Git and GitHub

Nicolás García Peñaloza

Incorporación en R y Python (Coming soon to Stata).

10 de marzo, 2025

Resumen

El desarrollo de proyectos en análisis de datos exige no solo el dominio de herramientas de vanguardia, sino también la aplicación rigurosa de buenas prácticas en medición. La calidad y el profesionalismo en este campo emergen de una sinergia entre el poder computacional, la organización meticulosa y una sólida gobernanza de datos. Así, el análisis de los datos deja de ser solo técnica y se convierte en una forma de pensamiento estratégico.

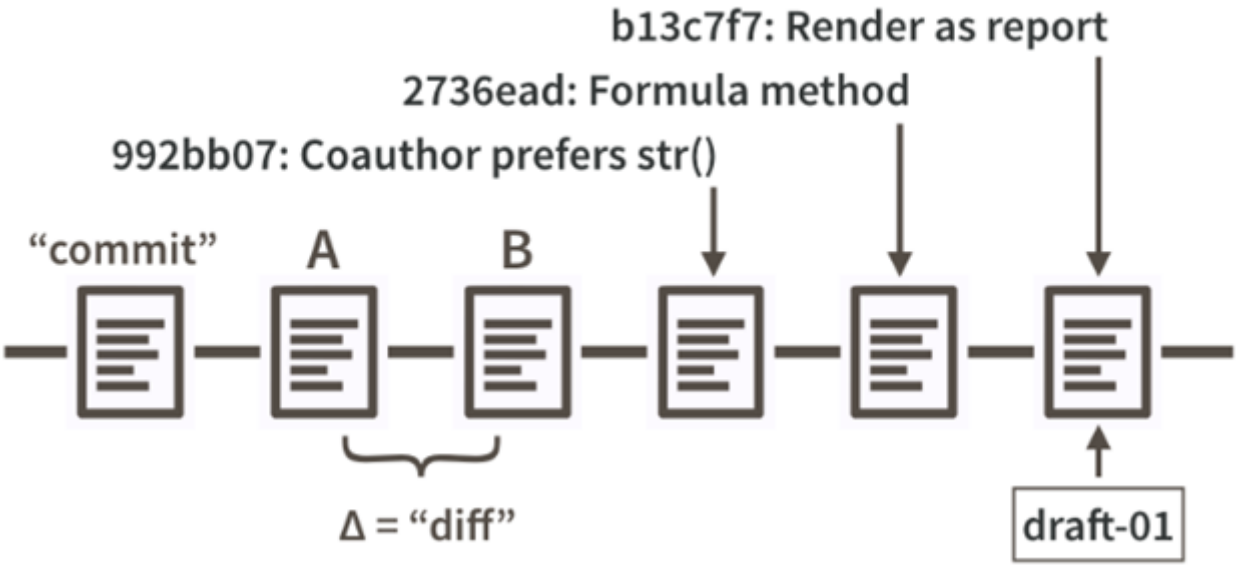
Índice

1. Git and GitHub

El desarrollo de proyectos que implican la programación de aplicaciones, software o análisis de datos requiere un control riguroso de las versiones generadas. Esto garantiza una adecuada estructuración y documentación del código, además de facilitar el trabajo colaborativo. Para ello, es fundamental incorporar un sistema de control de versiones estándar y moderno. Git, un sistema de control de versiones distribuido, estándar en el mundo del desarrollo, es una excelente opción que permite gestionar de manera eficiente la evolución de un conjunto de archivos, llamado repositorio, de forma ordenada y coherente.

Git es un proyecto de código abierto robusto, mantenido de manera activa, desarrollado por Linus Torvalds, el creador del kernel de Linux. Su principal función es almacenar el historial de versiones de un proyecto, lo que facilita tanto el seguimiento de cambios como la colaboración entre equipos.

