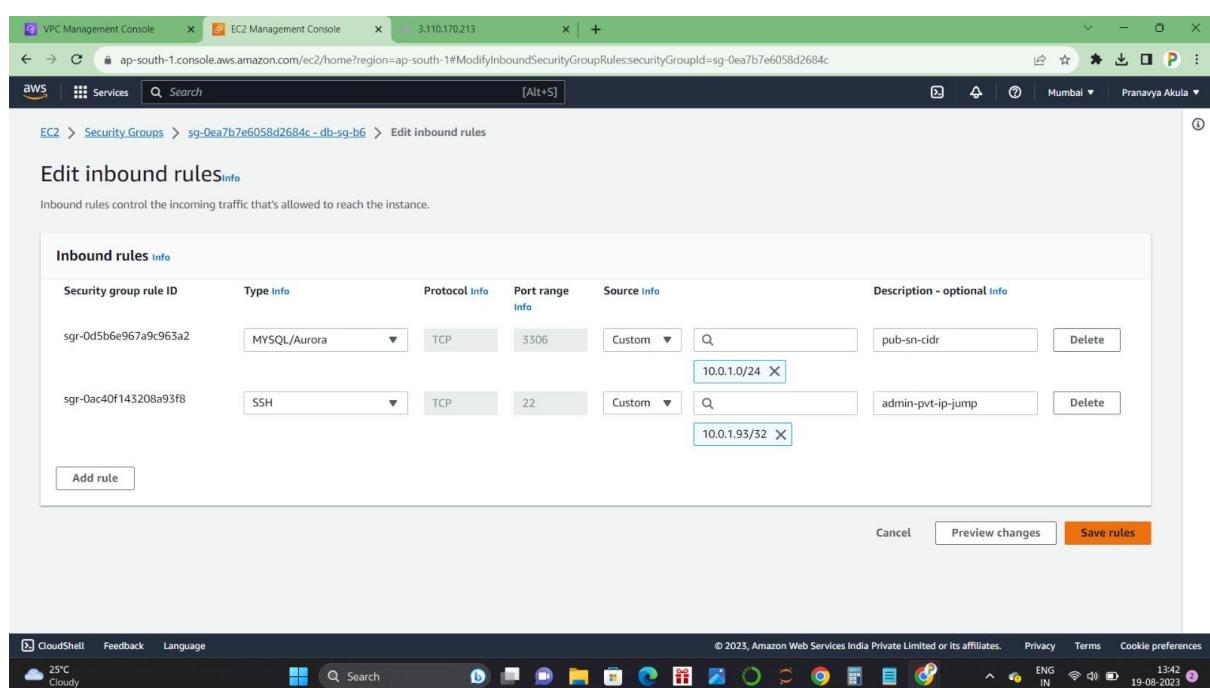
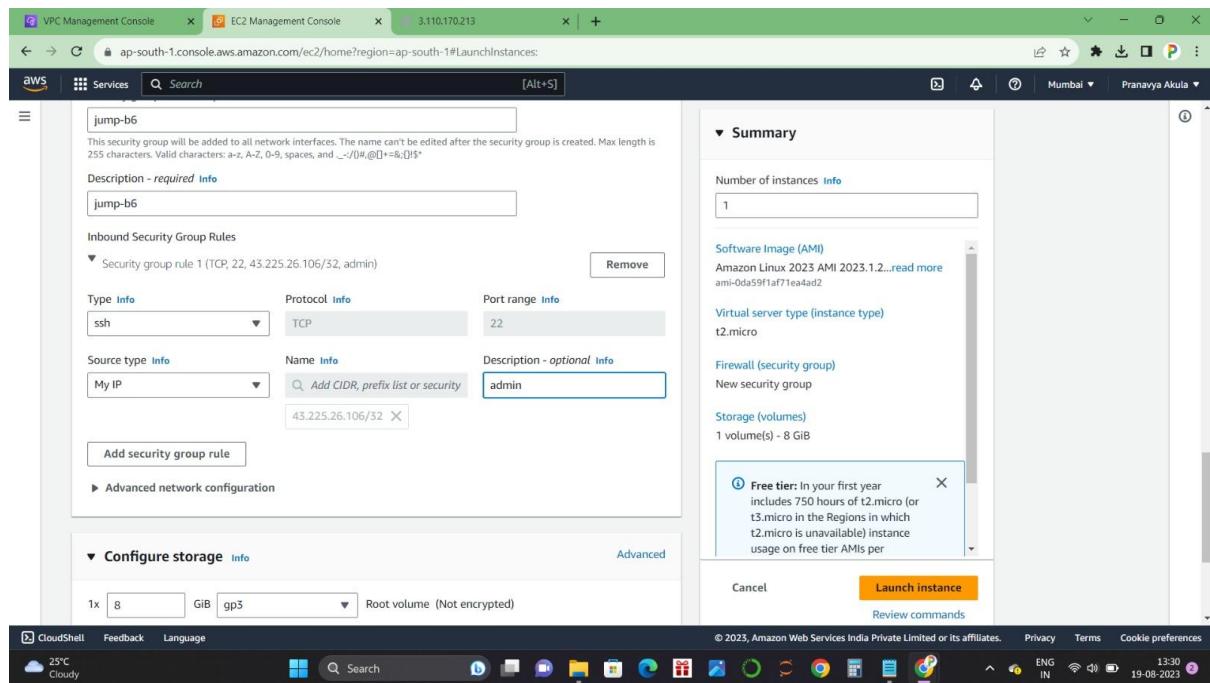
When prompted, store the private key in a secure and accessible location on your computer. You will need it later to connect to your instance. Learn more

Cancel Create key pair

© 2023, Amazon Web Services India Private Limited or its affiliates. Privacy Terms Cookie preferences ENG IN 13:29 19-08-2023

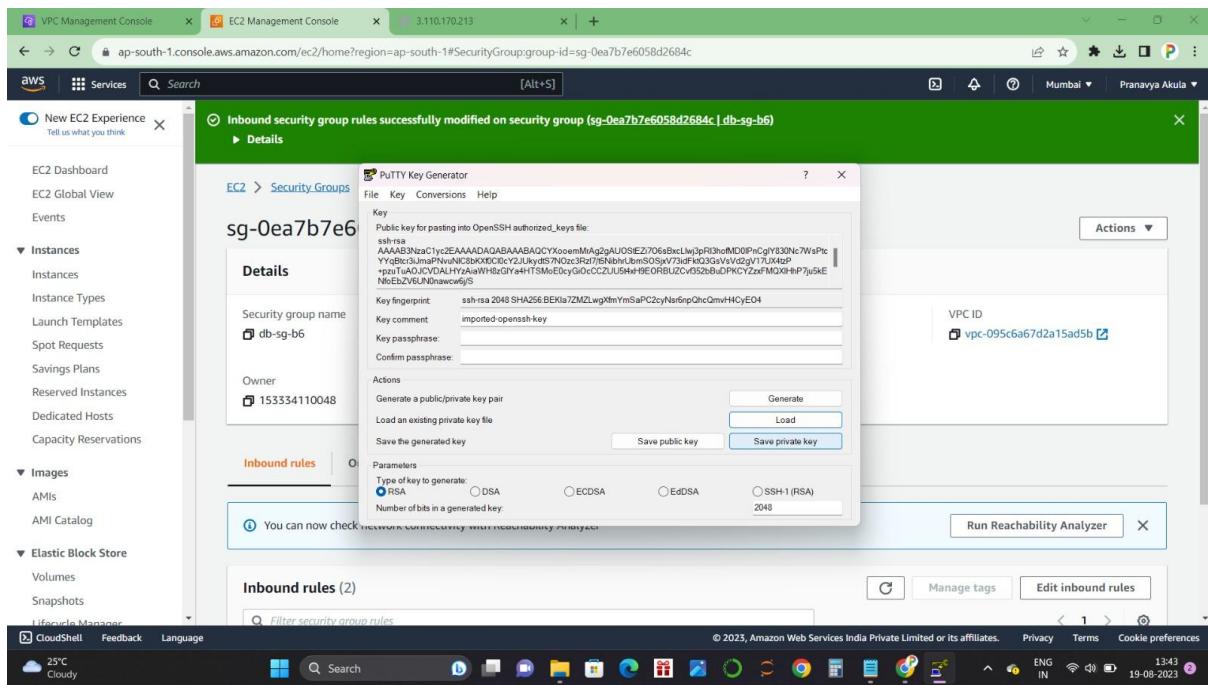
CloudShell Feedback Language 25°C Cloudy





