



**Tecnológico
de Monterrey**

Instituto Tecnológico y de Estudios Superiores de Monterrey

AWS E2C y Sparks

Diego Alberto Baños Lopez

A01275100

Monterrey, Nuevo León, México 29 de Octubre del 2023

Inteligencia artificial avanzada para la ciencia de datos II (Gpo 501)

Panel de EC2

Vista global de EC2

Eventos

▼ Instancias

Instancias

Tipos de Instancia

Plantillas de lanzamiento

Solicitudes de spot

Savings Plans

Instancias reservadas

Alojamientos dedicados

Reservas de capacidad

▼ Imágenes

AMI

Catálogo de AMI

▼ Elastic Block Store

Volúmenes

Instantáneas

Administrador del ciclo de vida

▼ Red y seguridad

Security Groups

Direcciones IP elásticas

Grupos de ubicación

Pares de claves

Interfases de red

▼ Equilibrio de carga

Instancias (1) Información

Buscar instancia por atributo o etiqueta (case-sensitive)

Conectar Estado de la instancia Acciones Lanzar instancias

<input type="checkbox"/>	Name	ID de la instancia	Estado de la instancia	Tipo de inst...	Comprobación ...	Estado de la ...	Zona de dispon...	DNS de IPv4 pública
<input type="checkbox"/>	escuelaInstancia	i-0e4f12370ff71a104	En ejecución	t2.micro	2/2 comprobador	Sin alarmas +	us-east-2a	ec2-5-137-210-35-us-east-2.compute.amazonaws.com

Seleccione una instancia

© 2023 Amazon Web Services, Inc. o sus filiales. Privacidad Términos Preferencias de cookies

```
(ubuntu) ec2-3-137-210-35-us-east-2.compute.amazonaws.com — Console

Nueva pestaña Dividir vista Copiar Pegar Buscar

* Management: https://landscape.canonical.com
* Support: https://ubuntu.com/advantage

System information as of Mon Oct 30 04:42:50 UTC 2023

System load: 0.0 Processes: 95
Usage of /: 6.0% of 25.99GB Users logged in: 0
Memory usage: 21% IPv4 address for eth0: 172.31.1.244
Swap usage: 0%

* Ubuntu Pro delivers the most comprehensive open source security and
  compliance features.
  https://ubuntu.com/aws/pro

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-172-31-1-244:~$ ls -l
total 0
ubuntu@ip-172-31-1-244:~$
```

Panel de EC2

Vista global de EC2

Eventos

Instancias

Instancias

Tipos de instancia

Plantillas de lanzamiento

Solicitudes de spot

Savings Plans

Instancias reservadas

Alojamientos dedicados

Reservas de capacidad

Imágenes

AMI

Catálogo de AMI

Elastic Block Store

Volúmenes

Instantáneas

Administrador del ciclo de vida

Red y seguridad

Security Groups

Direcciones IP elásticas

Grupos de ubicación

Pares de claves

Interfases de red

Equilibrio de carga

EC2 > Instancias > i-0e4f12370ff71a104

Resumen de instancia de i-0e4f12370ff71a104 (escuelainstancia) Información

Conectar

Estado de la instancia

Acciones

Se ha actualizado hace less than a minute

ID de la instancia

1-0e4f12370ff71a104 (escuelainstancia)

Dirección IPv4 pública

3.137.210.35 dirección abierta

Direcciones IPv4 privadas

172.31.1.244

Dirección IPv6

-

Estado de la instancia

En ejecución

DNS de IPv4 pública

ec2-3-137-210-35.us-east-2.compute.amazonaws.com dirección abierta

Tipo de nombre de anfitrión

Nombre de IP: ip-172-31-1-244.us-east-2.compute.internal

Nombre DNS de IP privada (solo IPv4)

ip-172-31-1-244.us-east-2.compute.internal

Direcciones IP elásticas

-

Responder al nombre DNS de recurso privado IPv4 (A)

-

Tipo de instancia

t2.micro

Hallazgo de AWS Compute Optimizer

Suscribirse a AWS Compute Optimizer para recibir recomendaciones. Más información

Dirección IP asignada automáticamente

3.137.210.35 [IP pública]