Para comprender el funcionamiento de Git, es esencial familiarizarse con tres conceptos clave: repositorios, commits y diffs. Los repositorios son los contenedores donde se almacena toda la información del proyecto. Los commits son puntos específicos del historial que documentan los cambios realizados, junto con una descripción clara de los mismos. Por último, las diffs muestran las diferencias entre versiones, destacando los cambios realizados de una versión a otra.



Ahora, exploremos GitHub, una plataforma que actúa tanto como red social para desarrolladores como un servicio de gestión de proyectos en la nube. GitHub permite almacenar y gestionar proyectos de manera remota, con la opción de hacerlos públicos o privados según lo decida el contribuyente. Es comparable a servicios de almacenamiento como Dropbox, pero específicamente diseñado para el desarrollo y la colaboración en proyectos de programación, con funcionalidades mucho más avanzadas.

GitHub utiliza el sistema de control de versiones de Git, lo que significa que puedes trasladar tu proyecto desde un entorno local a la nube sin perder la capacidad de gestionar versiones, colaborar con otros y realizar un seguimiento detallado de los cambios.

Este servicio ofrece una versión gratuita que incluye almacenamiento de hasta 1 GB para repositorios, lo que lo convierte en una excelente opción para proyectos pequeños y medianos. ¿Ahora Git y GitHub son lo mismo? No. Aunque están estrechamente relacionados, no son lo mismo. Git es la herramienta que gestiona versiones localmente, mientras que GitHub es una plataforma que extiende sus funcionalidades al entorno colaborativo en la nube. En conjunto, forman una poderosa combinación para el desarrollo moderno.

Git	GitHub
Es un software.	Es un servicio.
Es un sistema de control de versiones para gestionar el historial del código fuente.	Es un servicio de alojamiento para repositorios Git.
Es mantenido por linux.	Es mantenido por Microsoft.
Se instala localmente en el sistema.	Está alojado en la web.
No tiene ninguna función de gestión de usuarios.	Incorpora una función de gestión de usuarios.
Open-source.	Incluye un nivel gratuito y un nivel de pago por uso.
Tiene una configuración mínima de herramientas externas.	Tiene un mercado activo para la integración de herramientas.

Un vez tengamos instalados los programas llámese R (1),Python (1,2,3,4,5) y sus respectivas IDE con las que se sienta a gusto para trabajar, debemos instalar Git y crear una cuenta en GitHub En el terminal del sistema operativo podemos actualizar Git con (git update) para ver la versión puede escribir (git - -version)

*Recuerde:***Inicialmente instalamos Git y debemos tener una cuenta en GitHub**

ver estos videos para entender de fondo GIT: Link, Link 2, Link 3

1. Abrimos en el equipo el simbolo del sistema (cmd y/o terminal) o buscamos en el equipo *git bash*

o *git CMD* una vez se abré alguno de los dos escribimos *git – –version* para conocer la versión con la que contamos, luego introducirnos a Git:

git config - -global user.name “NicolasGP01”

git config - -global user.email “nicolasgp0109@gmail.com”