To convert a PEM file to a PPK file using PuttyGen, follow these steps:

1. Open PuttyGen and load the "jump.pem" file.
2. Once loaded, click on the "Save private key" button to save the private key in PPK format.



To connect to the jumpbox, follow these steps:

1. Copy the DNS address of the jumpbox.
2. Open Putty and paste the DNS address in the "Hostname" field.
3. In the "Authentication" section under SSH, browse and select the saved "jump.ppk" file.
4. Accept any prompts, if necessary, to establish the connection, and you will access the command-line interface.

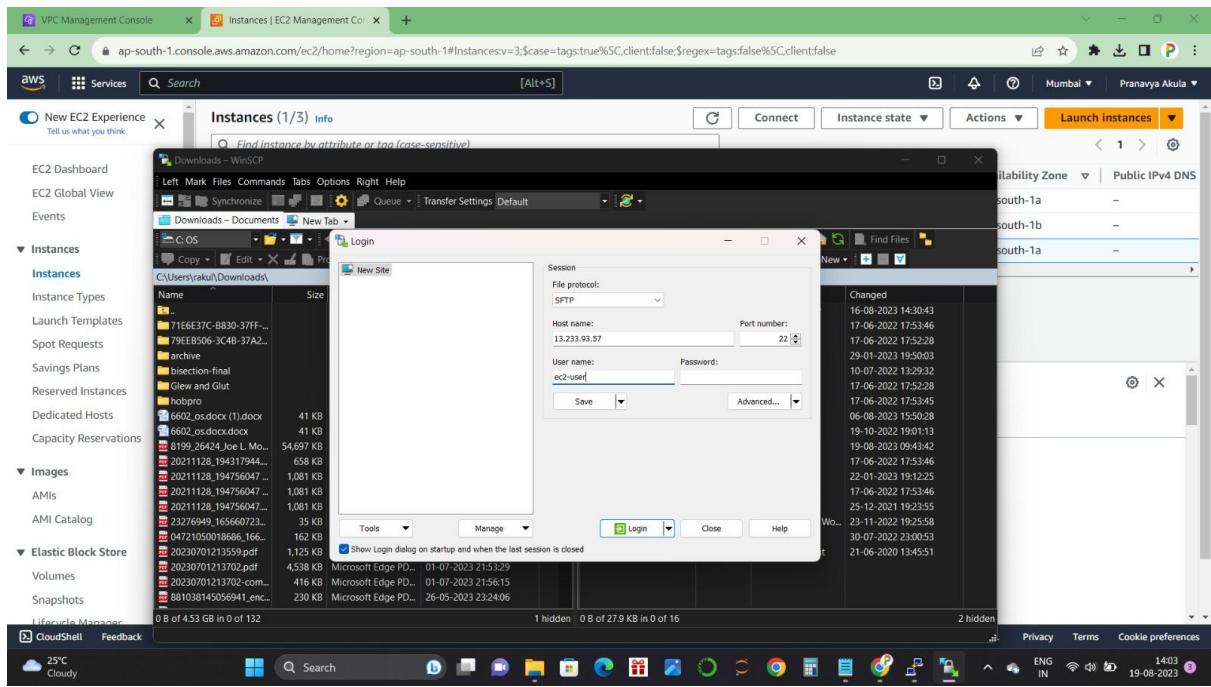
```

root@ip-10-0-1-93:~# Using username "ec2-user".
root@ip-10-0-1-93:~# Authenticating with public key "imported-openssh-key"
root@ip-10-0-1-93:~# Amazon Linux 2023
root@ip-10-0-1-93:~# https://aws.amazon.com/linux/amazon-linux-2023
root@ip-10-0-1-93:~# 
root@ip-10-0-1-93:~# Last login: Sat Aug 19 08:19:47 2023 from 43.225.26.106
[root@ip-10-0-1-93 ~]# sudo su
[root@ip-10-0-1-93 ~]# yum update -y
Last metadata expiration check: 0:22:56 ago on Sat Aug 19 08:01:38 2023.
Dependencies resolved.
Nothing to do.
Complete!
[root@ip-10-0-1-93 ~]# ssh -i "db-sg-b6.pem" ec2-user@10.0.2.176
Warning: Identity file db-sg-b6.pem not accessible: No such file or directory.
ec2-user@10.0.2.176: Permission denied (publickey,gssapi-keyex,gssapi-with-mic).
[root@ip-10-0-1-93 ~]#

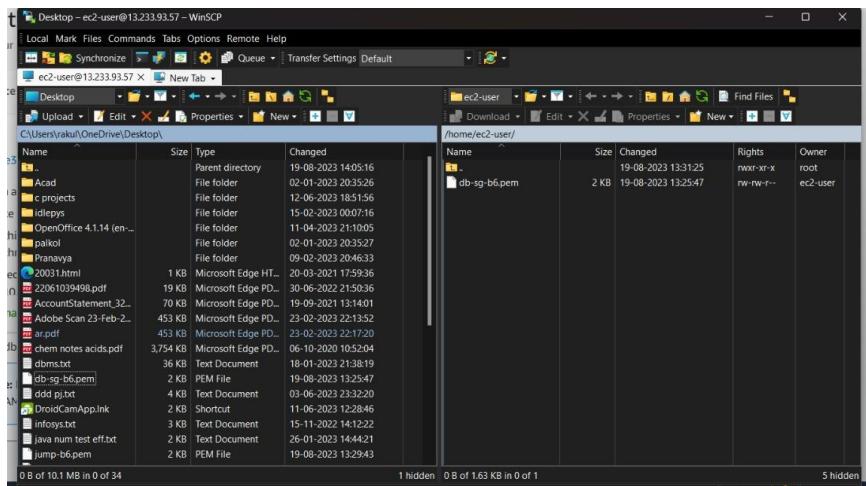
```

To connect to the database, follow these steps:

1. Move the ".pem" file of the database to the virtual environment.
2. Utilize the "WinSCP" application for this purpose. Download and open WinSCP, then input the public DNS of the jumpbox in the hostname field.
3. In the advanced settings of "authentication," upload the "jump.ppk" file and click "login."
4. Upon successful login, you will have access to a dedicated environment for storing files related to the virtual environment.