ID de VPC

vpc-02c6c71f6e2b92d6c

Nombre del grupo de Auto Scaling

-

Rol de IAM

-

ID de subred

subnet-0b1b59dab015c8ad9

IMDSv2

Optional

Detalles

Seguridad

Redes

Almacenamiento

Comprobaciones de estado

Monitoreo

Etiquetas

Detalles de la instancia Información

Plataforma

Ubuntu (inferido)

Monitoreo

desactivado

Detalles de la plataforma

Linux/UNIX

Protección de terminación

desactivado

Detener la protección

desactivado

Horario de lanzamiento

Sun Oct 29 2023 17:03:46 GMT-0600 (hora estándar central) (about 6 hours)

Ubicación de AMI

amazon/ubuntu/images/hvm-ssd/ubuntu-jammy-22.04-amd64-server-20230919

CloudShell

Comentarios

© 2023, Amazon Web Services, Inc. o sus filiales. Privacidad Términos Preferencias de cookies

```
base) ubuntu@ip-172-31-1-244:~$ jupyter notebook
[05:15:24.888 NotebookApp] Writing notebook server cookie secret to /home/ubuntu/.local/share/jupyter/runtime/notebook_cookie_secret

[05:15:24.888 NotebookApp]
[05:15:24.888 NotebookApp]
[05:15:24.888 NotebookApp]

Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions.
https://jupyter-notebook.readthedocs.io/en/latest/migrate_to_notebook7.html

Please note that updating to Notebook 7 might break some of your extensions.

[05:15:24.589 NotebookApp] WARNING: The notebook server is listening on all IP addresses and not using encryption. This is not recommended.
[05:15:25.436 NotebookApp] Loading JupyterLab as a classic notebook (v6) extension.
[05:15:25.436 NotebookApp] 'ip' has moved from NotebookApp to ServerApp. This config will be passed to ServerApp. Be sure to update your config before our next release.
[05:15:25.436 NotebookApp] 'port' has moved from NotebookApp to ServerApp. This config will be passed to ServerApp. Be sure to update your config before our next release.
[05:15:25.436 NotebookApp] 'port' has moved from NotebookApp to ServerApp. This config will be passed to ServerApp. Be sure to update your config before our next release.
[05:15:25.442 NotebookApp] JupyterLab extension loaded from /home/ubuntu/anaconda3/lib/python3.11/site-packages/jupyterlab
[05:15:25.442 NotebookApp] JupyterLab application directory is /home/ubuntu/anaconda3/share/jupyterlab
[05:15:27.889 NotebookApp] Serving notebook from local directory: /home/ubuntu
[05:15:27.889 NotebookApp] Jupyter Notebook 6.5.4 is running at:
[05:15:27.889 NotebookApp] http://ip-172-31-1-244:8888/?token=ddc659abf85fd8c4ff08c2f53f6f9afb3c98e2a0b186ff78
[05:15:27.889 NotebookApp] or http://127.0.0.1:8888/?token=ddc659abf85fd8c4ff08c2f53f6f9afb3c98e2a0b186ff78
[05:15:27.889 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).

[C 05:15:27.812 NotebookApp]

To access the notebook, open this file in a browser:
file:///home/ubuntu/.local/share/jupyter/runtime/nbserver-21135-open.html
Or copy and paste one of these URLs:
http://ip-172-31-1-244:8888/?token=ddc659abf85fd8c4ff08c2f53f6f9afb3c98e2a0b186ff78
or http://127.0.0.1:8888/?token=ddc659abf85fd8c4ff08c2f53f6f9afb3c98e2a0b186ff78
```

No seguro | ec2-3-137-210-35.us-east-2.compute.amazonaws.com:8888/tree/ejemplos_notebooks

UPDATE

Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions. - Please note that updating to Notebook 7 might break some of your extensions.

Don't show anymore

jupyter

Quit Logout

Files Running Clusters

Select items to perform actions on them.

