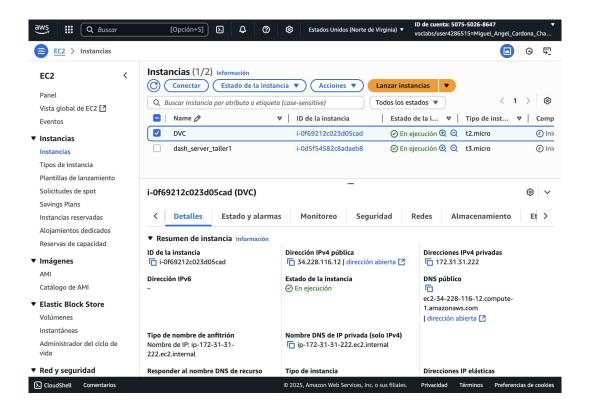
Data Version Control

Miguel Angel Cardona Chamorro

Lanzamiento de instancia DVC



Instalación de Python

```
    master — ec2-user@ip-172-31-31-222:~ — ssh -i ./Downloads/dvc.keys.pem ec2-user@34.228.116.12 — 164×48

Last login: Wed Aug 13 22:03:04 on ttys000
[master@MacBook-Air-de-master ~ % ssh -i ./Downloads/dvc.kevs.pem ec2-user@34.228.116.12
       ####
                    Amazon Linux 2023
  ~~ \ #####\
        \###1
                    https://aws.amazon.com/linux/amazon-linux-2023
          \#/
           V~' '->
Last login: Thu Aug 14 03:03:29 2025 from 191.104.216.209
[[ec2-user@ip-172-31-31-222 ~]$ sudo yum update -y
Last metadata expiration check: 0:01:16 ago on Thu Aug 14 03:03:41 2025.
Dependencies resolved.
Nothing to do.
Complete!
[[ec2-user@ip-172-31-31-222 ~]$ sudo yum install python3-pip
Last metadata expiration check: 0:01:24 ago on Thu Aug 14 03:03:41 2025.
Dependencies resolved.
 Package
                         Architecture Version
                                                                        Repository
                                                                                             Size
Installing:
                                       21.3.1-2.amzn2023.0.13
                                                                                            1.8 M
 python3-pip
                         noarch
                                                                        amazonlinux
Installing weak dependencies:
                                                                        amazonlinux
 libxcrypt-compat
                                        4.4.33-7.amzn2023
                                                                                             92 k
Transaction Summary
Install 2 Packages
Total download size: 1.9 M
Installed size: 11 M
[Is this ok [y/N]: y
Downloading Packages:
(1/2): libxcrypt-compat-4.4.33-7.amzn2023.x86_64.rpm
                                                                  1.4 MB/s | 92 kB
                                                                                        00:00
(2/2): python3-pip-21.3.1-2.amzn2023.0.13.noarch.rpm
                                                                   24 MB/s | 1.8 MB
                                                                                        00:00
Total
                                                                   18 MB/s | 1.9 MB
                                                                                        00:00
Running transaction check
Transaction check succeeded.
Running transaction test
Transaction test succeeded.
Running transaction
  Preparing
                                                                                             1/1
                                                                                             1/2
  Installing
                  : libxcrypt-compat-4.4.33-7.amzn2023.x86_64
```

env-dvc + git

```
● ■ master — ec2-user@ip-172-31-31-222:~ — ssh -i ./Downloads/dvc.keys.pem ec2-user@34.228.116.12 — 164×48

[((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$ git --version
git version 2.58.1
((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$ ■
```

