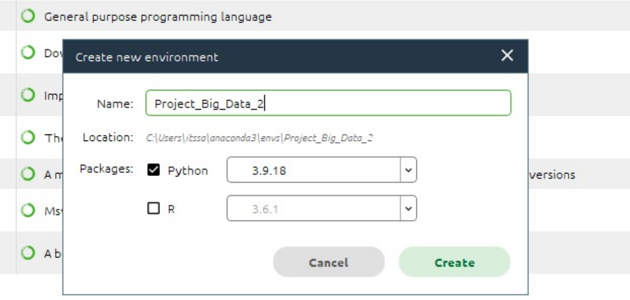
Github documentation

Made by: Pedro Torrijos, Ignacio Donderis and Miguel Santiago

Index

- We **prepare anaconda** for the entire process.



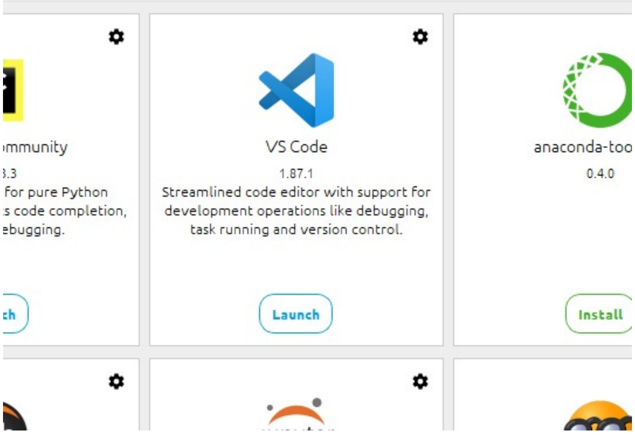
Create new environment

Name: Project_Big_Data_2

Location: C:\Users\itssa\anaconda3\envs\Project_Big_Data_2

Packages: ☑ Python 3.9.18

☐ R 3.6.1

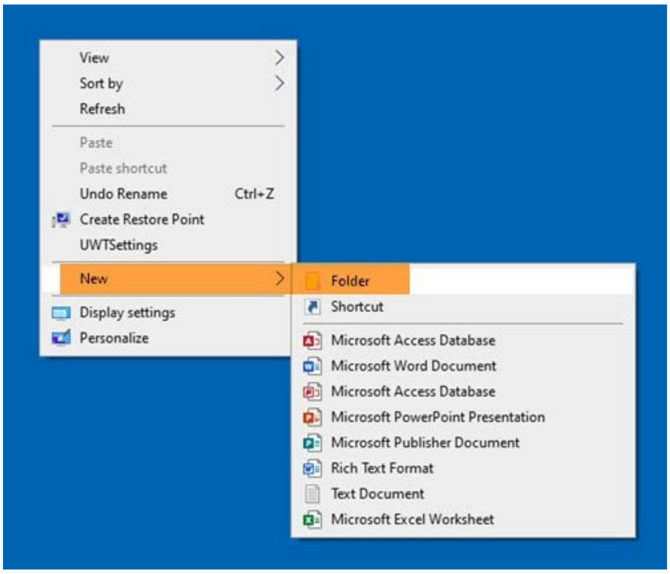


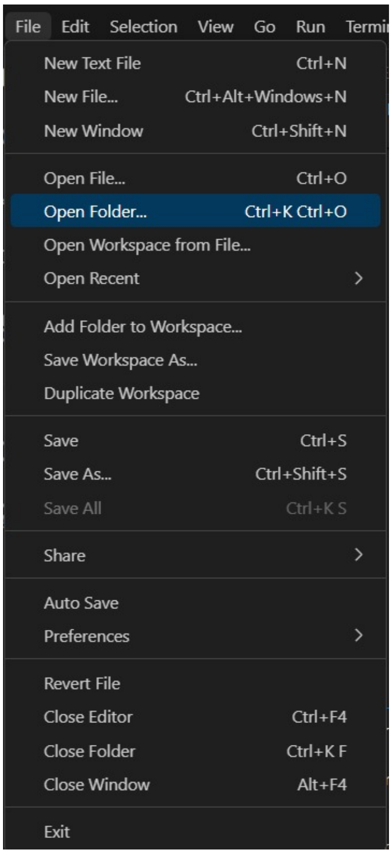
VS Code

1.87.1

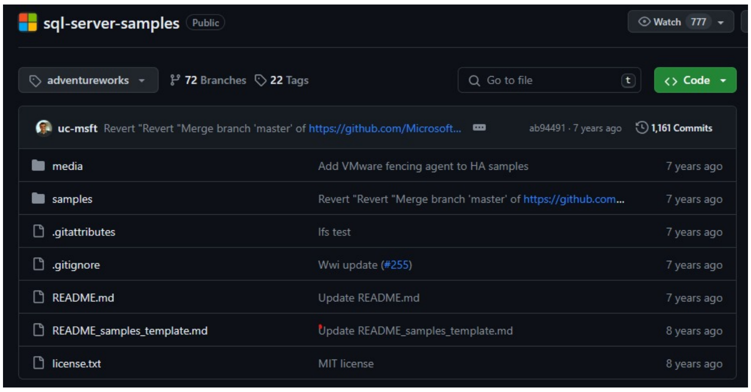
Streamlined code editor with support for development operations like debugging, task running and version control.

- We **create the folder** where we are going to host the git repository locally and **open with VS code.**



