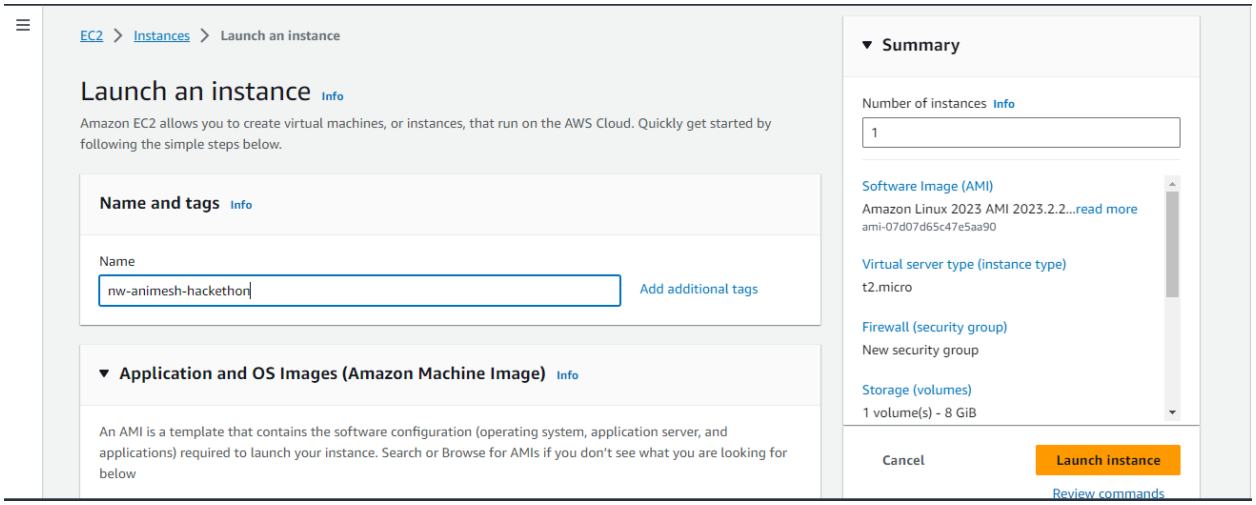
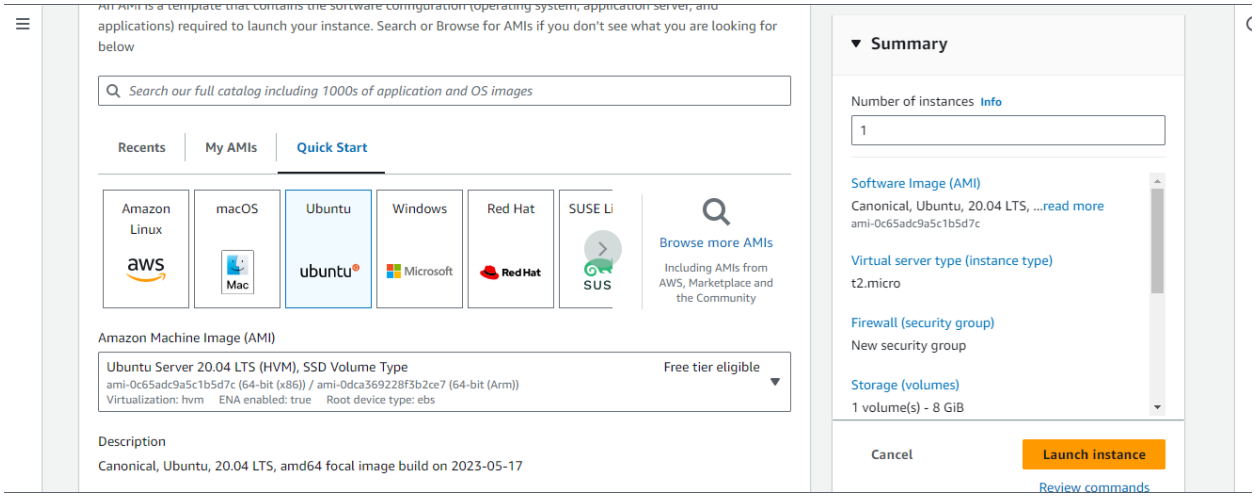
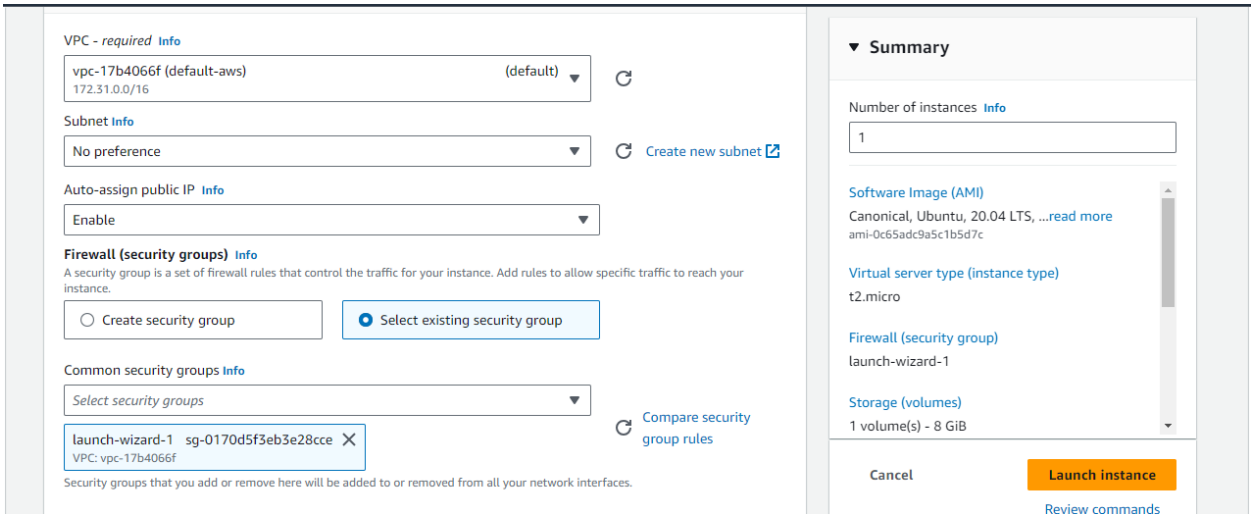
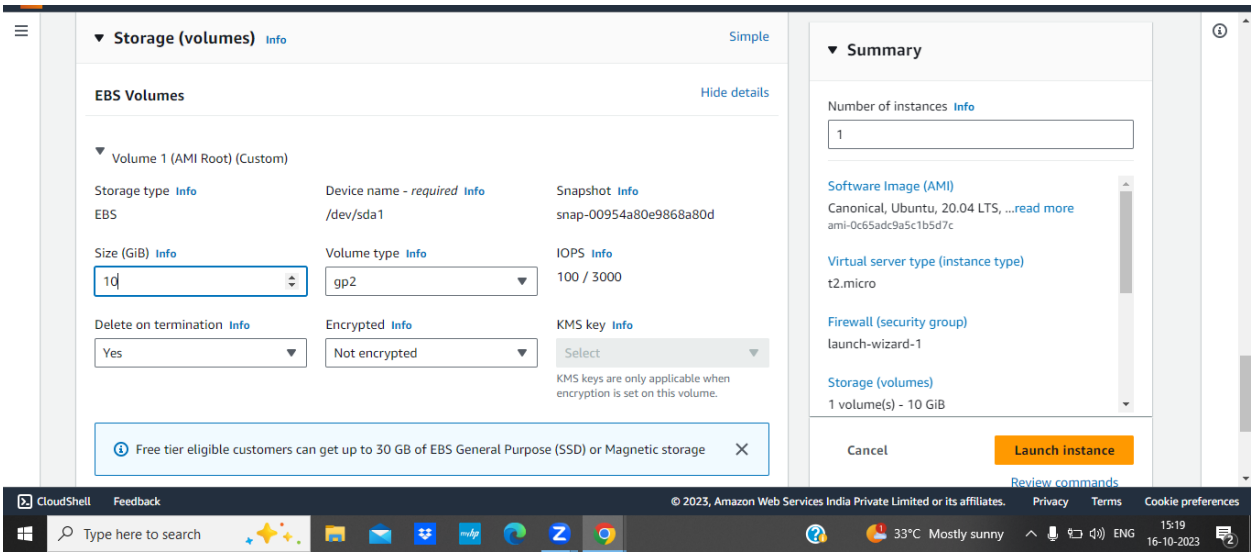
Hackathon