Proyecto en DVC

```
o o master — ec2-user@ip-172-31-31-222:~/dvc-proj — ssh -i ./Downloads/dvc.keys.pem ec2-user@34.228.116.12 —...
((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$ pwd
/home/ec2-user
((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$ mkdir dvc-proj
((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$ cd dvc-proj/
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ git init
hint: Using 'master' as the name for the initial branch. This default branch name
hint: is subject to change. To configure the initial branch name to use in all
hint: of your new repositories, which will suppress this warning, call:
hint: git config --global init.defaultBranch <name>
hint: Names commonly chosen instead of 'master' are 'main', 'trunk' and
hint: 'development'. The just-created branch can be renamed via this command:
hint: git branch -m <name>
hint: Disable this message with "git config set advice.defaultBranchName false"
Initialized empty Git repository in /home/ec2-user/dvc-proj/.git/
((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ git branch -m main
((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ dvc init
Initialized DVC repository.
You can now commit the changes to git.
         DVC has enabled anonymous aggregate usage analytics.
      Read the analytics documentation (and how to opt-out) here:
               <https://dvc.org/doc/user-guide/analytics>
What's next?
- Check out the documentation: <a href="https://dvc.org/doc">https://dvc.org/doc</a>
- Get help and share ideas: <a href="https://dvc.org/chat">https://dvc.org/chat</a>
- Star us on GitHub: <a href="https://github.com/iterative/dvc">https://github.com/iterative/dvc</a>
((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ git status
On branch main
No commits vet
Changes to be committed:
  (use "git rm --cached <file>..." to unstage)
       new file: .dvc/.gitignore
new file: .dvc/config
        new file: .dvcignore
((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$
```

Dataset iris.csv

```
((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ git commit -m " Inicializacion de DVC "
[main (root-commit) 3074bc6] Inicializacion de DVC
Committer: EC2 Default User <ec2-user@ip-172-31-31-222.ec2.internal>
Your name and email address were configured automatically based
on your username and hostname. Please check that they are accurate.
You can suppress this message by setting them explicitly. Run the
following command and follow the instructions in your editor to edit
your configuration file:
   git config --global --edit
After doing this, you may fix the identity used for this commit with:
   git commit --amend --reset-author
3 files changed, 6 insertions(+)
 create mode 100644 .dvc/.gitignore
create mode 100644 .dvc/config
create mode 100644 .dvcignore
((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ mkdir -p data && curl https://raw.githubusercontent.com/mwaskom/seaborn-data/master/iris.csv -o data/data .csv
 % Total % Received % Xferd Average Speed Time Time Time Current
                             Dload Upload Total Spent Left Speed
100 3858 100 3858 0 0 189k
                                       0 --:--:- 198
curl: (6) Could not resolve host: .csv
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ ls -la
total 4
drwxr-xr-x. 5 ec2-user ec2-user 60 Aug 14 03:17
drwx-----. 8 ec2-user ec2-user 169 Aug 14 03:15
drwxr-xr-x. 3 ec2-user ec2-user 49 Aug 14 03:15 .dvc
-rw-r--r-. 1 ec2-user ec2-user 139 Aug 14 03:15 .dvcignore
drwxr-xr-x. 7 ec2-user ec2-user 150 Aug 14 03:16 .git
drwxr-xr-x. 2 ec2-user ec2-user 18 Aug 14 03:17 data
((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ cd data/
[((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ ls
[((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ cat data
sepal_length, sepal_width, petal_length, petal_width, species
5.1,3.5,1.4,0.2, setosa
4.9.3.0.1.4.0.2.setosa
4.7,3.2,1.3,0.2, setosa
4.6,3.1,1.5,0.2, setosa
5.0,3.6,1.4,0.2, setosa
5.4,3.9,1.7,0.4, setosa
4.6.3.4.1.4.0.3.setosa
5.0.3.4.1.5.0.2.setosa
4.4.2.9.1.4.0.2.setosa
4.9,3.1,1.5,0.1, setosa
5.4,3.7,1.5,0.2, setosa
4.8,3.4,1.6,0.2, setosa
4.8,3.0,1.4,0.1, setosa
4.3,3.0,1.1,0.1,setosa
```