Drag and drop the database ".pem" file into the virtual environment. Next, open the Command Line Interface and paste the address of the Database instance into the interface.



To establish internet connectivity for the private subnet, we followed these :

1. Create a NAT Gateway:

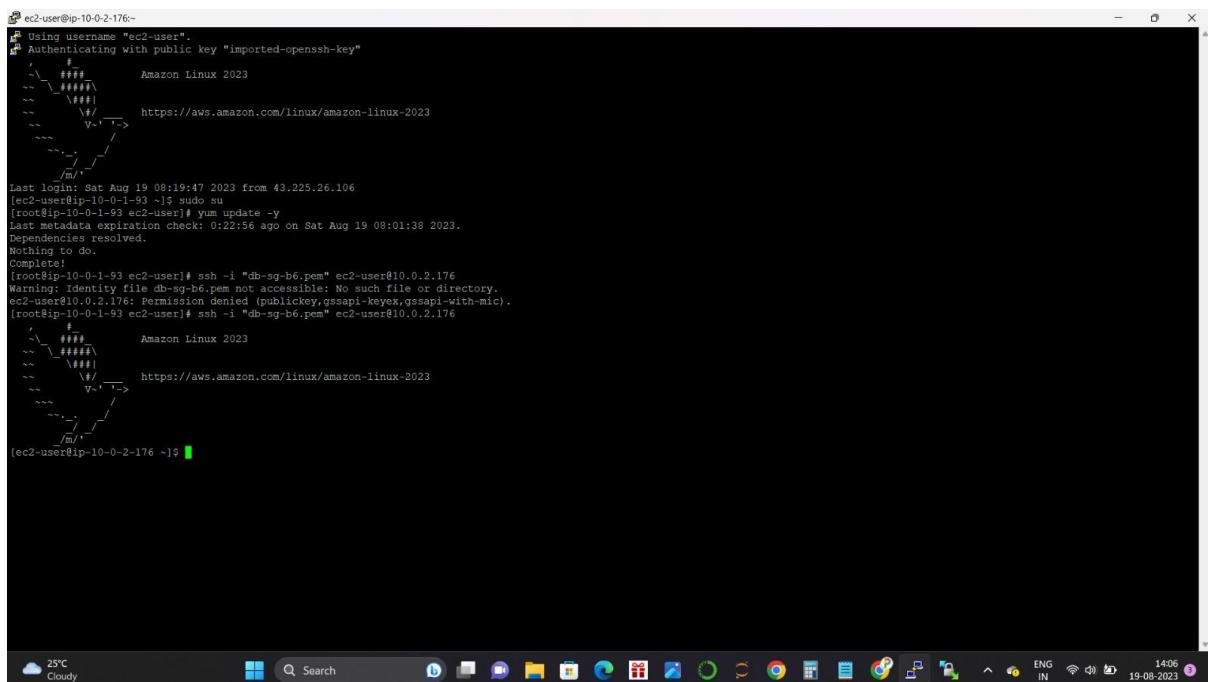
- Navigate to the NAT Gateways section in VPC .
- Click on "Create NAT Gateway."
- Enter a suitable name for the NAT Gateway.
- Choose the public subnet where we want to deploy the NAT Gateway.
- Allocate an Elastic IP address to ensure a static public IP for the NAT Gateway.
- Click "Create" to initiate the NAT Gateway creation process.

2. Connect NAT Gateway to Private Subnet:

- Access VPC's route tables section.
- Identify the route table associated with private subnet.
- Edit the route table and add a new route to direct internet-bound traffic (0.0.0.0/0) to the newly created NAT Gateway.
- Save the changes to the route table.

3. Verification:

- Confirmed that the private subnet is now using the NAT Gateway as its route to the internet.
- Instances in the private subnet can now access the internet indirectly through the NAT Gateway.



The screenshot shows a terminal window on an Amazon Linux 2023 system. The user is connected via SSH from their local machine (IP 43.225.26.106) to the EC2 instance (IP 10.0.1.93). The session starts with the user logging in as 'ec2-user'. They run 'sudo su' to become root. Then, they attempt to update the package list with 'yum update -y', but receive a warning about a missing identity file 'db-sg-b6.pem' and a permission denied error for 'ssh -i "db-sg-b6.pem"'. Finally, they run another 'ssh' command with the same issue. The terminal window has a dark background with light-colored text. The bottom of the screen shows the Windows taskbar with various icons and system status information.

```
ec2-user@ip-10-0-2-176:~$ Using username "ec2-user".
Authenticating with public key "imported-openssh-key"
Amazon Linux 2023
https://aws.amazon.com/linux/amazon-linux-2023
Last login: Sat Aug 19 08:19:47 2023 from 43.225.26.106
[ec2-user@ip-10-0-1-93 ~]$ sudo su
[root@ip-10-0-1-93 ec2-user]# yum update -y
Last metadata expiration check: 0:22:56 ago on Sat Aug 19 08:01:38 2023.
Dependencies resolved.
Nothing to do.
Complete!
[root@ip-10-0-1-93 ec2-user]# ssh -i "db-sg-b6.pem" ec2-user@10.0.2.176
Warning: Identity file db-sg-b6.pem not accessible: No such file or directory.
ec2-user@10.0.2.176: Permission denied (publickey,gssapi-keyex,gssapi-with-mic).
[root@ip-10-0-1-93 ec2-user]# ssh -i "db-sg-b6.pem" ec2-user@10.0.2.176
Amazon Linux 2023
https://aws.amazon.com/linux/amazon-linux-2023
[ec2-user@ip-10-0-2-176 ~]$
```

Network Access Control List (NACL) is the additional layer of security for controlling the in and out traffic at the subnet levels in AWS .

Connecting NACL to both the subnets :

- Open Amazon VPC service in the AWS console and click on ‘Network ACLs’ on the navigator pane.
- Create a new Network ACL and specify a name and VPC association.
- Define inbound and outbound variables by specifying the protocols, ports and IP ranges.
- Name the default NACL associated with the public subnet as main.

Name	Subnet ID	Associated with	Availability Zone	IPv4 CIDR	IPv6 CIDR
<input checked="" type="checkbox"/> 10.0.1.0/24-1a-Pub	subnet-0fbaf21bc9196a958	acl-0a7f5b3c335f50f82 / NACL-Custom	ap-south-1a	10.0.1.0/24	-
<input type="checkbox"/> 10.0.2.0/24-1b-Pvt	subnet-0c60e3e81a325c108	acl-0a7f5b3c335f50f82 / NACL-Custom	ap-south-1b	10.0.2.0/24	-

Name	Net...	Associate...	Default	VPC ID	Inbound rules count	Outbound r...
<input type="checkbox"/> NACL-Default	acl-0a7f...	subnet-0c60...	Yes	vpc-095c6a67d2a15ad5b / Batch-6-VPC	2 Inbound rules	2 Outbound r...
<input type="checkbox"/> -	acl-0e4...	-	Yes	vpc-058197d168ae19ee	2 Inbound rules	2 Outbound r...
<input checked="" type="checkbox"/> NACL-Custom	acl-038...	subnet-0fbaf...	No	vpc-095c6a67d2a15ad5b / Batch-6-VPC	1 Inbound rule	1 Outbound r...