git config - -global - -list

git config - -global user.name

git config - -global user.email

**Utilizamos generalmente las credenciales del nombre y correo que tenemos en GitHub*

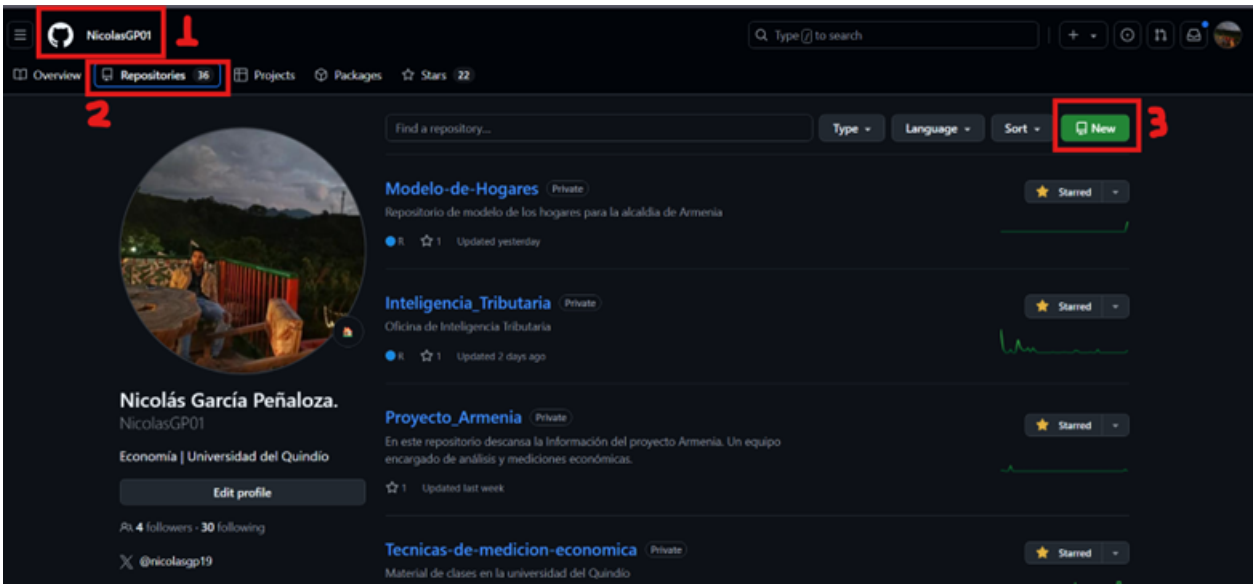
2. R en Rstudio

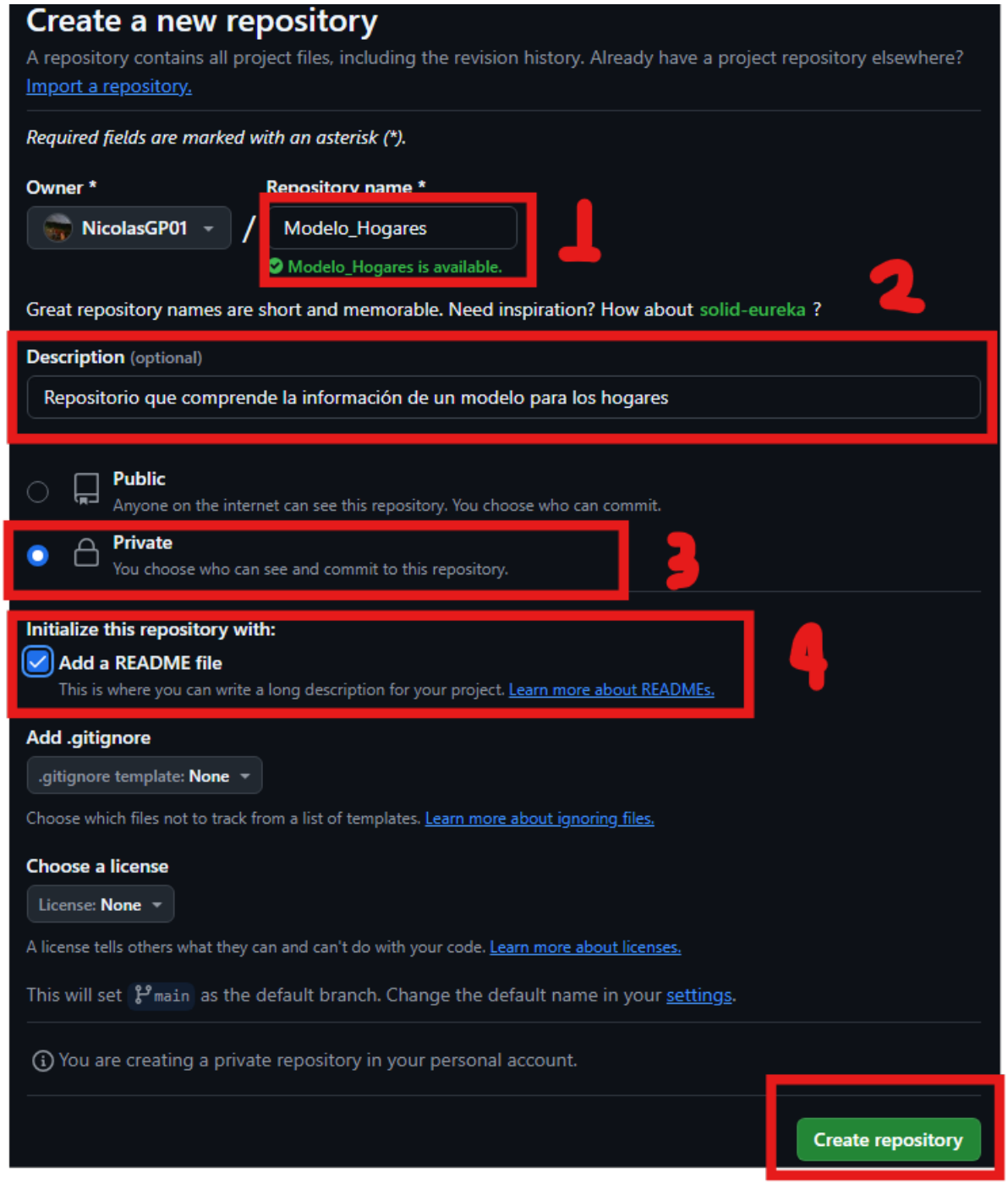
Cómo lo integramos con R en Rstudio: En este caso debemos contar con R y Rstudio instalados en nuestros equipos. URL