dvc add data

```
💿 🔵 📄 master — ec2-user@ip-172-31-31-222:~/dvc-proj/data — ssh -i ./Downloads/dvc.keys.pem ec2-user@34.228.116....
[((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ pwd
/home/ec2-user/dvc-proj/data
[((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ dvc add data
100% Adding...|
                                                                                                                                  |1/1 [00:00, 24.74file/s]
To track the changes with git, run:
       git add .gitignore data.dvc
To enable auto staging, run:
       dvc config core.autostage true
[((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ ls -la
total 12
drwxr-xr-x. 2 ec2-user ec2-user 52 Aug 14 03:19
drwxr-xr-x. 5 ec2-user ec2-user 60 Aug 14 03:17 ...
-rw-r--r-. 1 ec2-user ec2-user 6 Aug 14 03:19 .gitignore
-rw-r--r-. 1 ec2-user ec2-user 3858 Aug 14 03:19 data
-rw-r--r-. 1 ec2-user ec2-user 84 Aug 14 03:19 data.dvc
[((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ cat .gitignore
/data
((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$
```

Push to remote repository

https://github.com/micardona96/DVC

```
master — ec2-user@ip-172-31-31-222:~/dvc-proj — ssh -i ./Downloads/dvc.keys.pem ec2-user@34.228.116.12 —...

[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ git push -u origin main

Enumerating objects: 10% (11/11), done.

Counting objects: 100% (11/11), done.

Compressing objects: 100% (8/8), done.

Writing objects: 100% (11/11), 869 bytes | 869.00 KiB/s, done.

Total 11 (delta 2), reused 0 (delta 0), pack-reused 0 (from 0)

remote: Resolving deltas: 100% (2/2), done.

To github.com:micardona96/DVC.git

* [new branch] main -> main

branch | main' > set up to track 'origin/main'.

((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$
```

Datos con fuente local

```
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ git push -u origin main
Enumerating objects: 11, done.
Counting objects: 100% (11/11), done.
Compressing objects: 100% (8/8), done.
Writing objects: 100% (11/11), 869 bytes | 869.00 KiB/s, done.
Total 11 (delta 2), reused 0 (delta 0), pack-reused 0 (from 0)
remote: Resolving deltas: 100% (2/2), done.
To github.com:micardona96/DVC.git
 * [new branch]
                     main -> main
branch 'main' set up to track 'origin/main'.
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ cd
[((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$ mkdir /tmp/dvcstore
[((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$ cd dvc-proj/
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ dvc remote add -d fuentelocal /tmp/dvcstore
Setting 'fuentelocal' as a default remote.
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ cat .dvc/config
[core]
    remote = fuentelocal
['remote "fuentelocal"']
    url = /tmp/dvcstore
((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$
```

Contenido de la carpeta /tmp/dvcstore/

```
[((env-dvc) ) [ec2-user@ip-172-31-31-222 /]$ cd tmp/dvcstore/
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvcstore]$ ls -la
total 0
drwxr-xr-x. 3 ec2-user ec2-user 60 Aug 14 03:28 .
drwxrwxrwt. 12 root root 360 Aug 14 03:29 .
drwxr-xr-x. 3 ec2-user ec2-user 60 Aug 14 03:28 files
((env-dvc) ) [ec2-user@ip-172-31-31-222 dvcstore]$ .
```