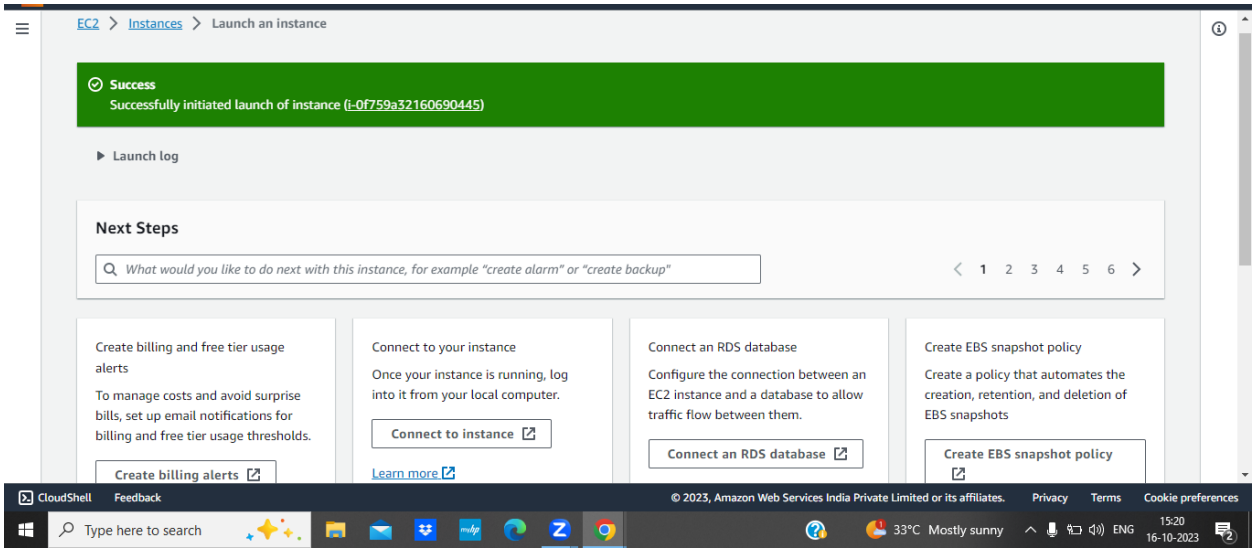
Launch an instance

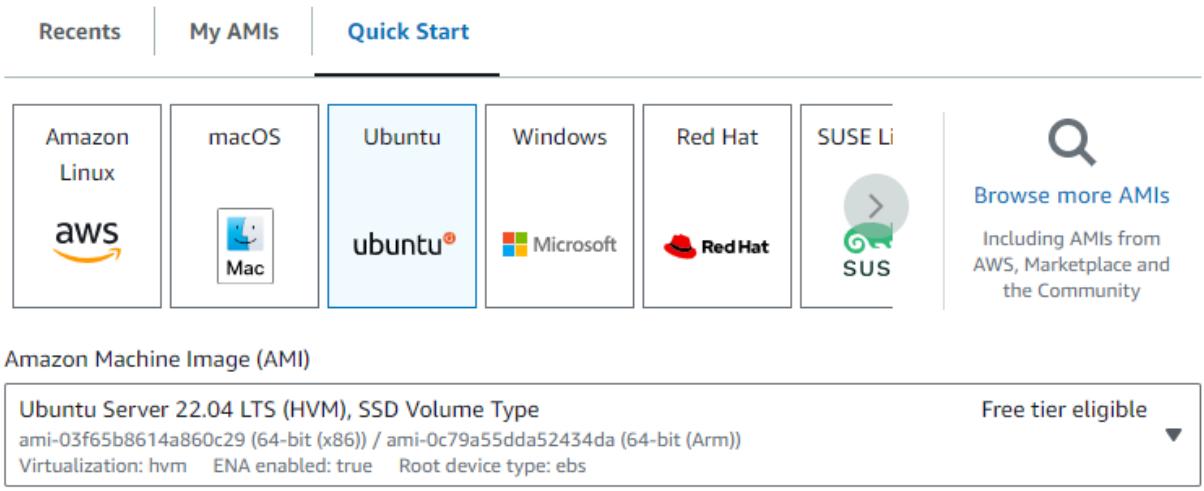


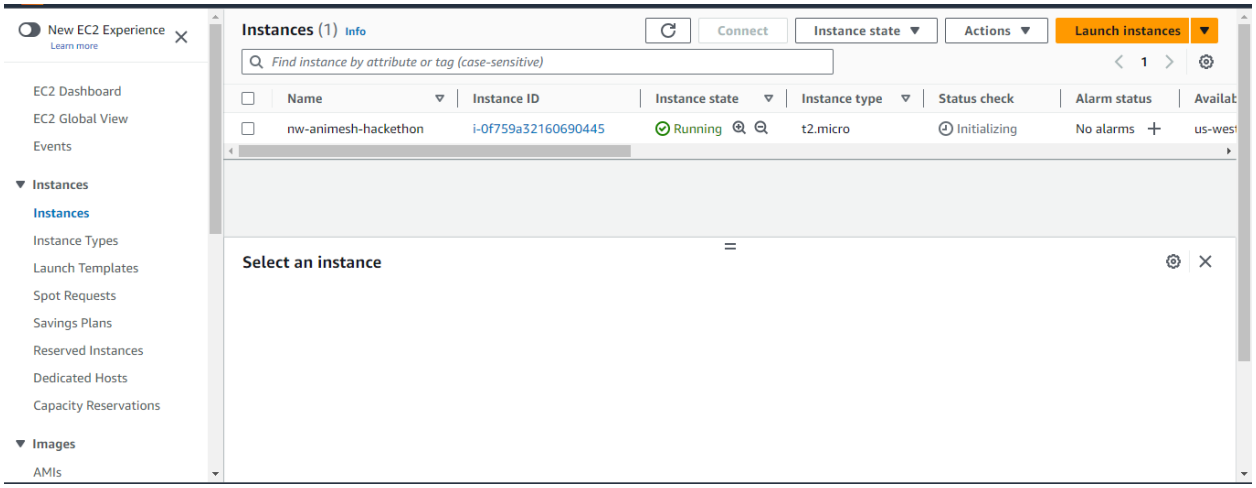


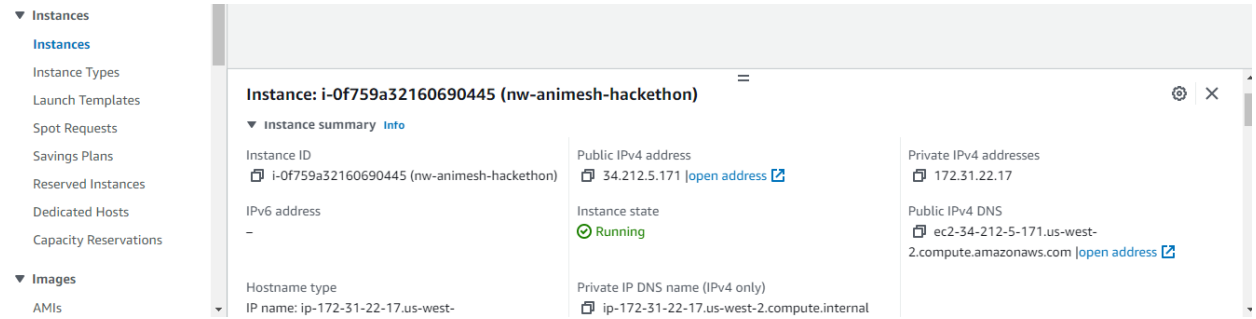

VPC - required Info
vpc-17b4066f (default-aws) (default)
172.31.0.0/16
Subnet Info
No preference
Create new subnet
Auto-assign public IP Info
Enable
Firewall (security groups) Info
A security group is a set of firewall rules that control the traffic for your instance. Add rules to allow specific traffic to reach your instance.
Create security group
Select existing security group
Common security groups Info
Select security groups
launch-wizard-1 sg-0170d5f3eb3e28cce
VPC: vpc-17b4066f
Compare security group rules
Security groups that you add or remove here will be added to or removed from all your network interfaces.
Summary
Number of instances Info
1
Software Image (AMI)
Canonical, Ubuntu, 20.04 LTS, ...read more
ami-0c65adc9a5c1b5d7c
Virtual server type (instance type)
t2.micro
Firewall (security group)
launch-wizard-1
Storage (volumes)
1 volume(s) - 8 GiB
Cancel
Launch instance
Review commands

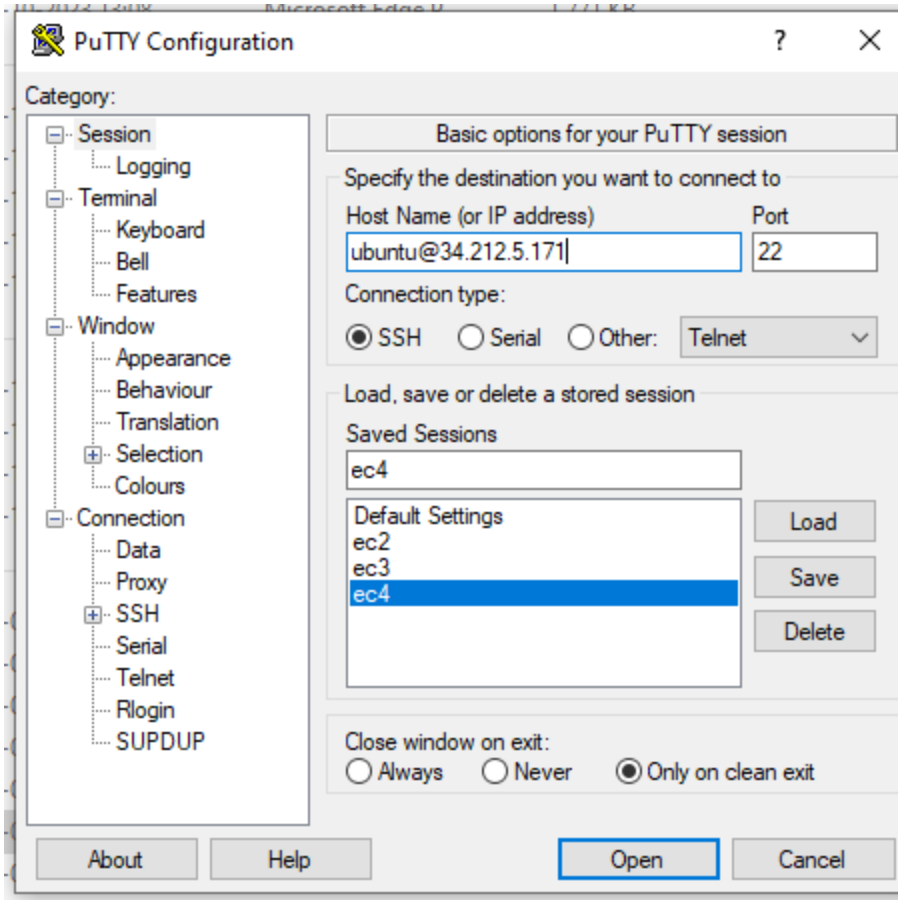

Storage (volumes) Info
Simple
EBS Volumes
Hide details
Volume 1 (AMI Root) (Custom)
Storage type Info
EBS
Device name - required Info
/dev/sda1
Snapshot Info
snap-00954a80e9868a80d
Size (GiB) Info
10
Volume type Info
gp2
IOPS Info
100 / 3000
Delete on termination Info
Yes
Encrypted Info
Not encrypted
KMS key Info
Select
KMS keys are only applicable when encryption is set on this volume.
Free tier eligible customers can get up to 30 GB of EBS General Purpose (SSD) or Magnetic storage
Summary
Number of instances Info
1
Software Image (AMI)
Canonical, Ubuntu, 20.04 LTS, ...read more
ami-0c65adc9a5c1b5d7c
Virtual server type (instance type)
t2.micro
Firewall (security group)
launch-wizard-1
Storage (volumes)
1 volume(s) - 10 GiB
Cancel
Launch instance
Review commands
CloudShell
Feedback
© 2023, Amazon Web Services India Private Limited or its affiliates.
Privacy
Terms
Cookie preferences
Type here to search
33°C Mostly sunny
ENG
15:19
16-10-2023