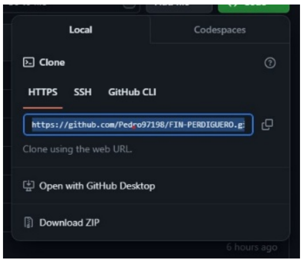


- First of all you cloned a repository with the **git clone** command.

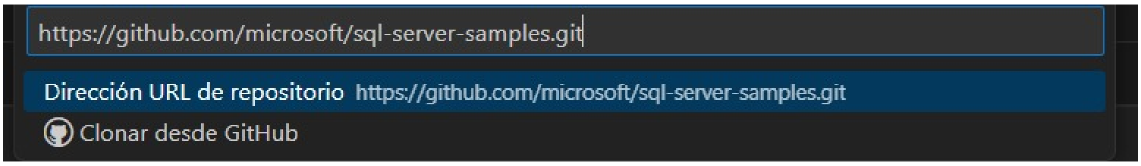


uc-msft Revert "Revert "Merge branch 'master' of https://github.com/Microsoft... ab94491 · 7 years ago 1,161 Commits

media	Add VMware fencing agent to HA samples	7 years ago
samples	Revert "Revert "Merge branch 'master' of https://github.com...	7 years ago
.gitattributes	lfs test	7 years ago
.gitignore	Wwi update (#255)	7 years ago
README.md	Update README.md	7 years ago
README_samples_template.md	Update README_samples_template.md	8 years ago
license.txt	MIT license	8 years ago



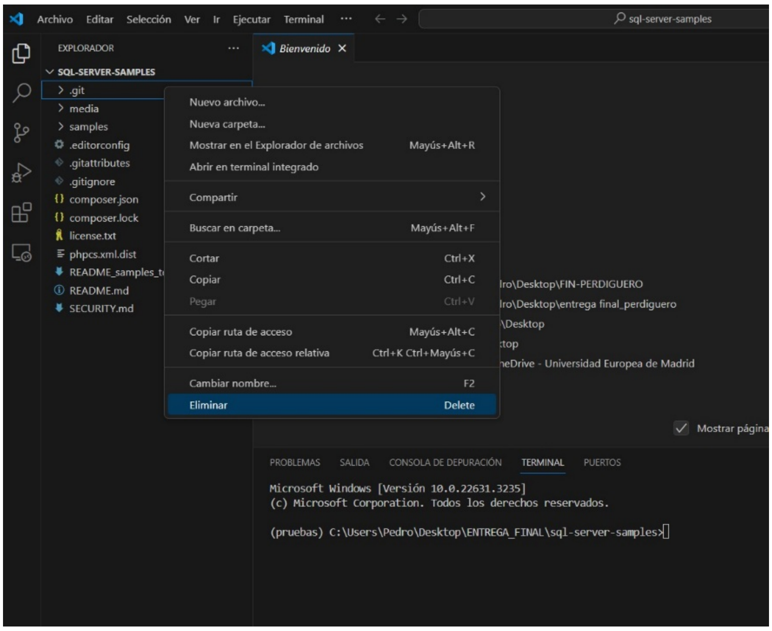
```
https://github.com/Pedro97198/FIN-PERDIGUERO.gi
```