VPC Management Console EC2 Management Console

ap-south-1.console.aws.amazon.com/vpc/home?region=ap-south-1>EditInboundRules:networkAclId=acl-03824d81215ec1a4b

Services ec2 Mumbai Pranavya Akula

VPC > Network ACLs > acl-03824d81215ec1a4b / NAACL-Custom > Edit inbound rules

Edit inbound rules Info

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

Rule number	Type	Protocol	Port range	Source	Allow/Deny	Action
100	SSH (22)	TCP (6)	22	43.225.26.106/32	Allow	<button>Remove</button>
200	HTTP (80)	TCP (6)	80	0.0.0.0/0	Allow	<button>Remove</button>
300	Custom TCP	TCP (6)	1024 - 65535	0.0.0.0/0	Allow	<button>Remove</button>
*	All traffic	All	All	0.0.0.0/0	Deny	

Add new rule Sort by rule number

Cancel Preview changes Save changes

CloudShell Feedback Language © 2023, Amazon Web Services India Private Limited or its affiliates. Privacy Terms Cookie preferences 14:31 19-08-2023 ENG IN

Network ACLs | VPC Management Console EC2 Management Console

ap-south-1.console.aws.amazon.com/vpc/home?region=ap-south-1#acls:

VPC dashboard Services ec2

EC2 Global View New

Filter by VPC: Select a VPC

Virtual private cloud Your VPCs New 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 groups Network ACLs

You have successfully updated inbound rules for acl-03824d81215ec1a4b / NACL-Custom

Network ACLs (1/3) Info

Name Net... Associate... Default VPC ID Inbound rules count Outbound rules count

Name	Net...	Associate...	Default	VPC ID	Inbound rules count	Outbound rules count
NACL-Default	acl-0a7f...	subnet-0c60...	Yes	vpc-095c6a67d2a15ad5b / Batch-6-VPC	2 Inbound rules	2 Outbound rules
-	acl-0e4...	-	Yes	vpc-058197d1686ae19ee	2 Inbound rules	2 Outbound rules
<input checked="" type="checkbox"/> NACL-Custom	acl-038...	subnet-0fbaf...	No	vpc-095c6a67d2a15ad5b / Batch-6-VPC	4 Inbound rules	1 Outbound rule

Find resources by attribute or tag

Filter inbound rules

Rule number	Type	Protocol	Port range	Source	Allow/Deny
100	SSH (22)	TCP (6)	22	43.225.26.106/32	Allow
200	HTTP (80)	TCP (6)	80	0.0.0.0/0	Allow
300	Custom TCP	TCP (6)	1024 - 65535	0.0.0.0/0	Allow
*	All traffic	All	All	0.0.0.0/0	Deny

CloudShell Feedback Language © 2023, Amazon Web Services India Private Limited or its affiliates. Privacy Terms Cookie preferences 14:30 IN 19-08-2023

Network ACLs | VPC Management Console Instance details | EC2 Management Console

3.110.170.213 Not secure | 3.110.170.213

26°C Cloudy Search ENG IN 14:42 19-08-2023

Rule number	Type	Protocol	Port range	Destination	Allow/Deny
100	SSH (22)	TCP (6)	22	43.225.26.106/32	Allow
200	HTTP (80)	TCP (6)	80	0.0.0.0/0	Allow
300	Custom TCP	TCP (6)	1024 - 65535	0.0.0.0/0	Allow
400	HTTP (80)	TCP (6)	80	43.225.26.106/32	Deny
*	All traffic	All	All	0.0.0.0/0	Deny

Edit outbound rules Info
Outbound rules control the outgoing traffic that's allowed to leave the VPC.

Add new rule Sort by rule number

Cancel Preview changes Save changes

TASK 2:

PROBLEM STATEMENT

Reducing the wastage of medicine by accurate prediction of the quantity required of medicine by developing a ML model

LEVEL 2

Feature Engineering

Feature engineering is the process of selecting, transforming, or creating relevant features from raw data to improve the performance of machine learning algorithms. Effective feature engineering can significantly impact the model's ability to learn patterns, generalise, and make accurate predictions.

There are 4 steps involved in feature engineering:

1) Feature Selection: In this step we focused on selecting a subset of most relevant features from original list to reduce dimensionality and improve model performance

2) Feature Transformation: We tried creating new features from existing features for more explanatory power.

3)Feature creation(Encoding and binning):Creating new features by transforming categorical values of necessary features to numerical ones(encoding) and converting continuous numerical variables into categorical variables by grouping their values into discrete intervals

4)Feature Extraction:We reduced the redundant data from the data set for better yields on applying ML algorithms.

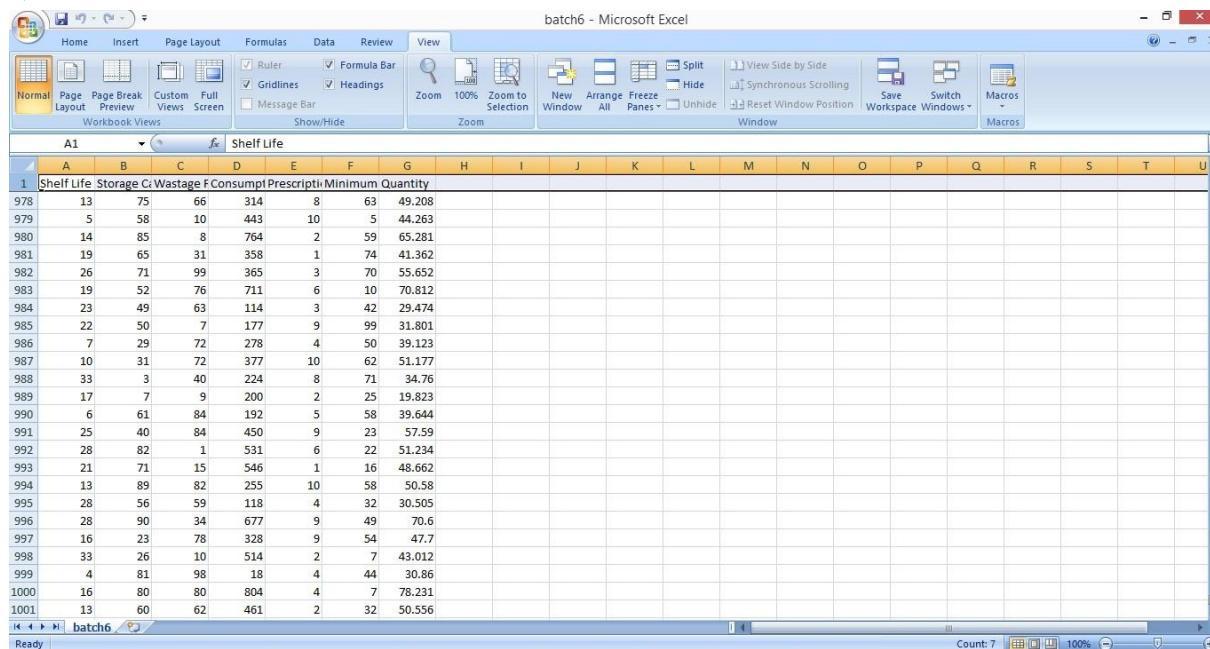
A	B	C	D	E	F	G	H	I	
PROBLEM STATEMENT: Reducing the wastage of medicine by accurate prediction of the quantity required of medicine by developing an ML model									
	FEATURES	FEATURE SELECTION	FEATURE EXTRACTION						
(Availability)	1 Shelf life 2 Inventory Turnover Ratio 3 Price of drug 4 Cold storage 5 Transportation delay 6 Drug counterfeiting 7 Seasonal demand 8 Patient demographic 9 Demand fluctuations 10 Prescription Frequency 11 recommendations 12 Geographical Factors 13 Time taken by Supplier 14 Budget 15 Health trends 16 Rate of allergy 17 Consumption rate 18 Inflation 19 Distribution rate 20 Government policies 21 Minimum stock 22 Tax rates 23 Marketing 24 Favoured Prescription 25 Feedback of users 26 Environment of storage 27 Climate 28 Competition 29 Dosage of Medicine recommended 30 Patient Inflow and Outflow 31 Return of medicine 32 Size and Type of Hospital or pharma store 33 Cost of Maintenance 34 Relations with Suppliers 35 Availability of Medicine 36 Handling drug waste 37 Composition of medicine 38 Emergence of New Diseases 39 Temperature 40 Availability of Experts 41 Historical Data of Sales 42 Damage Rate	Shelf Life Inventory Turnover Ratio Budget Consumption Rate=(Amount of medicines consumed)/time Distribution Rate Demand Fluctuations Minimum Stock Patient Inflow and Outflow Historical Data of Sales Feedback of users Prescription Frequency = how many times it was prescribed by doctors Wastage Rate= the medicines which couldn't be sold due to expiry Damaged good= the good received that were already damaged	Shelf life Consumption Rate Wastage Rate Storage Capacity Prescription Frequency Minimum Stock						