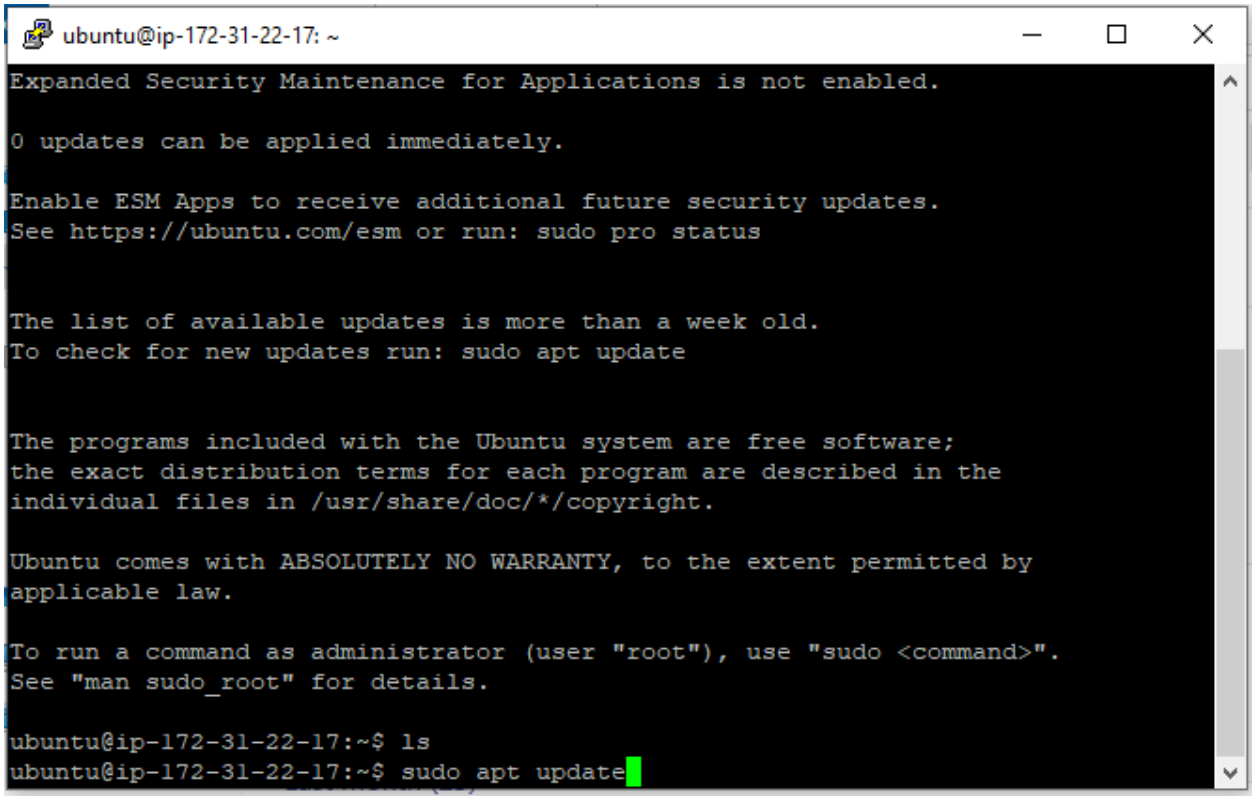

EC2 > Instances > Launch an instance
Success
Successfully initiated launch of instance (i-0f759a32160690445)
Launch log
Next Steps
What would you like to do next with this instance, for example "create alarm" or "create backup"
1 2 3 4 5 6
Create billing and free tier usage alerts
To manage costs and 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 EBS snapshot policy
Create a policy that automates the creation, retention, and deletion of EBS snapshots
Create EBS snapshot policy
CloudShell
Feedback
© 2023, Amazon Web Services India Private Limited or its affiliates.
Privacy
Terms
Cookie preferences
Type here to search
33°C Mostly sunny
ENG
15:20
16-10-2023

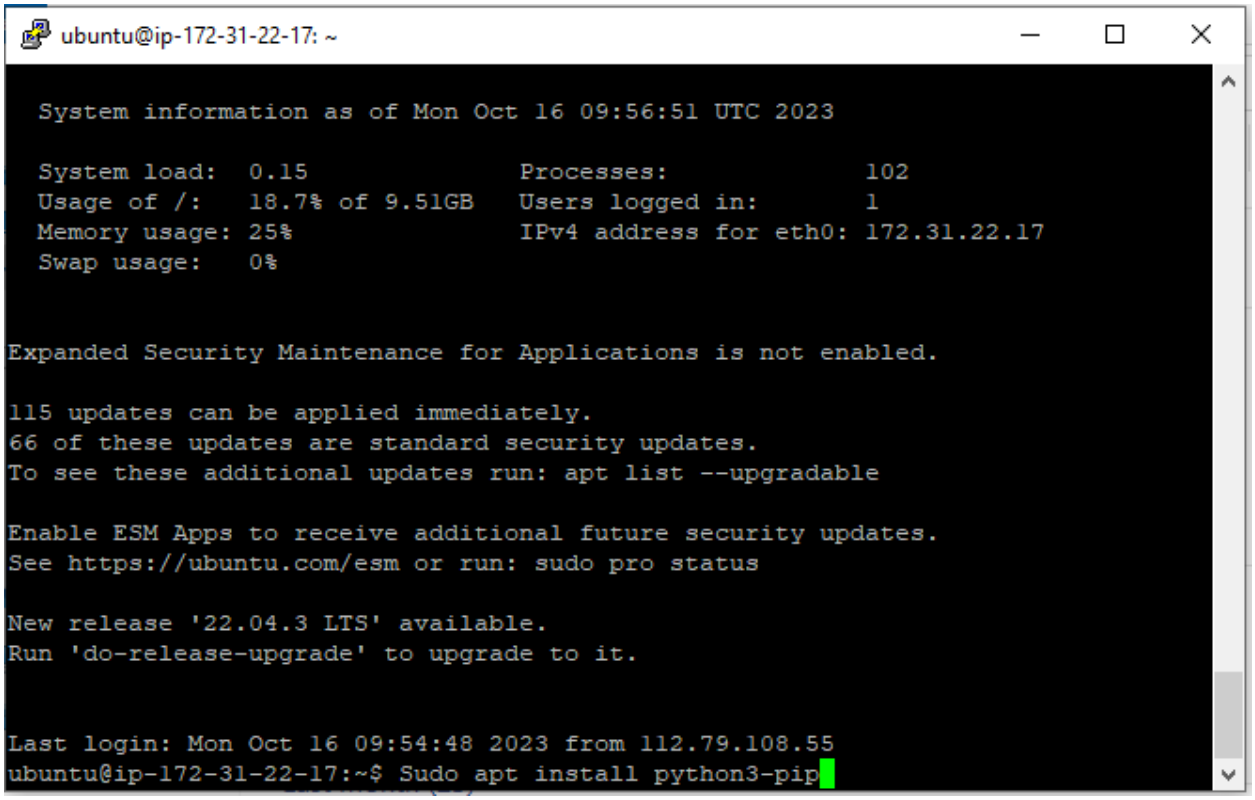

Recents
My AMIs
Quick Start
Amazon Linux
aws
macOS
Mac
Ubuntu
ubuntu
Windows
Microsoft
Red Hat
Red Hat
SUSE Li
SUS
Browse more AMIs
Including AMIs from AWS, Marketplace and the Community
Amazon Machine Image (AMI)
Ubuntu Server 22.04 LTS (HVM), SSD Volume Type
Free tier eligible
ami-03f65b8614a860c29 (64-bit (x86)) / ami-0c79a55dda52434da (64-bit (Arm))
Virtualization: hvm ENA enabled: true Root device type: ebs

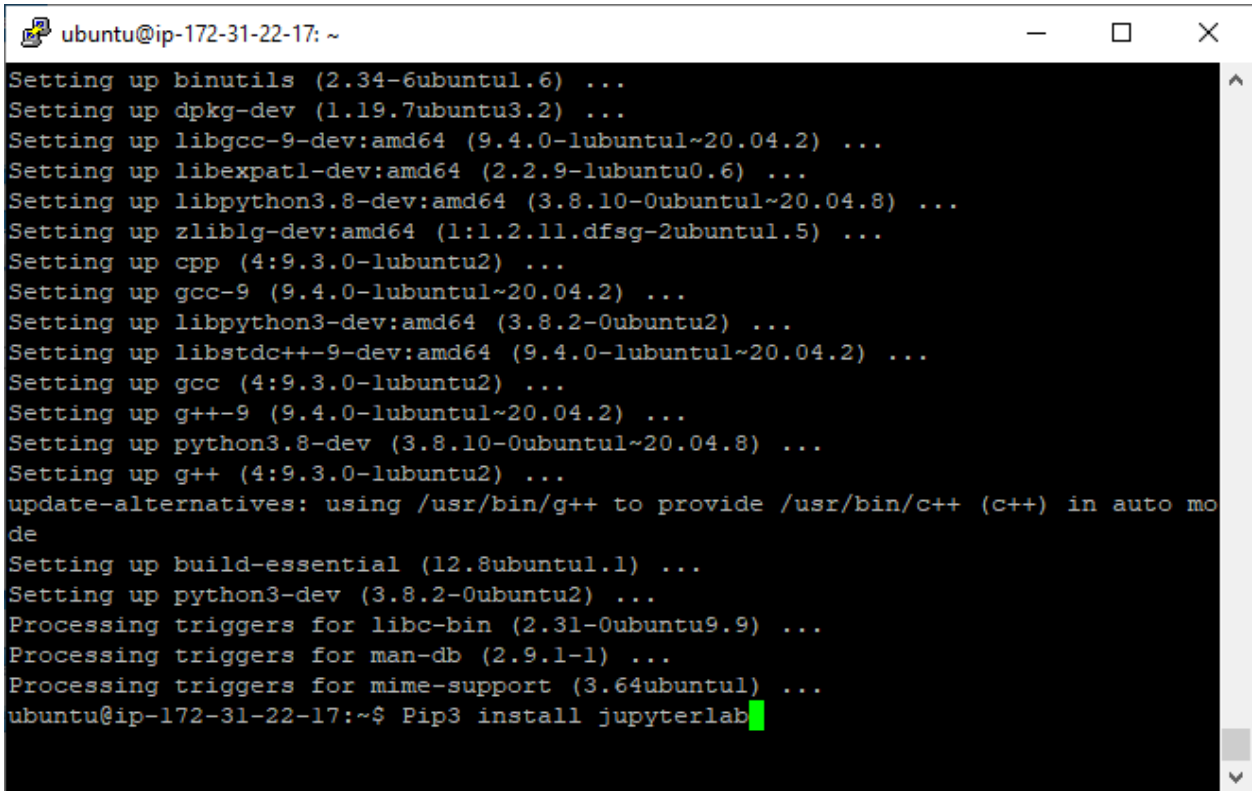

New EC2 Experience
Learn more
EC2 Dashboard
EC2 Global View
Events
Instances
Instances
Instance Types
Launch Templates
Spot Requests
Savings Plans
Reserved Instances
Dedicated Hosts
Capacity Reservations
Images
AMIs
Instances (1) Info
Connect
Instance state
Actions
Launch instances
Find instance by attribute or tag (case-sensitive)
Name
Instance ID
Instance state
Instance type
Status check
Alarm status
nw-animesh-hackethon
i-0f759a32160690445
Running
t2.micro
Initializing
No alarms
Select an instance