Dirección URL de repositorio https://github.com/microsoft/sql-server-samples.git
Clonar desde GitHub

```
https://github.com/microsoft/sql-server-samples.git
```

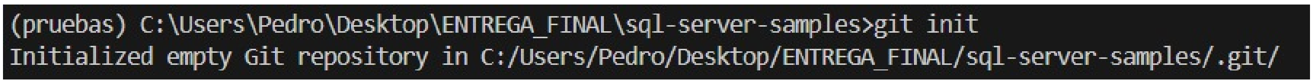
- We **close the remote connection** with the git-hub repository.



```
Microsoft Windows [Versión 10.0.22631.3235]
(c) Microsoft Corporation. Todos los derechos reservados.

(pruebas) C:\Users\Pedro\Desktop\ENTREGA_FINAL\sql-server-samples>
```

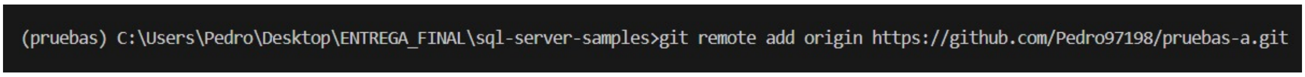
- We execute the **git init command**.



```
(pruebas) C:\Users\Pedro\Desktop\ENTREGA_FINAL\sql-server-samples>git init
Initialized empty Git repository in C:/Users/Pedro/Desktop/ENTREGA_FINAL/sql-server-samples/.git/
```

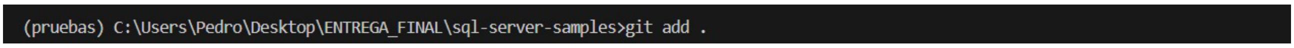
- We **created a github repository**.

- We execute the command **git remote add repository url**.



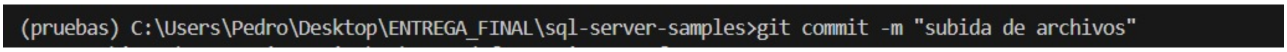
```
(pruebas) C:\Users\Pedro\Desktop\ENTREGA_FINAL\sql-server-samples>git remote add origin https://github.com/Pedro97198/pruebas-a.git
```

- We do a **git add**.

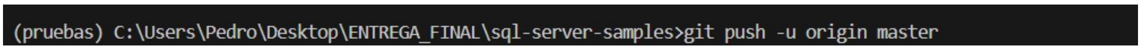


```
(pruebas) C:\Users\Pedro\Desktop\ENTREGA_FINAL\sql-server-samples>git add .
```

- We do the **commit and finally the push**.



```
(pruebas) C:\Users\Pedro\Desktop\ENTREGA_FINAL\sql-server-samples>git commit -m "subida de archivos"
```



```
(pruebas) C:\Users\Pedro\Desktop\ENTREGA_FINAL\sql-server-samples>git push -u origin master
```

- **Clean** of the data: We proceed to eliminate the items and folders that we are not going to use, and understand the rest of the data.

- We've recently set up an SQL server to manage our data, and as part of this process, we've incorporated the AdventureWorks database. This involved placing the AdventureWorks .bak file inside our local SQL server for easy access and management.

With the SQL server up and running, we seamlessly connected it to Power BI, leveraging its powerful data visualization capabilities. Using Power BI, we imported all the data from AdventureWorks, allowing us to delve deep into its insights and metrics.

After importing the data, we took it a step further by creating an entity-relationship diagram. This diagram visually represents the relationships between different entities (such as tables) in the AdventureWorks database. It's a crucial step in understanding the data model and how various pieces of information are connected.