LEVEL 3

Data frame creation and Github

Data frame creation:

- 1)We created a sample dataset and performed necessary calculations.
- 2)Used cramer's rule to solve a system of linear equations to obtain weights that can be applied further.
- 3)We generated data set by using pandas library in Jupyter notebook.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Shelf Life	Storage	Wastage	Consumpti	Prescripti	Minimum	Quantity														
978	13	75	66	314	8	63	49.208														
979	5	58	10	443	10	5	44.263														
980	14	85	8	764	2	59	65.281														
981	19	65	31	358	1	74	41.362														
982	26	71	99	365	3	70	55.652														
983	19	52	76	711	6	10	70.812														
984	23	49	63	114	3	42	29.474														
985	22	50	7	177	9	99	31.801														
986	7	29	72	278	4	50	39.123														
987	10	31	72	377	10	62	51.177														
988	33	3	40	224	8	71	34.76														
989	17	7	9	200	2	25	19.823														
990	6	61	84	192	5	58	39.644														
991	25	40	84	450	9	23	57.59														
992	28	82	1	531	6	22	51.234														
993	21	71	15	546	1	16	48.662														
994	13	89	82	255	10	58	50.58														
995	28	56	59	118	4	32	30.505														
996	28	90	34	677	9	49	70.6														
997	16	23	78	328	9	54	47.7														
998	33	26	10	514	2	7	43.012														
999	4	81	98	18	4	44	30.86														
1000	16	80	80	804	4	7	78.231														
1001	13	60	62	461	2	32	50.556														

GitHub

We created a github repository ,uploaded our dataset file and copied the link for further purposes

LEVEL 4

Notebook instance and ML Model

AWS Sagemaker : Amazon SageMaker is a fully managed machine learning service where we can quickly and easily build and train machine learning models, and then directly deploy them into a production-ready hosted environment.

CREATING A NOTEBOOK INSTANCE:

Go to Amazon SageMaker service from the AWS Management Console.Create a Notebook instance by clicking on “Notebook instances” in the left menu and then click “Create notebook instance”.

The screenshot shows the 'Create notebook instance' page in the AWS SageMaker console. The top navigation bar includes the AWS logo, 'Services' button, a search bar, and user information for 'Pranavya Akula'. The main content area has a title 'Create notebook instance' and a sub-instruction: 'Amazon SageMaker provides pre-built fully managed notebook instances that run Jupyter notebooks. The notebook instances include example code for common model training and hosting exercises. [Learn more](#)'.

Notebook instance settings

- Notebook instance name:** Batch6
- Notebook instance type:** ml.t3.medium
- Platform identifier:** Amazon Linux 2, Jupyter Lab 3
- Additional configuration:** A link to expand.

Permissions and encryption

- IAM role:** AmazonSageMaker-FullAccess (selected)
- Platform identifier:** Amazon Linux 2, Jupyter Lab 3
- Additional configuration:** A link to expand.

At the bottom, there are links for CloudShell, Feedback, Language, and a footer with copyright information: '© 2023, Amazon Web Services India Private Limited or its affiliates.' and links for Privacy, Terms, and Cookie preferences.

AWS Identity and Access Management (IAM) is a web service that helps you securely control access to AWS resources. With IAM, you can centrally manage permissions that control which AWS resources users can access. You use IAM to control who is authenticated (signed in) and authorized (has permissions) to use resources.

This screenshot shows the same 'Create notebook instance' page but with expanded sections for 'Permissions and encryption' and 'Network'.

Permissions and encryption

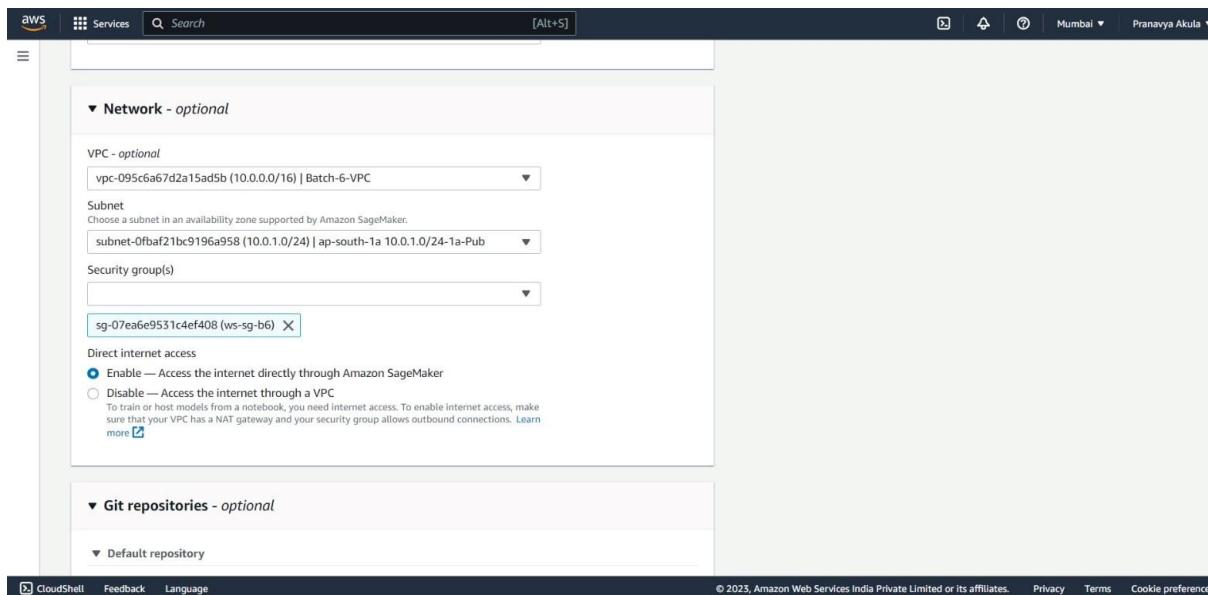
- IAM role:** AmazonSageMaker-ExecutionRole-20230819T205091
- Create role using the role creation wizard:** A link to expand.
- Root access - optional:**
 - Enable - Give users root access to the notebook
 - Disable - Don't give users root access to the notebook
Lifecycle configurations always have root access
- Encryption key - optional:** No Custom Encryption