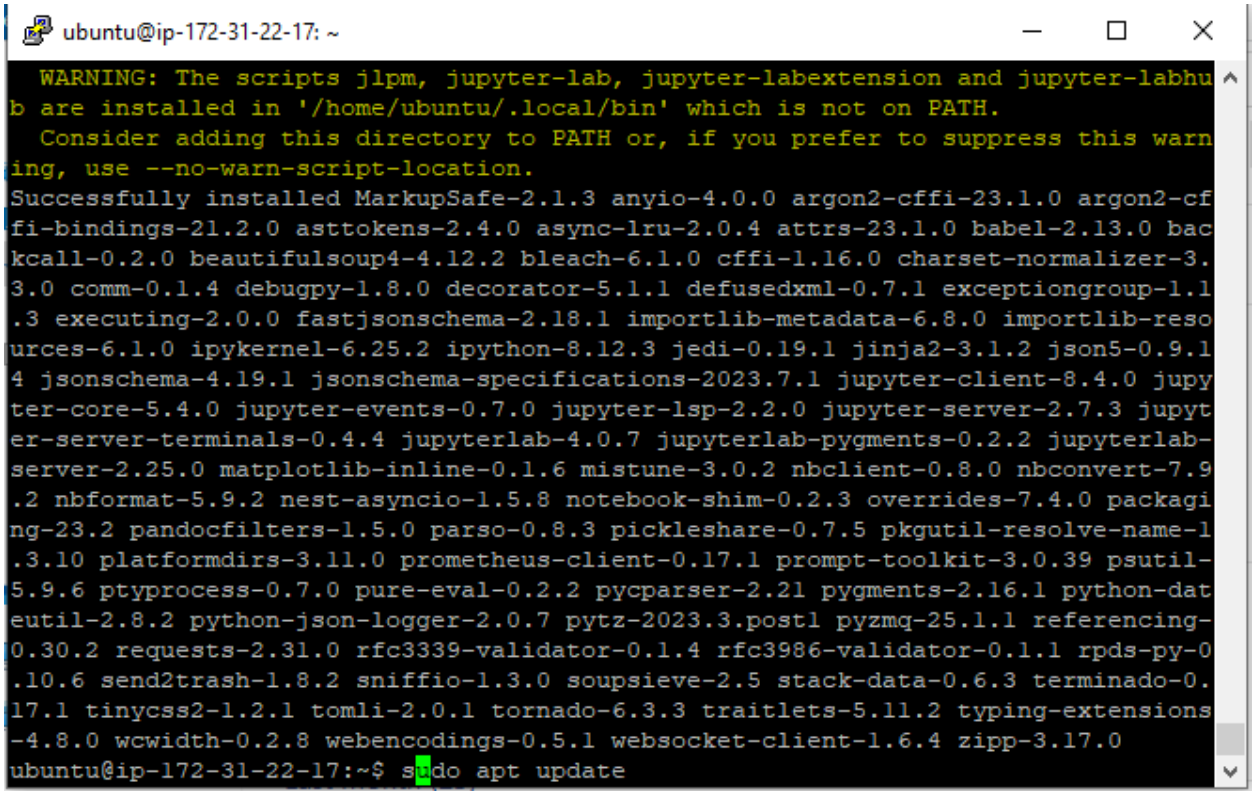

Instances
Instances
Instance Types
Launch Templates
Spot Requests
Savings Plans
Reserved Instances
Dedicated Hosts
Capacity Reservations
Images
AMIs
Instance: i-0f759a32160690445 (nw-animesh-hackethon)
Instance summary Info
Instance ID
i-0f759a32160690445 (nw-animesh-hackethon)
Public IPv4 address
34.212.5.171 |open address
Private IPv4 addresses
172.31.22.17
IPv6 address
–
Instance state
Running
Public IPv4 DNS
ec2-34-212-5-171.us-west-2.compute.amazonaws.com |open address
Hostname type
IP name: ip-172-31-22-17.us-west-
Private IP DNS name (IPv4 only)
ip-172-31-22-17.us-west-2.compute.internal

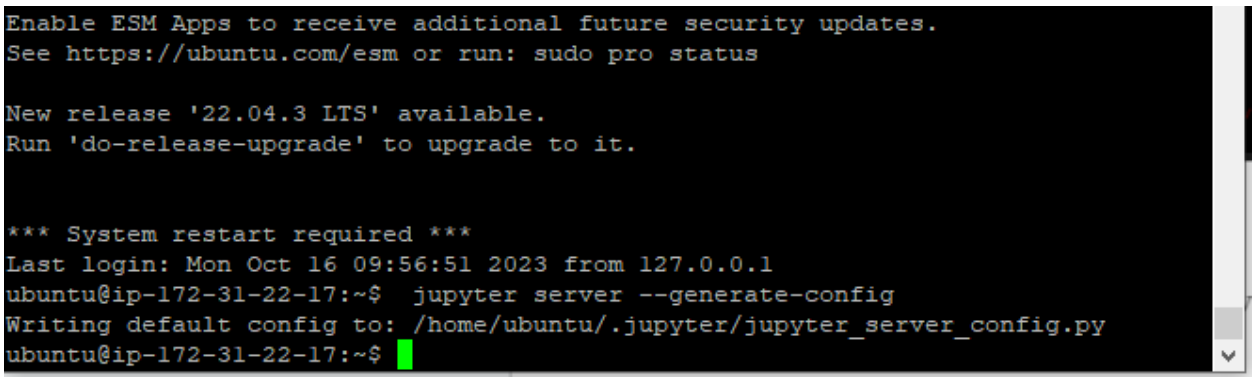

PuTTY Configuration
Category:
Session
Logging
Terminal
Keyboard
Bell
Features
Window
Appearance
Behaviour
Translation
Selection
Colours
Connection
Data
Proxy
SSH
Serial
Telnet
Rlogin
SUPDUP
Basic options for your PuTTY session
Specify the destination you want to connect to
Host Name (or IP address)
Port
ubuntu@34.212.5.171
22
Connection type:
SSH
Serial
Other:
Telnet
Load, save or delete a stored session
Saved Sessions
ec4
Default Settings
ec2
ec3
ec4
Load
Save
Delete
Close window on exit:
Always
Never
Only on clean exit
About
Help
Open
Cancel

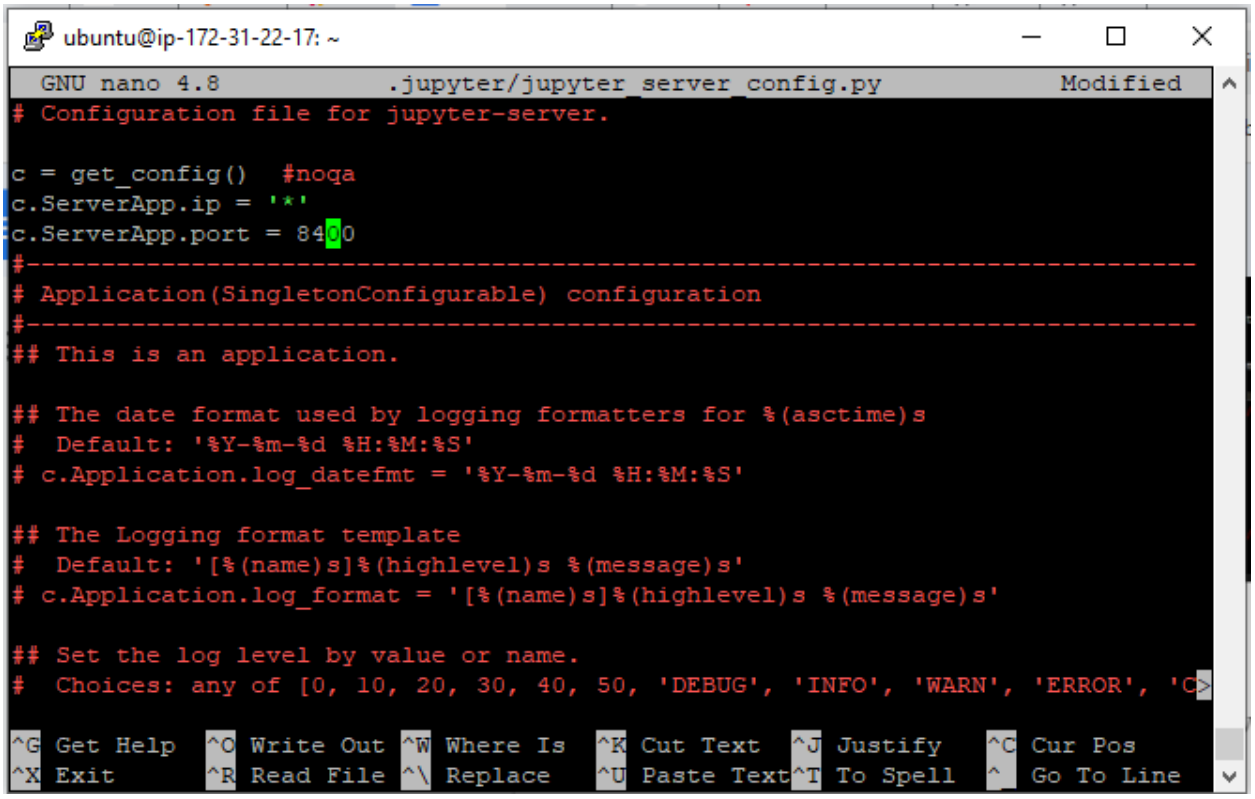

ubuntu@ip-172-31-22-17: ~
Expanded Security Maintenance for Applications is not enabled.
0 updates can be applied immediately.
Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status
The list of available updates is more than a week old.
To check 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-172-31-22-17:~$ ls
ubuntu@ip-172-31-22-17:~$ sudo apt update