dvc pull

```
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ pwd
/home/ec2-user/dvc-proi
((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proi]$ rm -rf .dvc/cache/
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ ls
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ rm -f data
rm: cannot remove 'data': Is a directory
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ rm -f data/data
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ ls -la
total 4
drwxr-xr-x. 5 ec2-user ec2-user 60 Aug 14 03:17
drwx-----. 8 ec2-user ec2-user 169 Aug 14 03:15 ...
drwxr-xr-x. 3 ec2-user ec2-user 49 Aug 14 03:31 .dvc
-rw-r--r-. 1 ec2-user ec2-user 139 Aug 14 03:15 .dvcignore
drwxr-xr-x. 7 ec2-user ec2-user 150 Aug 14 03:26 .git
drwxr-xr-x. 2 ec2-user ec2-user 40 Aug 14 03:32 data
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ cd data/
[((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ ls -la
total 8
drwxr-xr-x. 2 ec2-user ec2-user 40 Aug 14 03:32
drwxr-xr-x. 5 ec2-user ec2-user 60 Aug 14 03:17
-rw-r--r-. 1 ec2-user ec2-user 6 Aug 14 03:19 .gitignore
-rw-r--r-. 1 ec2-user ec2-user 84 Aug 14 03:19 data.dvc
((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ dvc pull
Collecting
Fetching
Building workspace index
Comparing indexes
Applying changes
       data
1 file added and 1 file fetched
total 12
drwxr-xr-x. 2 ec2-user ec2-user 52 Aug 14 03:32
drwxr-xr-x. 5 ec2-user ec2-user 60 Aug 14 03:17
-rw-r--r-. 1 ec2-user ec2-user 6 Aug 14 03:19 .gitignore
-rw-r--r-. 1 ec2-user ec2-user 3858 Aug 14 03:32 data
-rw-r--r-. 1 ec2-user ec2-user 84 Aug 14 03:19 data.dvc
((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$
```

última versión de OpenSSL

```
[((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ pip install pyopenssl --upgrade
Collecting pyopenssl
Using cached pyopenssl-25.1.0-py3-none-any.whl.metadata (17 kB)
Requirement already satisfied: cryptography<46,>=41.0.5 in /home/ec2-user/env-dvc/lib64/python3.12/site-packages (from pyopenssl) (45.0.6)
Requirement already satisfied: typing-extensions>=4.9 in /home/ec2-user/env-dvc/lib64/python3.12/site-packages (from pyopenssl) (4.14.1)
Requirement already satisfied: cffi>=1.14 in /home/ec2-user/env-dvc/lib64/python3.12/site-packages (from cryptography<46,>=41.0.5->pyopenssl) (1.17.1)
Requirement already satisfied: pycparser in /home/ec2-user/env-dvc/lib64/python3.12/site-packages (from cffi>=1.14->cryptography<46,>=41.0.5->pyopenssl) (2.2
Downloading pyopenssl-25.1.0-py3-none-any.whl (56 kB)
Installing collected packages: pyopenssl
Successfully installed pyopenssl-25.1.0
((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$
```

Descarga de aws_cli

```
master — ec2-user@ip-172-31-31-222:~ — ssh
total 60536
drwx----. 8 ec2-user ec2-user
                                 189 Aug 14 03:36
drwxr-xr-x. 3 root
                     root
                                  22 Aug 14 02:51
-rw-----. 1 ec2-user ec2-user
                                 140 Aug 14 03:04 .bash_history
-rw-r--r--. 1 ec2-user ec2-user
                                  18 Jan 28 2023 .bash_logout
                                 141 Jan 28 2023 .bash profile
-rw-r--r--. 1 ec2-user ec2-user
                                 492 Jan 28 2023 .bashrc
-rw-r--r-. 1 ec2-user ec2-user
                                  17 Aug 14 03:06 .cache
drwxr-xr-x. 3 ec2-user ec2-user
drwxr-xr-x. 3 ec2-user ec2-user
                                  23 Aug 14 03:15 .config
drwx---- 4 ec2-user ec2-user
                                  28 Aug 14 03:06 .local
drwx---- 2 ec2-user ec2-user
                                 111 Aug 14 03:26 .ssh
-rw-r--r-. 1 ec2-user ec2-user 61970604 Aug 14 03:36 awscliv2.zip
drwxr-xr-x. 5 ec2-user ec2-user
                                  60 Aug 14 03:17 dvc-proj
drwxr-xr-x. 5 ec2-user ec2-user
                                  74 Aug 14 03:10 env-dvc
((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$ ■
```

aws cli

```
[((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$ ls

aws awscliv2.zip dvc-proj env-dvc

[((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$ sudo aws/install

You can now run: /usr/local/bin/aws --version

[((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$ aws --version

aws-cli/2.28.9 Python/3.13.4 Linux/6.1.147-172.266.amzn2023.x86_64 exe/x86_64.amzn.2023

((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$ ■
```