Otra alternativa para establer nuestros alias en GIT es desde Rstudio con el paquete “usethis” que puede establecer su nombre de usuario Git y correo electrónico desde dentro de R:

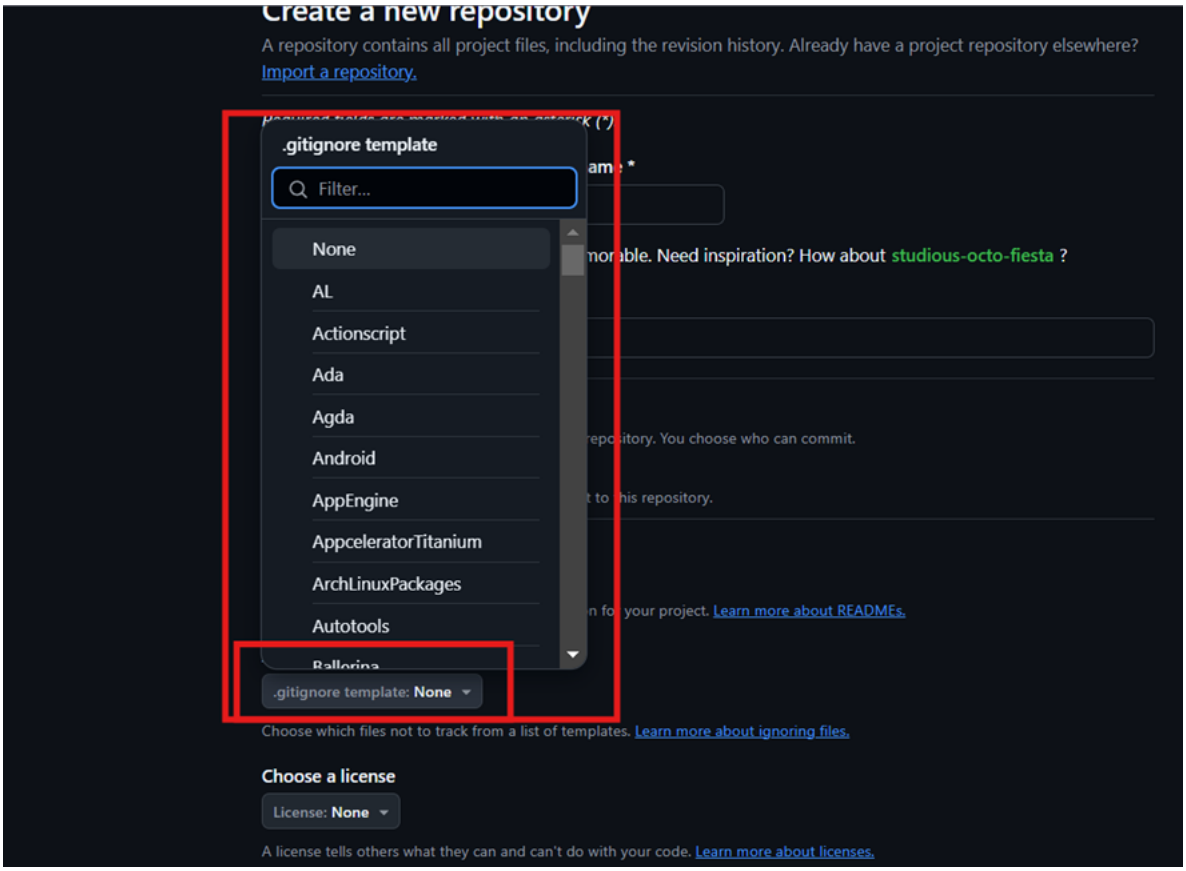
```
library(usethis)
usethis::use_git_config(user.name = "NicolasGP01" , user.email = "
    nicolasgp0109@gmail.com")
```