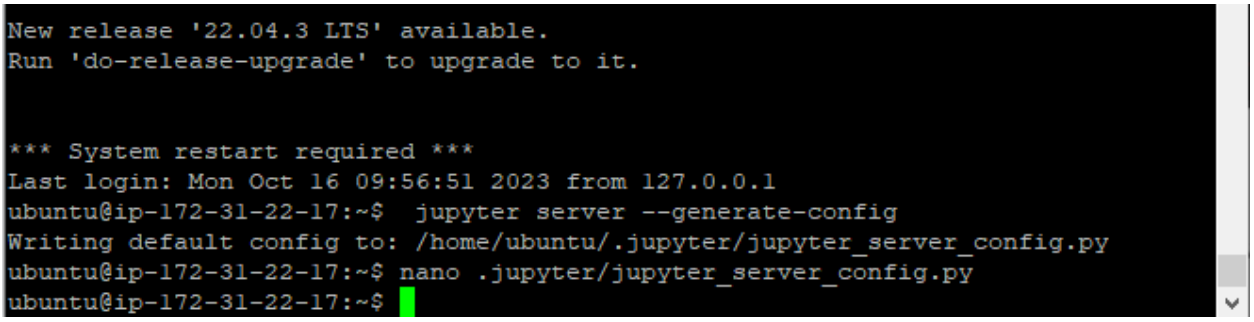

ubuntu@ip-172-31-22-17: ~
System information as of Mon Oct 16 09:56:51 UTC 2023
System load: 0.15 Processes: 102
Usage of /: 18.7% of 9.51GB Users logged in: 1
Memory usage: 25% IPv4 address for eth0: 172.31.22.17
Swap usage: 0%
Expanded Security Maintenance for Applications is not enabled.
115 updates can be applied immediately.
66 of these updates are standard security updates.
To see these additional updates run: apt list --upgradable
Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status
New release '22.04.3 LTS' available.
Run 'do-release-upgrade' to upgrade to it.
Last login: Mon Oct 16 09:54:48 2023 from 112.79.108.55
ubuntu@ip-172-31-22-17:~$ Sudo apt install python3-pip

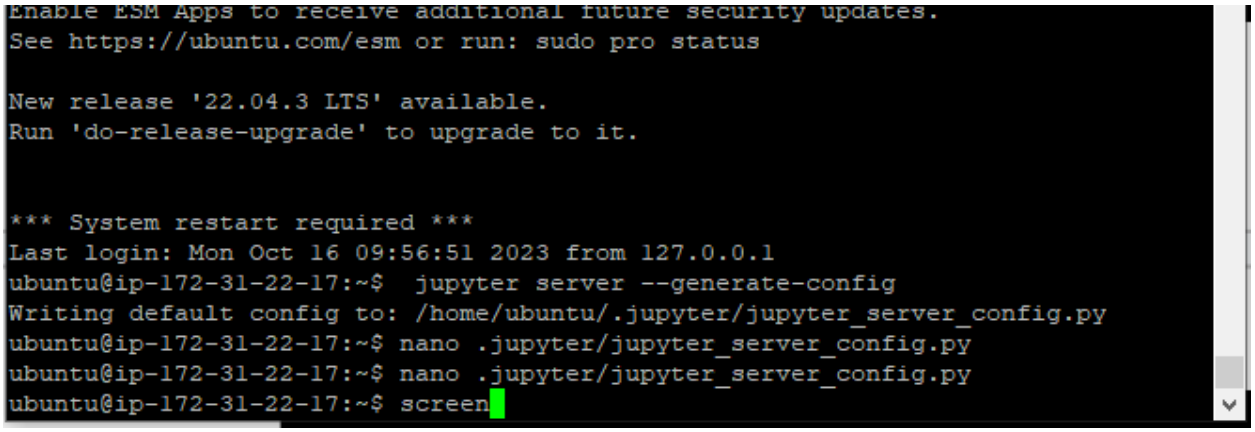

ubuntu@ip-172-31-22-17: ~
Setting up binutils (2.34-6ubuntu1.6) ...
Setting up dpkg-dev (1.19.7ubuntu3.2) ...
Setting up libgcc-9-dev:amd64 (9.4.0-1ubuntu1~20.04.2) ...
Setting up libexpat1-dev:amd64 (2.2.9-1ubuntu0.6) ...
Setting up libpython3.8-dev:amd64 (3.8.10-0ubuntu1~20.04.8) ...
Setting up zlib1g-dev:amd64 (1:1.2.11.dfsg-2ubuntu1.5) ...
Setting up cpp (4:9.3.0-1ubuntu2) ...
Setting up gcc-9 (9.4.0-1ubuntu1~20.04.2) ...
Setting up libpython3-dev:amd64 (3.8.2-0ubuntu2) ...
Setting up libstdc++-9-dev:amd64 (9.4.0-1ubuntu1~20.04.2) ...
Setting up gcc (4:9.3.0-1ubuntu2) ...
Setting up g++-9 (9.4.0-1ubuntu1~20.04.2) ...
Setting up python3.8-dev (3.8.10-0ubuntu1~20.04.8) ...
Setting up g++ (4:9.3.0-1ubuntu2) ...
update-alternatives: using /usr/bin/g++ to provide /usr/bin/c++ (c++) in auto mo
de
Setting up build-essential (12.8ubuntu1.1) ...
Setting up python3-dev (3.8.2-0ubuntu2) ...
Processing triggers for libc-bin (2.31-0ubuntu9.9) ...
Processing triggers for man-db (2.9.1-1) ...
Processing triggers for mime-support (3.64ubuntu1) ...
ubuntu@ip-172-31-22-17:~$ Pip3 install jupyterlab