AWS Credentials

```
[((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$ cat aws/credentials
[default]
aws_access_key_id=ASSIAXMLC2PTTXBBH6NIJ
aws_access_key_id=ASSIAXMLC2PTTXBBH6NIJ
aws_secret_access_key=+Qn+Rck2vWR1561hQgp007QkC71nVRIe9Lc7ag/o
aws_session_token=IQoJb3JpZ2luX2vjEPP///////wEaCXvzLXdlc3QtMiJHMEUCIQCoduwjeuvXWnjVJFEFjaCS8Y0hrNWTprBauwn7RnVWtgIgbaszXsp14sGnRPsTjZI3v9RSkqh/1jRytyzFyM5Owgoqtw
IIPBAAGgw1MDc1NTAyNjg2NDciDJtBGhzyGWTgBVGMBCQUAUR+guTnXy222PQEMgAQZWTJEGxwwWpUNLheOhUq/7XU6AzNicZKDRmY6UMRyzZL7u64gsM7wxf234ZMHV0x9S6cybfudsVUvhrG71UGgNjo8ZmARpZFtMD
CHbawifLneDkwE0g6g+RzQhW0hNtEi121HhMOCTPbfurC3k1j379XgExSiwMxL5BfgBfCSgQMi0kKCVckif7OuJ7XhU0yHcD2AHY2FShkcNXTMgyzt3tCJ6IaTRiEifU4OXeJfFnOJHW4LbUeLfUz6PFddsp7m1G7AmT2
EMQQB921/xubdzaitIT4x9N0w9n1n0DV7Rdc6souU92uWfVnauJSMMGgmSe2W3LfBB2HAtWi405Fqum43Kr0JGxDD0nPXEBjqdAc8ODFhnmWt82maOecjia5//3suA7gvc3F3cBtxqcTkpBpMLFdXLqfh2UFd1sf8TOh
X9hQ3LkWy2xHqUq8B8W3E7i1ItCKkI49+TRIJJyJ1fZ1xLp1sX+JCjgmCtJ4z10eJJj8dclTCfvDQHVXzktBgYm4mjhZYlsqCUfHkRzxxq0jTlsQ0YsrQfPxdJ6JeyB0GWPHyWwNI9aaBNOSEM=
((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$
```

aws iam list - users

```
[((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$ aws iam list-users
{
    "Users": []
}
((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$
```

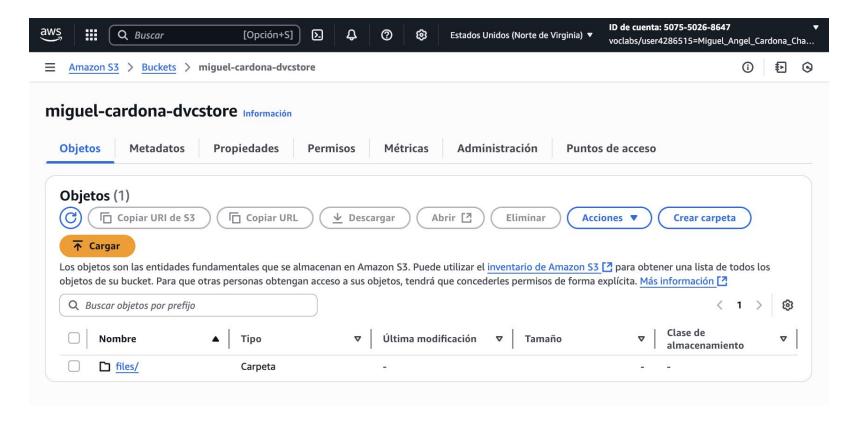
Bucket S3 + dvc

```
[((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$ aws s3 mb s3://miguel-cardona-dvcstore
make_bucket: miguel-cardona-dvcstore
[((env-dvc) ) [ec2-user@ip-172-31-31-222 ~]$ cd dvc-proj/
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ dvc remote add -d aws-remote s3://miguel-cardona-dvcstore
Setting 'aws-remote' as a default remote.
  ((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ ...
```