Network - optional

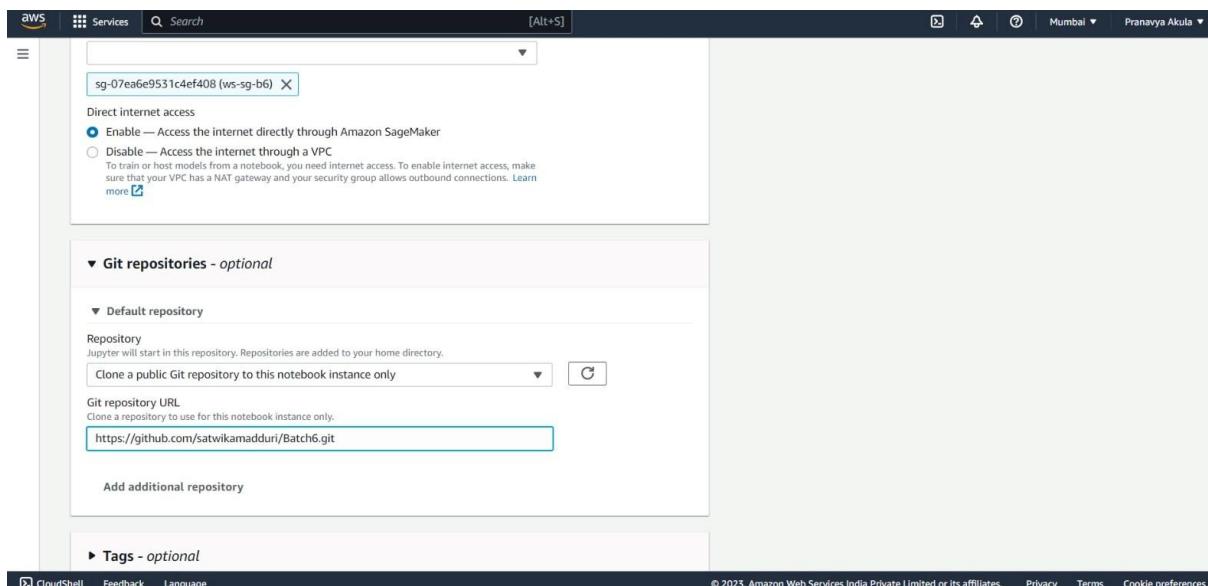
VPC - optional

At the bottom, there are links for CloudShell, Feedback, Language, and a footer with copyright information: '© 2023, Amazon Web Services India Private Limited or its affiliates.' and links for Privacy, Terms, and Cookie preference.

Select the VPC and public subnet created initially



Clone the github repository and paste link the address



Data Models we have obtained from the generated data

1. Linear Regression

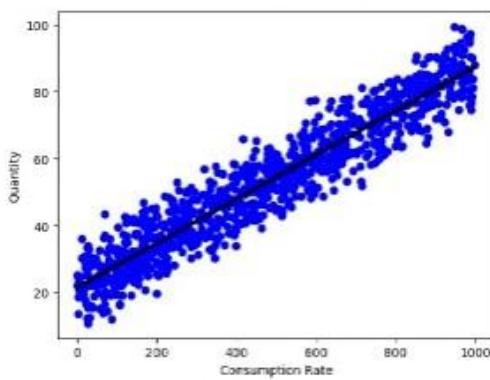
Linear regression is a simple regression algorithm that models the relationship between the dependent variable and one or more independent variables. It assumes a linear relationship and tries to find the best-fit straight line that minimizes the sum of squared differences between the observed and predicted values.

```
In [33]: #LINEAR REGRESSION
X = data[["Consumption Rate"]]
Y = data["Quantity"]
x_train,x_test,y_train,y_test=train_test_split(X,Y,test_size=0.2,random_state=1)
mdl = LinearRegression()
mdl.fit(x_train, y_train)
pred = mdl.predict(x_test)
print("Predicted value (LR): ",pred[0])
print("Accuracy (LR): ",mdl.score(X[:100], Y[:100])*100)

mse=mean_squared_error(y_test,pred)
r2=r2_score(y_test,pred)
mae=mean_absolute_error(y_test,pred)
print("Mean Square Error of Linear Regression",mse)
print("R2 score of Linear Regression",r2)
print("Mean absolute Error of Linear Regression",mae)

plt.scatter(X['Consumption Rate'], Y, color='b')
plt.plot(X['Consumption Rate'], mdl.predict(X),color='black', linewidth=3)
plt.xlabel('Consumption Rate')
plt.ylabel('Quantity')
plt.show()

Predicted value (LR): 23.74385734968174
Accuracy (LR): 91.4839766775333
Mean Square Error of Linear Regression 42.62547487485745
R2 score of Linear Regression 0.8851188385098868
Mean absolute Error of Linear Regression 5.337882812520429
```

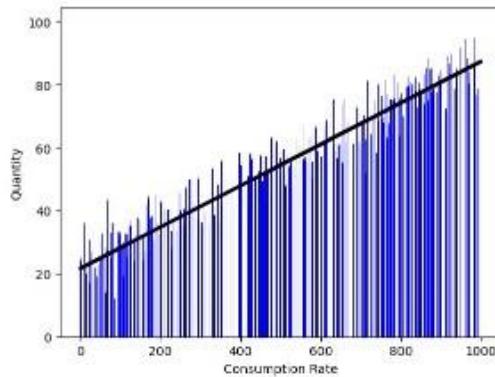


```
In [41]: #LINEAR REGRESSION
X = data[["Consumption Rate"]]
Y = data["Quantity"]
x_train,x_test,y_train,y_test=train_test_split(X,Y,test_size=0.2,random_state=1)
mdl = LinearRegression()
mdl.fit(x_train, y_train)
pred = mdl.predict(x_test)
print("Predicted value (LR): ",pred[0])
print("Accuracy (LR): ",mdl.score(X[:100], Y[:100])*100)

mse=mean_squared_error(y_test,pred)
r2=r2_score(y_test,pred)
mae=mean_absolute_error(y_test,pred)
print("Mean Square Error of Linear Regression",mse)
print("R2 score of Linear Regression",r2)
print("Mean absolute Error of Linear Regression",mae)

plt.bar(X['Consumption Rate'], Y, color='b')
plt.plot(X['Consumption Rate'], mdl.predict(X),color='black', linewidth=3)
plt.xlabel('Consumption Rate')
plt.ylabel('Quantity')
plt.show()

Predicted value (LR): 23.74385734968174
Accuracy (LR): 91.4839766775333
Mean Square Error of Linear Regression 42.62547487485745
R2 score of Linear Regression 0.8851188385098868
Mean absolute Error of Linear Regression 5.337882812529429
```