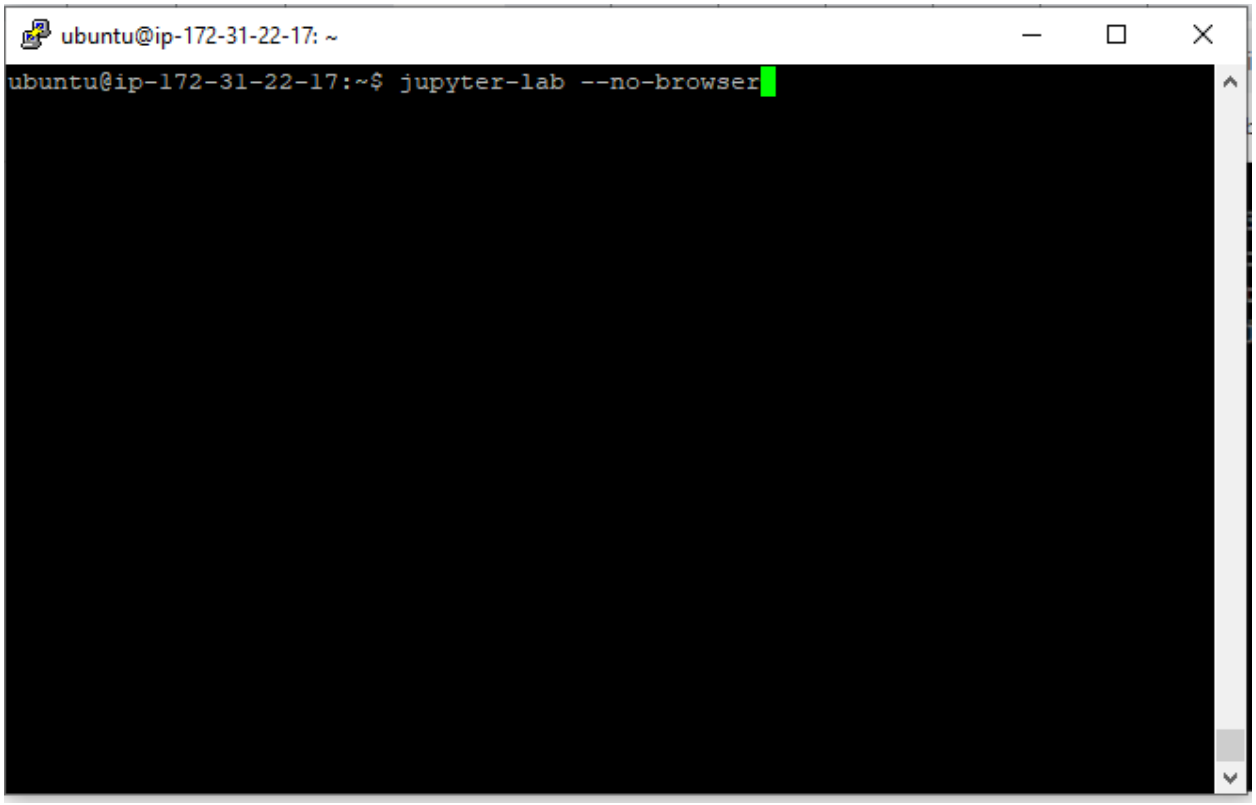

ubuntu@ip-172-31-22-17: ~
WARNING: The scripts jlpm, jupyter-lab, jupyter-labextension and jupyter-labhub are installed in '/home/ubuntu/.local/bin' which is not on PATH.
Consider adding this directory to PATH or, if you prefer to suppress this warning, use --no-warn-script-location.
Successfully installed MarkupSafe-2.1.3 anyio-4.0.0 argon2-cffi-23.1.0 argon2-cffi-bindings-21.2.0 asttokens-2.4.0 async-lru-2.0.4 attrs-23.1.0 babel-2.13.0 backcall-0.2.0 beautifulsoup4-4.12.2 bleach-6.1.0 cffi-1.16.0 charset-normalizer-3.3.0 comm-0.1.4 debugpy-1.8.0 decorator-5.1.1 defusedxml-0.7.1 exceptiongroup-1.1.3 executing-2.0.0 fastjsonschema-2.18.1 importlib-metadata-6.8.0 importlib-resources-6.1.0 ipykernel-6.25.2 ipython-8.12.3 jedi-0.19.1 jinja2-3.1.2 json5-0.9.14 jsonschema-4.19.1 jsonschema-specifications-2023.7.1 jupyter-client-8.4.0 jupyter-core-5.4.0 jupyter-events-0.7.0 jupyter-lsp-2.2.0 jupyter-server-2.7.3 jupyter-server-terminals-0.4.4 jupyterlab-4.0.7 jupyterlab-pygments-0.2.2 jupyterlab-server-2.25.0 matplotlib-inline-0.1.6 mistune-3.0.2 nbclient-0.8.0 nbconvert-7.9.2 nbformat-5.9.2 nest-asyncio-1.5.8 notebook-shim-0.2.3 overrides-7.4.0 packaging-23.2 pandocfilters-1.5.0 parso-0.8.3 pickleshare-0.7.5 pkgutil-resolve-name-1.3.10 platformdirs-3.11.0 prometheus-client-0.17.1 prompt-toolkit-3.0.39 psutil-5.9.6 ptyprocess-0.7.0 pure-eval-0.2.2 pycparser-2.21 pygments-2.16.1 python-dateutil-2.8.2 python-json-logger-2.0.7 pytz-2023.3.post1 pyzmq-25.1.1 referencing-0.30.2 requests-2.31.0 rfc3339-validator-0.1.4 rfc3986-validator-0.1.1 rpds-py-0.10.6 send2trash-1.8.2 sniffio-1.3.0 soupsieve-2.5 stack-data-0.6.3 terminado-0.17.1 tinycss2-1.2.1 tomli-2.0.1 tornado-6.3.3 traitlets-5.11.2 typing-extensions-4.8.0 wcwidth-0.2.8 webencodings-0.5.1 websocket-client-1.6.4 zipp-3.17.0
ubuntu@ip-172-31-22-17:~$ sudo apt update

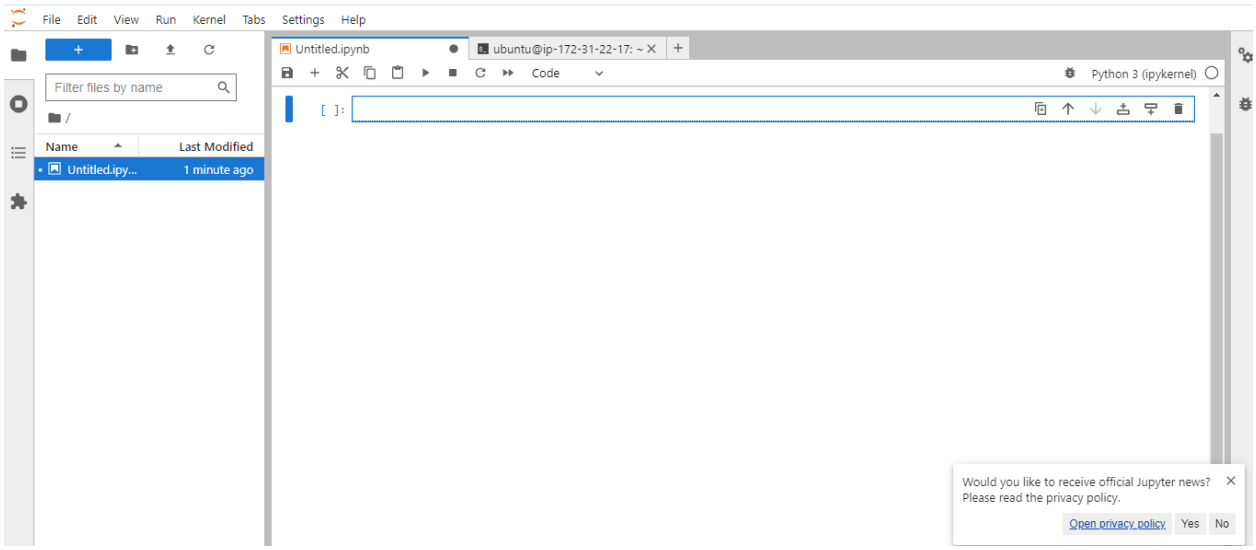

Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status
New release '22.04.3 LTS' available.
Run 'do-release-upgrade' to upgrade to it.
*** System restart required ***
Last login: Mon Oct 16 09:56:51 2023 from 127.0.0.1
ubuntu@ip-172-31-22-17:~$ jupyter server --generate-config
Writing default config to: /home/ubuntu/.jupyter/jupyter_server_config.py
ubuntu@ip-172-31-22-17:~$

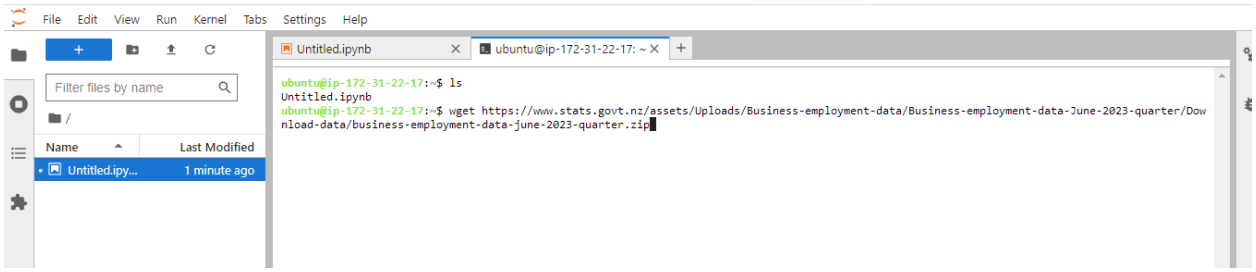

ubuntu@ip-172-31-22-17: ~
GNU nano 4.8 .jupyter/jupyter_server_config.py Modified
# Configuration file for jupyter-server.
c = get_config() #noqa
c.ServerApp.ip = '*'
c.ServerApp.port = 8400
# Application(SingletonConfigurable) configuration
This is an application.
The date format used by logging formatters for %(asctime)s
Default: '%Y-%m-%d %H:%M:%S'
c.Application.log_datefmt = '%Y-%m-%d %H:%M:%S'
The Logging format template
Default: '[%(name)s]%(highlevel)s %(message)s'
c.Application.log_format = '[%(name)s]%(highlevel)s %(message)s'
Set the log level by value or name.
Choices: any of [0, 10, 20, 30, 40, 50, 'DEBUG', 'INFO', 'WARN', 'ERROR', 'C
^G Get Help ^O Write Out ^W Where Is ^K Cut Text ^J Justify ^C Cur Pos
^X Exit ^R Read File ^\ Replace ^U Paste Text ^T To Spell ^_ Go To Line

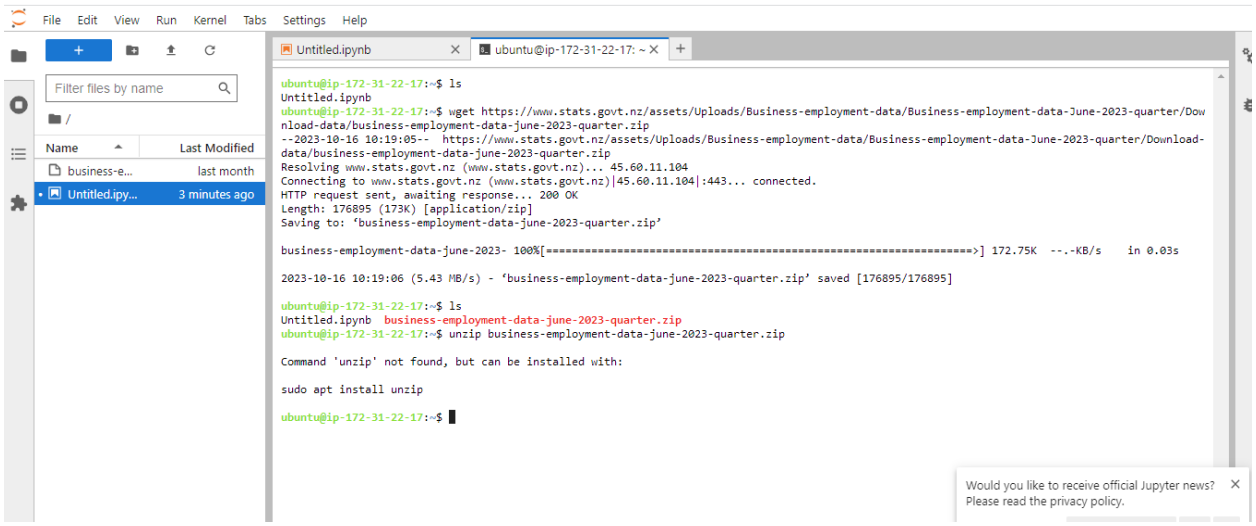

New release '22.04.3 LTS' available.
Run 'do-release-upgrade' to upgrade to it.
*** System restart required ***
Last login: Mon Oct 16 09:56:51 2023 from 127.0.0.1
ubuntu@ip-172-31-22-17:~$ jupyter server --generate-config
Writing default config to: /home/ubuntu/.jupyter/jupyter_server_config.py
ubuntu@ip-172-31-22-17:~$ nano .jupyter/jupyter_server_config.py
ubuntu@ip-172-31-22-17:~$


Enable ESM Apps to receive additional future security updates.
See https://ubuntu.com/esm or run: sudo pro status
New release '22.04.3 LTS' available.
Run 'do-release-upgrade' to upgrade to it.
*** System restart required ***
Last login: Mon Oct 16 09:56:51 2023 from 127.0.0.1
ubuntu@ip-172-31-22-17:~$ jupyter server --generate-config
Writing default config to: /home/ubuntu/.jupyter/jupyter_server_config.py
ubuntu@ip-172-31-22-17:~$ nano .jupyter/jupyter_server_config.py
ubuntu@ip-172-31-22-17:~$ nano .jupyter/jupyter_server_config.py
ubuntu@ip-172-31-22-17:~$ screen