Contenido del archivo .dvc/config

```
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ cat .dvc/config
[core]
    remote = aws-remote
['remote "fuentelocal"']
    url = /tmp/dvcstore
['remote "aws-remote"']
    url = s3://miguel-cardona-dvcstore
((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$
```

s3 con los files



git commit nuevo dvc/config

```
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ dvc push
Collecting
Pushing
1 file pushed
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ git commit .dvc/config -m "Adicionar S3 como remote"
[main 898afe1] Adicionar S3 como remote
 Committer: EC2 Default User <ec2-user@ip-172-31-31-222.ec2.internal>
Your name and email address were configured automatically based
on your username and hostname. Please check that they are accurate.
You can suppress this message by setting them explicitly. Run the
following command and follow the instructions in your editor to edit
your configuration file:
    git config --global --edit
After doing this, you may fix the identity used for this commit with:
    git commit --amend --reset-author
 1 file changed, 6 insertions(+)
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ git push
Enumerating objects: 7, done.
Counting objects: 100% (7/7), done.
Compressing objects: 100% (4/4), done.
Writing objects: 100% (4/4), 500 bytes | 500.00 KiB/s, done.
Total 4 (delta 0), reused 0 (delta 0), pack-reused 0 (from 0)
To github.com:micardona96/DVC.git
   f75602b..898afe1 main -> main
((env-dvc)) [ec2-user@ip-172-31-31-222 dvc-proj]$
```

Agregando datos duplicados al dvc

```
[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ dvc add data/data

100% Adding...|

To track the changes with git, run:

git add data/data.dvc

To enable auto staging, run:

dvc config core.autostage true

[((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$ dvc push

Collecting

Pushing

1 file pushed

((env-dvc) ) [ec2-user@ip-172-31-31-222 dvc-proj]$
```

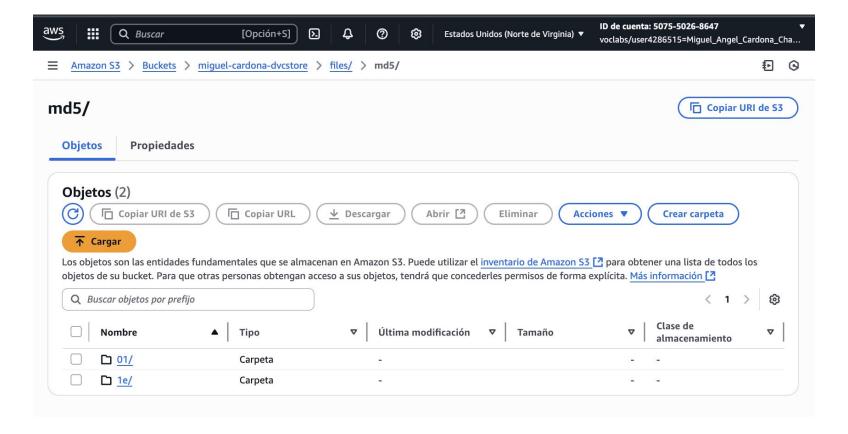
Actualizando git remoto

```
[((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ git add .
[((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ git commit -m " Dataset actualizado"
[main 657f9e4] Dataset actualizado
 Committer: EC2 Default User <ec2-user@ip-172-31-31-222.ec2.internal>
Your name and email address were configured automatically based
on your username and hostname. Please check that they are accurate.
You can suppress this message by setting them explicitly. Run the
following command and follow the instructions in your editor to edit
your configuration file:
    git config --global --edit
After doing this, you may fix the identity used for this commit with:
    git commit --amend --reset-author
 1 file changed, 2 insertions(+), 2 deletions(-)
Enumerating objects: 7, done.
Counting objects: 100% (7/7), done.
Compressing objects: 100% (4/4), done.
Writing objects: 100% (4/4), 436 bytes | 436.00 KiB/s, done.
Total 4 (delta 1), reused 0 (delta 0), pack-reused 0 (from 0)
remote: Resolving deltas: 100% (1/1), completed with 1 local object.
To github.com:micardona96/DVC.git
   898afe1..657f9e4 main -> main
((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$
```