2. Multiple Linear Regression

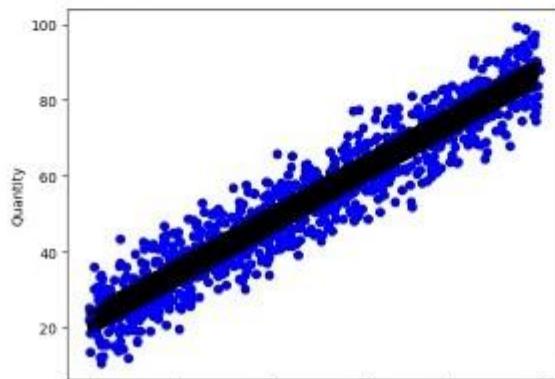
Multiple linear regression extends linear regression to multiple independent variables. It models the relationship between the dependent variable and multiple independent variables by finding the best-fit hyperplane in a higher-dimensional space.

```
In [39]: #MULTIPLE LINEAR REGRESSION
X = data[["Consumption Rate"]]
Y = data["Quantity"]
X = data[["Consumption Rate","Prescription Frequency"]]
x_train,x_test,y_train,y_test=train_test_split(X,Y,test_size=0.2,random_state=1)

mdl = LinearRegression()
mdl.fit(x_train,y_train)
pred = mdl.predict(x_train)
mse=mean_squared_error(y_test,pred[:200])
r2=r2_score(y_test,pred[:200])
mae=mean_absolute_error(y_test,pred[:200])
print("Predicted value (MLR): ",pred[0])
print("Accuracy (MLR): ",mdl.score(X[:180], Y[:180])*100)
print("Mean Square Error of Multiple Linear Regression",mse)
print("R2 score of Multiple linear Regression",r2)
print("Mean absolute Error of Multiple Regression",mae)

print("Predicted value (MLR): ",pred[8])
print("Accuracy (MLR): ",mdl.score(X[:180], Y[:180])*100)
plt.scatter(X['Consumption Rate'], Y, color='blue')
plt.plot(X['Consumption Rate'], mdl.predict(X),color='black', linewidth=3)
plt.xlabel('Consumption Rate')
plt.ylabel('Quantity')
plt.show()

Predicted value (MLR): 53.96436563471623
Accuracy (MLR): 91.3442737972953
Mean Square Error of Multiple Linear Regression 724.5598722800623
R2 score of Multiple Linear Regression -0.9529163269395946
Mean absolute Error of Multiple Regression 22.12496622783901
Predicted value (MLR): 53.96436563471623
Accuracy (MLR): 91.3442737972953
```