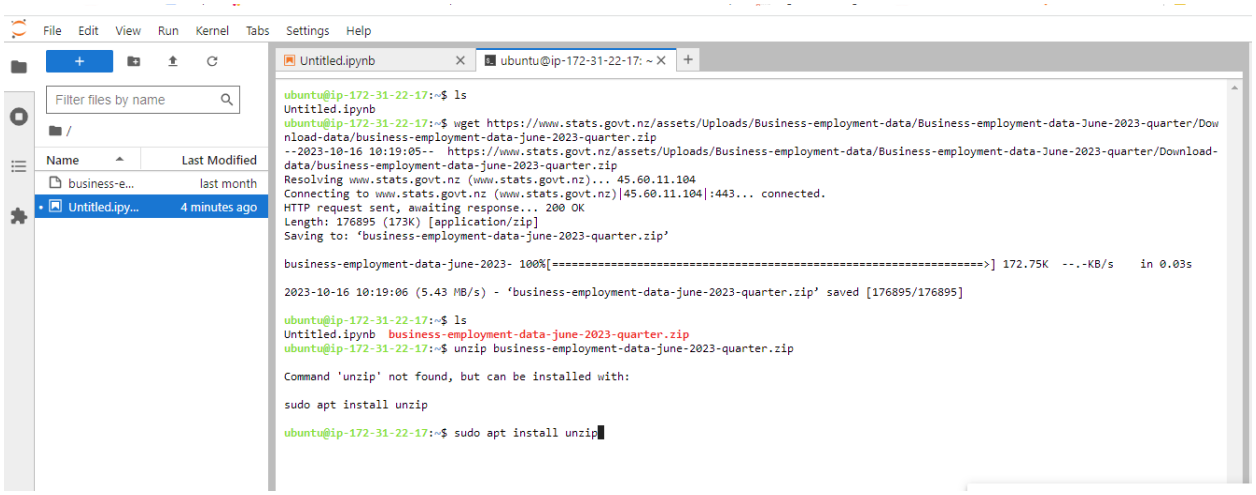

ubuntu@ip-172-31-22-17: ~
ubuntu@ip-172-31-22-17:~$ jupyter-lab --no-browser


File Edit View Run Kernel Tabs Settings Help
Filter files by name
/
Name
Last Modified
Untitled.ipy...
1 minute ago
Untitled.ipynb
ubuntu@ip-172-31-22-17: ~
Code
Python 3 (ipykernel)
[ ]:
Would you like to receive official Jupyter news?
Please read the privacy policy.
Open privacy policy
Yes
No

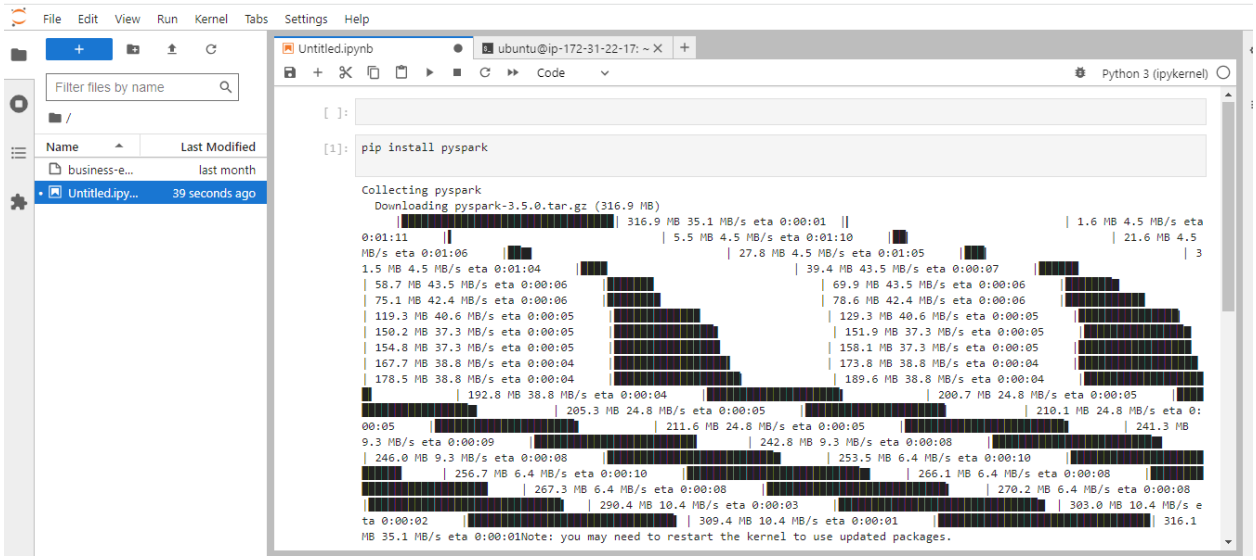




Downloading file



Installing py spark



Installing pandas

[14]:
```
import pandas as pd
p1=pd.read_csv('business-employment-data-june-2023-quarter.zip')
print(p1)
```

```
      Series_reference   Period  Data_value Suppressed STATUS   UNITS  \
0           BDCQ.SEA1AA  2011.06     80078.0        NaN      F  Number
1           BDCQ.SEA1AA  2011.09     78324.0        NaN      F  Number
2           BDCQ.SEA1AA  2011.12     85850.0        NaN      F  Number
3           BDCQ.SEA1AA  2012.03     90743.0        NaN      F  Number
4           BDCQ.SEA1AA  2012.06     81780.0        NaN      F  Number
...                 ...      ...         ...        ...    ...     ...
21797    BDCQ.SEE3999A  2017.06         NaN          Y      C  Number
21798    BDCQ.SEE3999A  2017.09         NaN          Y      C  Number
21799    BDCQ.SEE3999A  2017.12         NaN          Y      C  Number
21800    BDCQ.SEE3999A  2018.03         NaN          Y      C  Number
21801    BDCQ.SEE3999A  2018.06         NaN          Y      C  Number

       Magnitude                          Subject  \
0              0  Business Data Collection - BDC
1              0  Business Data Collection - BDC
2              0  Business Data Collection - BDC
3              0  Business Data Collection - BDC
4              0  Business Data Collection - BDC
...          ...                             ...
21797          0  Business Data Collection - BDC
21798          0  Business Data Collection - BDC
21799          0  Business Data Collection - BDC
21800          0  Business Data Collection - BDC
```

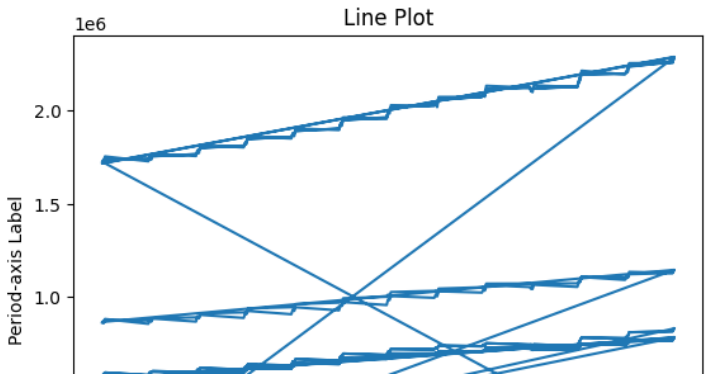
[15]:
```
df = pd.read_csv("business-employment-data-june-2023-quarter.zip")

print(df.head())
```

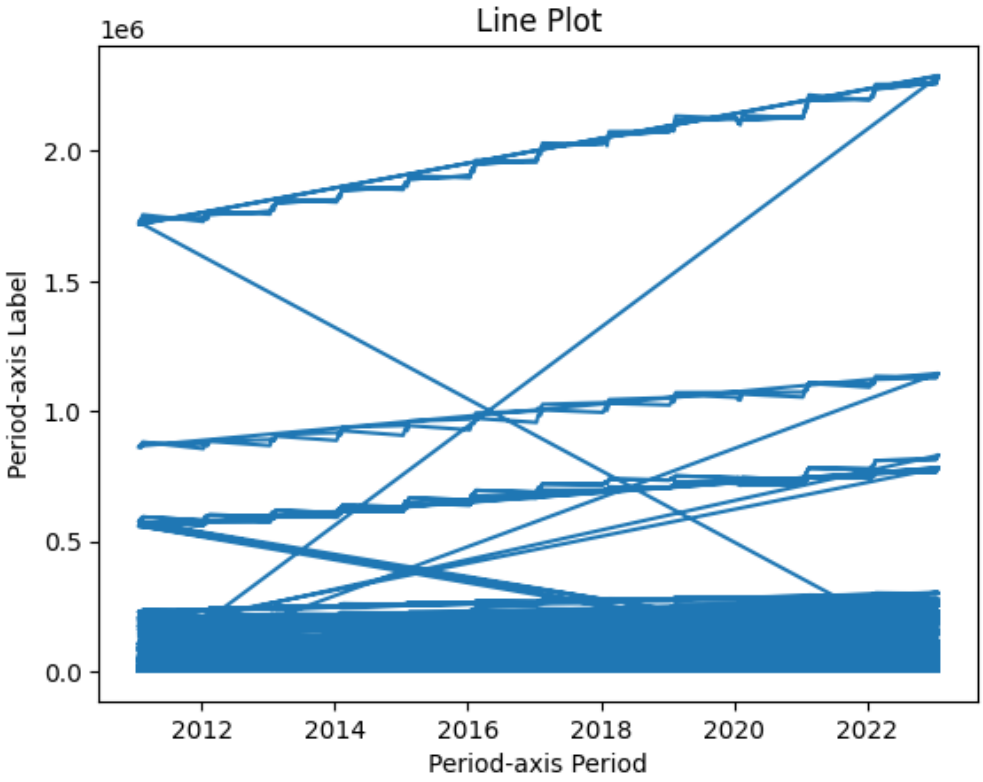
```
  Series_reference   Period  Data_value Suppressed STATUS   UNITS  Magnitude  \
0      BDCQ.SEA1AA  2011.06     80078.0        NaN      F  Number          0
1      BDCQ.SEA1AA  2011.09     78324.0        NaN      F  Number          0
2      BDCQ.SEA1AA  2011.12     85850.0        NaN      F  Number          0
3      BDCQ.SEA1AA  2012.03     90743.0        NaN      F  Number          0
4      BDCQ.SEA1AA  2012.06     81780.0        NaN      F  Number          0

                        Subject                            Group  \
0  Business Data Collection - BDC  Industry by employment variable
1  Business Data Collection - BDC  Industry by employment variable
2  Business Data Collection - BDC  Industry by employment variable
3  Business Data Collection - BDC  Industry by employment variable
4  Business Data Collection - BDC  Industry by employment variable
```