Upload New

0 / ejemplos_notebooks

hace unos segundos

hace 10 minutos 16.9 kB

hace 4 minutos 610 B

hace 10 minutos 44.6 kB

hace 10 minutos 4.24 kB

hace 10 minutos 4.33 kB

hace 10 minutos 18.2 kB

hace 10 minutos 26.7 kB

hace 10 minutos 8.94 kB

hace 10 minutos 5.92 kB

hace 10 minutos 17.1 kB

hace 10 minutos 1.82 kB

hace 10 minutos 24.8 kB

hace 10 minutos 9.91 kB

hace 10 minutos 13.1 kB

hace 10 minutos 12.5 kB

hace 10 minutos 10.1 kB

hace 10 minutos 192 kB

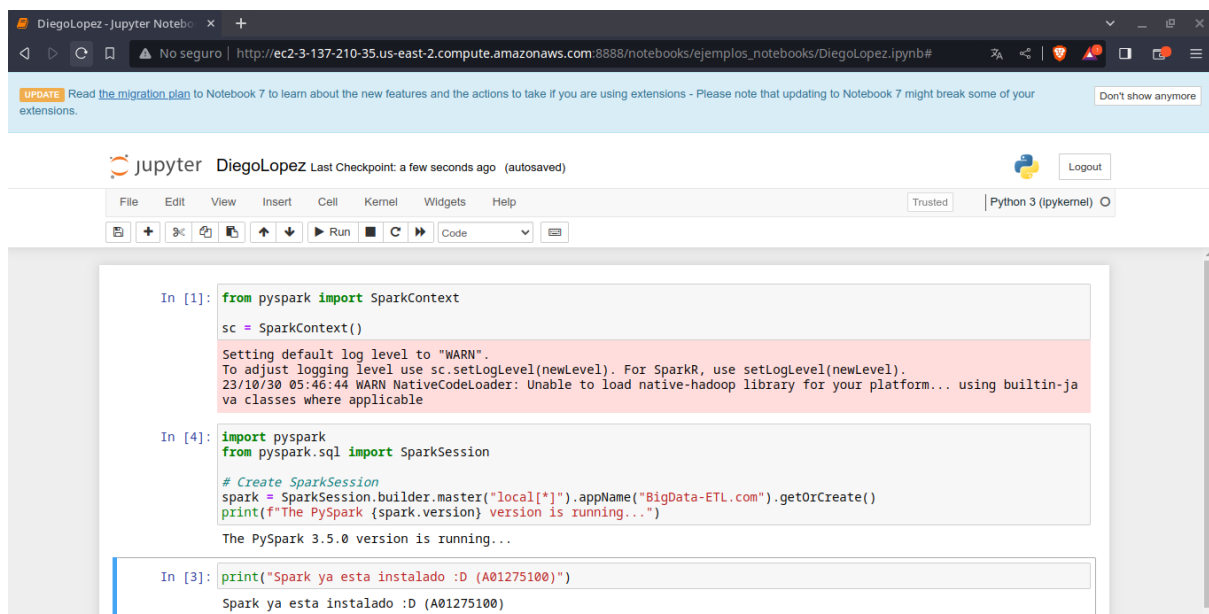
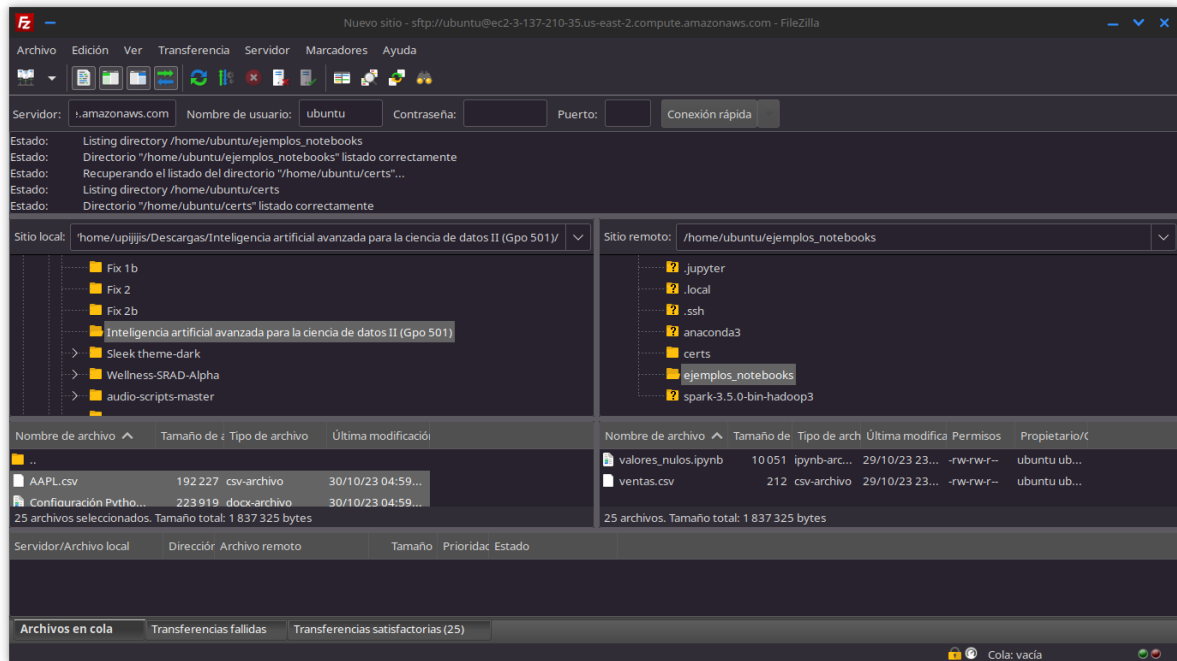
hace 10 minutos 224 kB