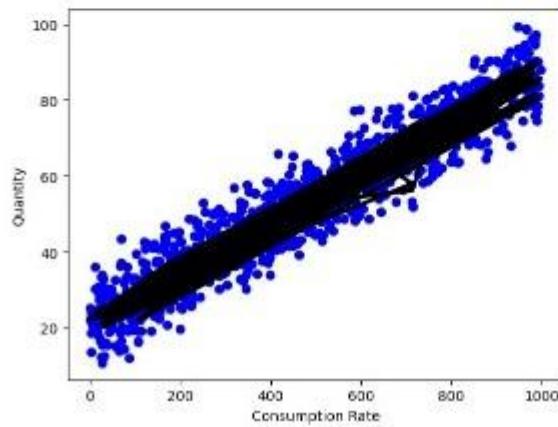
3. Random forest regression

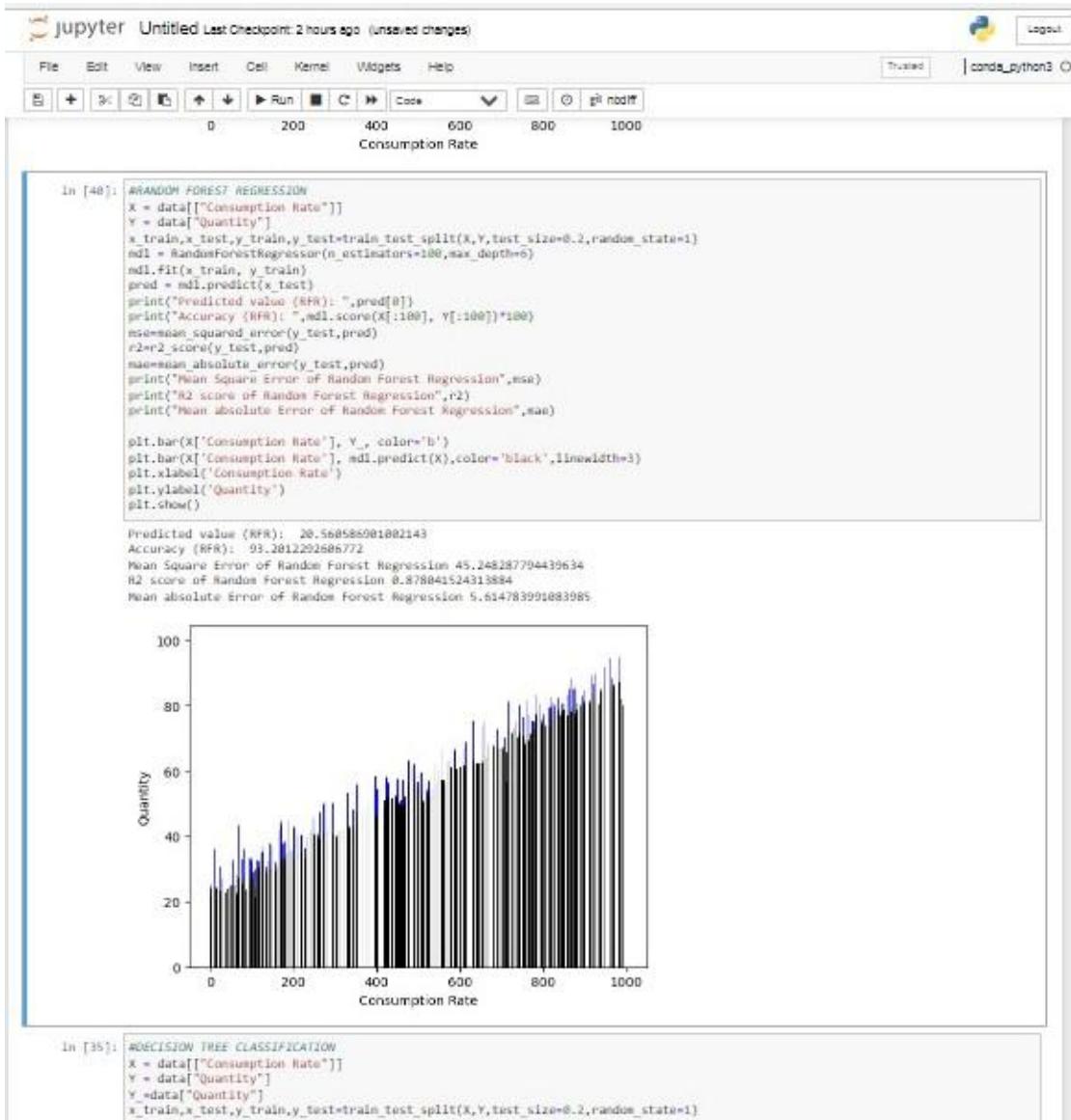
Random Forest Regression is an ensemble learning technique based on decision trees. It builds multiple decision trees and averages their predictions to improve prediction accuracy. It's particularly useful for handling complex relationships in data and mitigating overfitting.

```
In [34]: #RANDOM FOREST REGRESSION
X = data[["Consumption Rate"]]
Y = data["Quantity"]
X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.2,random_state=1)
mdl = RandomForestRegressor(n_estimators=100,max_depth=5)
mdl.fit(X_train, Y_train)
pred = mdl.predict(X_test)
print("Predicted value (RFR): ",pred[0])
print("Accuracy (RFR): ",mdl.score(X[:180], Y[:180])*100)
mse=mean_squared_error(Y_test,pred)
r2=r2_score(Y_test,pred)
mae=mean_absolute_error(Y_test,pred)
print("Mean Square Error of Random Forest Regression",mse)
print("R2 score of Random Forest Regression",r2)
print("Mean absolute Error of Random Forest Regression",mae)

plt.scatter(X['Consumption Rate'], Y_, color='b')
plt.plot(X['Consumption Rate'], pred,color="black", linewidth=3)
plt.xlabel('Consumption Rate')
plt.ylabel('Quantity')
plt.show()

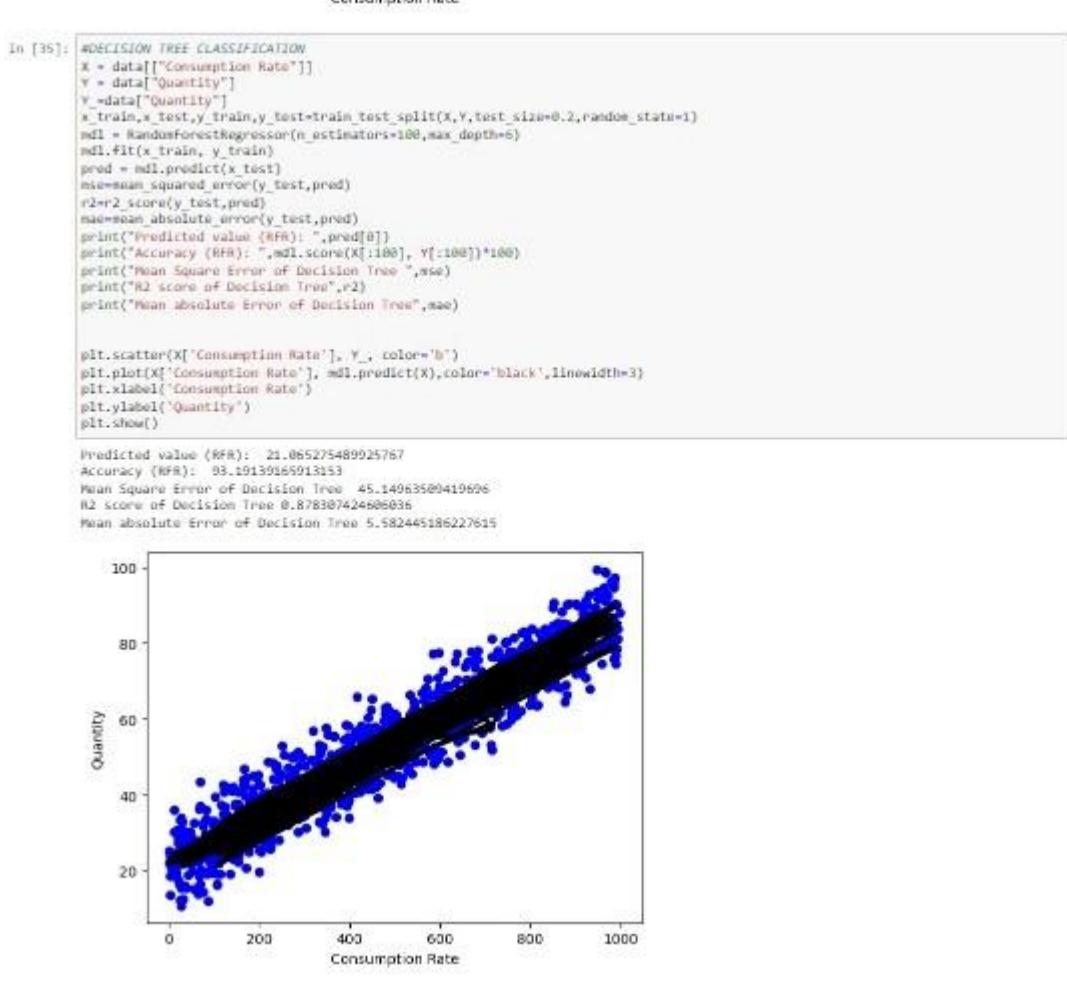
#Predicted value (RFR):  28.288661191839132
Accuracy (RFR):  93.05613409011441
Mean Square Error of Random Forest Regression 45.6557977658417
R2 score of Random Forest Regression  0.8759431557920988
Mean absolute Error of Random Forest Regression 5.628812520458646
```





4. Decision Tree regression and classification

Decision tree regression predicts continuous numerical values by recursively splitting the data into subsets based on features. It forms a tree structure where each internal node represents a decision based on a feature, and each leaf node provides a predicted value based on the training examples in that leaf's subset.



ANALYSIS: We deployed 4 machine learning models Linear Regression, Random Forest Regression, Decision tree Regression, Multiple Linear Regression and by evaluation matrix We obtained Mean Square error , R2 score, Mean Absolute Error for each model.

After obtaining the results we analysed all errors and found that, both Mean square error and Mean absolute error among 4 deployed models is less for Linear Regression and R2 score is nearly equal for all 4 models. Hence, we analysed that Linear Regression is best fit for our problem statement with minimum errors.

CONCLUSION: In conclusion, our project aimed to address the challenge of reducing medicine wastage by accurately predicting the quantity of medicine required using a machine learning model. Through a comprehensive analysis of various regression techniques, including linear regression, multiple regression, random forest, and decision tree models, we found that linear regression emerged as the most suitable algorithm for our specific dataset and prediction objective.

REFERENCES:

- <https://www.expresspharma.in/expiry-and-shelf-life-of-drug-products/>
- <https://timesofindia.indiatimes.com/blogs/voices/pharmaceutical-supply-chain-management-crises-and-innovations/>
- <https://viseven.com/pharmaceutical-industry-challenges/>
- <https://www.mdpi.com/2226-4787/8/4/221#:~:text=The%20aim%20of%20this%20study%20was%20to%20determine,healthcare%20bodies%20through%20the%20Circular%20Economy%20R%20principles>