[17]:
```
pip install matplotlib
```

```
Collecting matplotlib
  Downloading matplotlib-3.7.3-cp38-cp38-manylinux_2_12_x86_64.manylinux2010_x86_64.whl (9.2 MB)
     |████████████████████████████████| 9.2 MB 5.8 MB/s eta 0:00:01     |████████████████                | 5.3 MB 5.8 MB/s eta
0:00:01
Requirement already satisfied: numpy<2,>=1.20 in ./.local/lib/python3.8/site-packages (from matplotlib) (1.24.4)
Collecting fonttools>=4.22.0
  Downloading fonttools-4.43.1-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (4.6 MB)
     |████████████████████████████████| 4.6 MB 24.6 MB/s eta 0:00:01
Requirement already satisfied: python-dateutil>=2.7 in ./.local/lib/python3.8/site-packages (from matplotlib) (2.8.2)
Collecting kiwisolver>=1.0.1
  Downloading kiwisolver-1.4.5-cp38-cp38-manylinux_2_5_x86_64.manylinux1_x86_64.whl (1.2 MB)
     |████████████████████████████████| 1.2 MB 37.3 MB/s eta 0:00:01
Collecting pyparsing>=2.3.1
  Downloading pyparsing-3.1.1-py3-none-any.whl (103 kB)
     |████████████████████████████████| 103 kB 37.4 MB/s eta 0:00:01
Collecting pillow>=6.2.0
  Downloading Pillow-10.1.0-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (3.5 MB)
     |████████████████████████████████| 3.5 MB 20.0 MB/s eta 0:00:01
```

[22]:
```
import matplotlib.pyplot as plt

# Example: Plot 'x' against 'y' columns from your DataFrame
plt.plot(df['Period'], df['Data_value'])
plt.xlabel('Period-axis Period')
plt.ylabel('Period-axis Label')
plt.title('Line Plot')
plt.show()
```





Untitled.ipynb | Untitled1.ipynb | ubuntu@ip-172-31-22-17: ~ | Code | Python 3 (ipykernel)

```
[1]: pip install pyspark

Collecting pyspark
  Downloading pyspark-3.5.0.tar.gz (316.9 MB)
     |████████████████████████████████| 316.9 MB 34.3 MB/s eta 0:00:01     |▌                               | 13.1 MB 4.9 MB/s e
ta 0:01:02     |▋                               | 17.0 MB 4.9 MB/s eta 0:01:01     |▊                               | 26.3 MB
4.9 MB/s eta 0:00:59     |▉                               | 28.1 MB 4.9 MB/s eta 0:00:59     |▉
| 30.1 MB 4.9 MB/s eta 0:00:59     |▉                               | 39.5 MB 45.9 MB/s eta 0:00:07     |▉
| 45.0 MB 45.9 MB/s eta 0:00:06     |▉                               | 64.9 MB 45.9 MB/s eta 0:00:06     |▉
| 75.1 MB 41.4 MB/s eta 0:00:06     |▉                               | 88.0 MB 41.4 MB/s eta 0:00:06     |▉
| 89.6 MB 41.4 MB/s eta 0:00:06     |▉                               | 96.1 MB 41.4 MB/s eta 0:00:06     |▉
| 102.5 MB 43.3 MB/s eta 0:00:05     |▉                               | 105.7 MB 43.3 MB/s eta 0:00:05     |▉
| 107.4 MB 43.3 MB/s eta 0:00:05     |▉                               | 114.2 MB 43.3 MB/s eta 0:00:05     |▉
| 116.9 MB 43.3 MB/s eta 0:00:05     |▉                               | 127.0 MB 43.3 MB/s eta 0:00:05     |▉
| 132.8 MB 34.6 MB/s eta 0:00:06     |▉                               | 137.1 MB 34.6 MB/s eta 0:00:06     |▉
| 140.1 MB 34.6 MB/s eta 0:00:06     |▉                               | 144.9 MB 34.6 MB/s eta 0:00:05     |▉
| 148.2 MB 34.6 MB/s eta 0:00:05     |▉                               | 161.3 MB 12.5 MB/s eta 0:00:13     |▉
|▉                               | 171.4 MB 12.5 MB/s eta 0:00:12     |▉                               | 178.0 MB 12.5 MB/s eta 0:00:12     |▉
|▉                               | 192.5 MB 38.3 MB/s eta 0:00:04     |▉                               | 205.4 MB 38.3 MB/s eta
0:00:03     |▉                               | 215.3 MB 38.3 MB/s eta 0:00:03     |▉                               | 217.0 MB
38.3 MB/s eta 0:00:03     |▉                               | 222.0 MB 33.3 MB/s eta 0:00:03     |▉
| 228.5 MB 33.3 MB/s eta 0:00:03     |▉                               | 233.5 MB 33.3 MB/s eta 0:00:03     |▉
|▉                               | 236.7 MB 33.3 MB/s eta 0:00:03     |▉                               | 240.1 MB 33.3 MB/s eta 0:00:03     |▉
|▉                               | 243.5 MB 33.3 MB/s eta 0:00:03     |▉                               | 266.6 MB 18.9 MB/s eta 0:0
0:03     |▉                               | 268.2 MB 18.9 MB/s eta 0:00:03     |▉                               | 269.7 MB 18.
9 MB/s eta 0:00:03     |▉                               | 271.3 MB 18.9 MB/s eta 0:00:03     |▉
| 276.4 MB 18.9 MB/s eta 0:00:03     |▉                               | 277.6 MB 18.9 MB/s eta 0:00:03     |▉
```

[13]:
```
import pandas as pd
```

[14]:
```
import pyspark
p1=pd.read_csv('business-employment-data-june-2023-quarter.zip')
print(p1)
```

```
      Series_reference   Period  Data_value Suppressed STATUS   UNITS  \
0          BDCQ.SEA1AA  2011.06     80078.0        NaN      F  Number
1          BDCQ.SEA1AA  2011.09     78324.0        NaN      F  Number
2          BDCQ.SEA1AA  2011.12     85850.0        NaN      F  Number
3          BDCQ.SEA1AA  2012.03     90743.0        NaN      F  Number
4          BDCQ.SEA1AA  2012.06     81780.0        NaN      F  Number
...                ...      ...         ...        ...    ...     ...
21797    BDCQ.SEE3999A  2017.06         NaN          Y      C  Number
21798    BDCQ.SEE3999A  2017.09         NaN          Y      C  Number
21799    BDCQ.SEE3999A  2017.12         NaN          Y      C  Number
21800    BDCQ.SEE3999A  2018.03         NaN          Y      C  Number
21801    BDCQ.SEE3999A  2018.06         NaN          Y      C  Number

       Magnitude                          Subject  \
0              0  Business Data Collection - BDC
1              0  Business Data Collection - BDC
2              0  Business Data Collection - BDC
3              0  Business Data Collection - BDC
4              0  Business Data Collection - BDC
...          ...                             ...
```

[22]:
```
spark=SparkSession.builder.appName('zenlite') .getOrCreate()
```

[23]:
```
spark
```

[23]:

SparkSession - in-memory

SparkContext

Spark UI

Version	`v3.5.0`
Master	`local[*]`
AppName	`zenlite`

[ ]:

[29]:
```
f1=spark.read.csv('machine-readable-business-employment-data-jun-2023-quarter.csv')
```

[30]:
```
f1.show()
```

```
+----------------+-------+----------+----------+------+------+---------+--------------------+--------------------+----------
----+--------------------+--------------+--------------+--------------+
|             _c0|    _c1|       _c2|       _c3|   _c4|   _c5|      _c6|                 _c7|                 _c8|
_c9|                _c10|          _c11|          _c12|          _c13|
+----------------+-------+----------+----------+------+------+---------+--------------------+--------------------+----------
----+--------------------+--------------+--------------+--------------+
|Series_reference| Period|Data_value|Suppressed|STATUS| UNITS|Magnitude|             Subject|               Group|Series_tit
le_1|      Series_title_2|Series_title_3|Series_title_4|Series_title_5|
|     BDCQ.SEA1AA|2011.06|     80078|      NULL|     F|Number|        0|Business Data Col...|Industry by emplo...|   Filled
jobs|Agriculture, Fore...|        Actual|          NULL|          NULL|
|     BDCQ.SEA1AA|2011.09|     78324|      NULL|     F|Number|        0|Business Data Col...|Industry by emplo...|   Filled
jobs|Agriculture, Fore...|        Actual|          NULL|          NULL|
|     BDCQ.SEA1AA|2011.12|     85850|      NULL|     F|Number|        0|Business Data Col...|Industry by emplo...|   Filled
jobs|Agriculture, Fore...|        Actual|          NULL|          NULL|
|     BDCQ.SEA1AA|2012.03|     90743|      NULL|     F|Number|        0|Business Data Col...|Industry by emplo...|   Filled
jobs|Agriculture, Fore...|        Actual|          NULL|          NULL|
|     BDCQ.SEA1AA|2012.06|     81780|      NULL|     F|Number|        0|Business Data Col...|Industry by emplo...|   Filled
jobs|Agriculture, Fore...|        Actual|          NULL|          NULL|
|     BDCQ.SEA1AA|2012.09|     79261|      NULL|     F|Number|        0|Business Data Col...|Industry by emplo...|   Filled
jobs|Agriculture, Fore...|        Actual|          NULL|          NULL|
|     BDCQ.SEA1AA|2012.12|     87793|      NULL|     F|Number|        0|Business Data Col...|Industry by emplo...|   Filled
jobs|Agriculture, Fore...|        Actual|          NULL|          NULL|
|     BDCQ.SEA1AA|2013.03|     91571|      NULL|     F|Number|        0|Business Data Col...|Industry by emplo...|   Filled
```

[31]:
```
f2=spark.read.option('header', 'true').csv('machine-readable-business-employment-data-jun-2023-quarter.csv')
```

[32]:
```
f2.show()
```

```
+----------------+-------+----------+----------+------+------+---------+--------------------+--------------------+----------
----+--------------------+--------------+--------------+--------------+
|Series_reference| Period|Data_value|Suppressed|STATUS| UNITS|Magnitude|             Subject|               Group|Series_tit
le_1|      Series_title_2|Series_title_3|Series_title_4|Series_title_5|
+----------------+-------+----------+----------+------+------+---------+--------------------+--------------------+----------
----+--------------------+--------------+--------------+--------------+
|     BDCQ.SEA1AA|2011.06|     80078|      NULL|     F|Number|        0|Business Data Col...|Industry by emplo...|   Filled
jobs|Agriculture, Fore...|        Actual|          NULL|          NULL|
```