hace 10 minutos 308 kB

hace 10 minutos 86.9 kB

hace 10 minutos 55 B

hace 10 minutos 688 kB



```
In [4]: spark = SparkSession.builder.appName('DataFrame').getOrCreate()
```

```
23/11/25 20:19:59 WARN SparkSession: Using an existing Spark session; only runtime SQL configurations will take effect.
```

```
In [5]: df = spark.read.csv('ventas.csv', inferSchema = True, header = True)
```

```
In [6]: df.show()
```

```
+-----+-----+-----+
|Empresa|Persona|Ventas|
+-----+-----+-----+
|GOOG|Carlos|200|
|GOOG|Juan|120|
|GOOG|Felipe|340|
|MSFT|Tina|600|
|MSFT|Andrea|124|
|MSFT|Carla|243|
|FB|Sara|870|
|FB|Ignacio|350|
|APPL|Miguel|250|
|APPL|Oscar|130|
|APPL|Jorge|750|
|APPL|Ivan|350|
+-----+-----+-----+
```

```
In [7]: df.groupBy('Empresa').sum().show()
```

```
[Stage 3:>                                     (0 + 1) / 1]
```

```
+-----+-----+
|Empresa|sum(Ventas)|
+-----+-----+
|APPL|1480|
|GOOG|660|
|FB|1220|
|MSFT|967|
+-----+-----+
```

```
In [8]: df.agg({'Ventas': 'max'}).show()
```

```
+-----+
|max(Ventas)|
+-----+
|870|
+-----+
```

```
In [9]: agrupado = df.groupBy('Empresa')
agrupado.agg({'Ventas': 'max'}).show()
```

```
+-----+-----+
|Empresa|max(Ventas)|
+-----+-----+
|APPL|750|
|GOOG|340|
|FB|870|
|MSFT|600|
+-----+-----+
```

```
In [10]: agrupado.agg({'Ventas': 'min'}).show()
```

```
+-----+-----+
|Empresa|min(Ventas)|
+-----+-----+
|APPL|130|
|GOOG|120|
|FB|350|
|MSFT|124|
+-----+-----+
```

```

In [11]: df_agrupado = agrupado.agg({'Ventas': 'max'})
df_agrupado_pandas = df_agrupado.toPandas()

In [15]: df_agrupado_pandas = df_agrupado_pandas.sort_values('max(Ventas)', ascending=False)
import matplotlib.pyplot as plt

plt.figure(figsize=(10,6))
plt.bar(df_agrupado_pandas['Empresa'], df_agrupado_pandas['max(Ventas)'])
plt.xlabel('Empresa')
plt.ylabel('Máximo de Ventas')
plt.title('Máximo de Ventas por Empresa')
plt.xticks(rotation=45) # Rotar las etiquetas si son muchas o largas
plt.show()

```