data.dvc md5 local

```
[((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ cat data.dvc outs:
- md5: 1e0306f8f8af775862f5e5dafab5bc6f
    size: 7716
    hash: md5
    path: data
((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ ■
```

S3 objetos en md5



Clonado de Repo sin data

```
[((env-dvc) ) [ec2-user@ip-172-31-31-222 clone_git]$ pwd
/home/ec2-user/clone_git
[((env-dvc) ) [ec2-user@ip-172-31-31-222 clone_git]$ git clone git@github.com:micardona96/DVC.git
Cloning into 'DVC'...
remote: Enumerating objects: 19, done.
remote: Counting objects: 100% (19/19), done.
remote: Compressing objects: 100% (13/13), done.
remote: Total 19 (delta 3), reused 19 (delta 3), pack-reused 0 (from 0)
Receiving objects: 100% (19/19), done.
Resolving deltas: 100% (3/3), done.
[((env-dvc) ) [ec2-user@ip-172-31-31-222 clone git]$ ls
[((env-dvc) )        [ec2-user@ip-172-31-31-222        clone_git]$ cd DVC/
[((env-dvc)) [ec2-user@ip-172-31-31-222 DVC]$ ls
data
[((env-dvc) ) [ec2-user@ip-172-31-31-222 DVC]$ cd data/
((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$
```

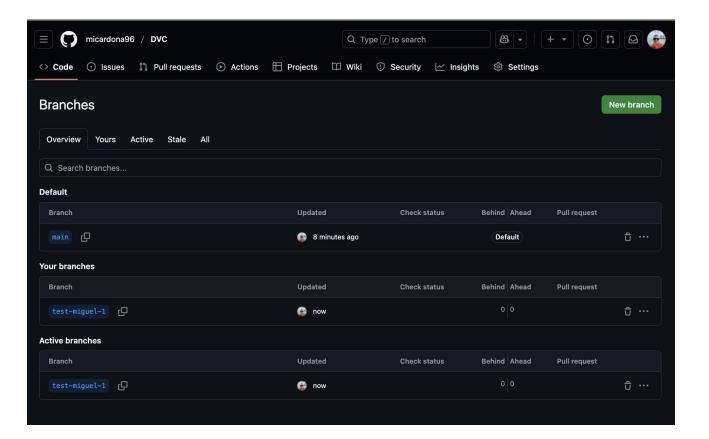
dvc pull en nueva carpeta

```
[((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ dvc pull
Collecting
Fetching
Building workspace index
Comparing indexes
Applying changes
A data
1 file added and 1 file fetched
[((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ ls
data data.dvc
((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ ls
```

Manejo de ramas en git

```
[((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ git checkout -b test-miguel-1
Switched to a new branch 'test-miguel-1'
  ((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ git checkout test-miguel-1
Already on 'test-miguel-1'
[((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$ dvc checkout
Building workspace index
Comparing indexes
Applying changes
  ((env-dvc) ) [ec2-user@ip-172-31-31-222 data]$
```

Ramas en Github



Repositorio con la evidencia + s3

https://github.com/micardona96/DVC/branches

s3://miguel-cardona-dvcstore/files/md5/