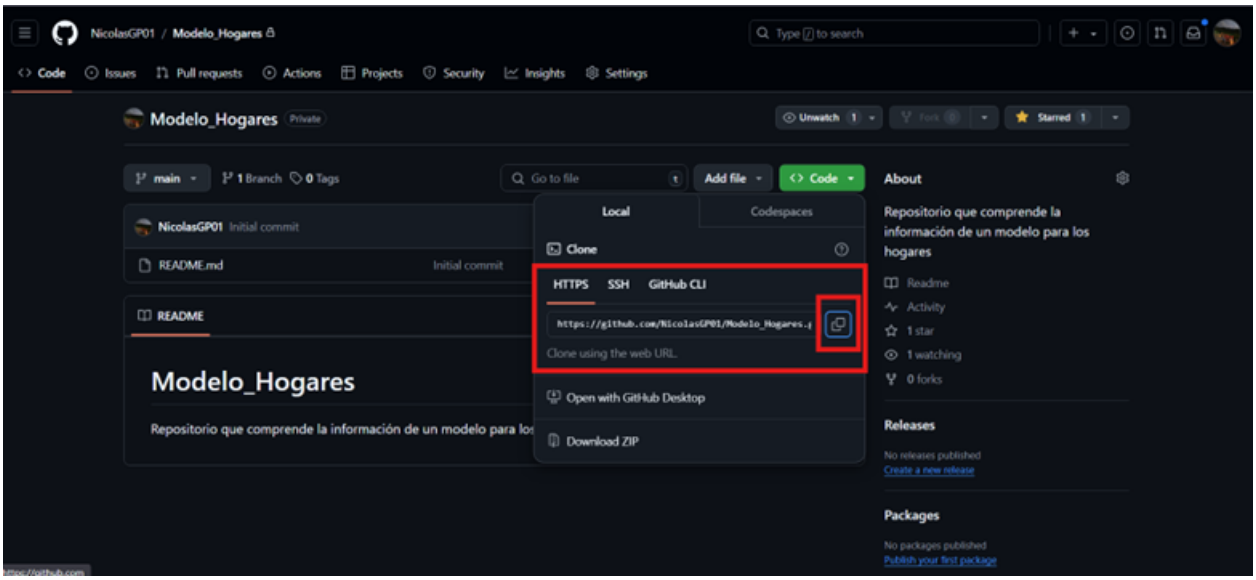
Creando la copia del repositorio GITHUB

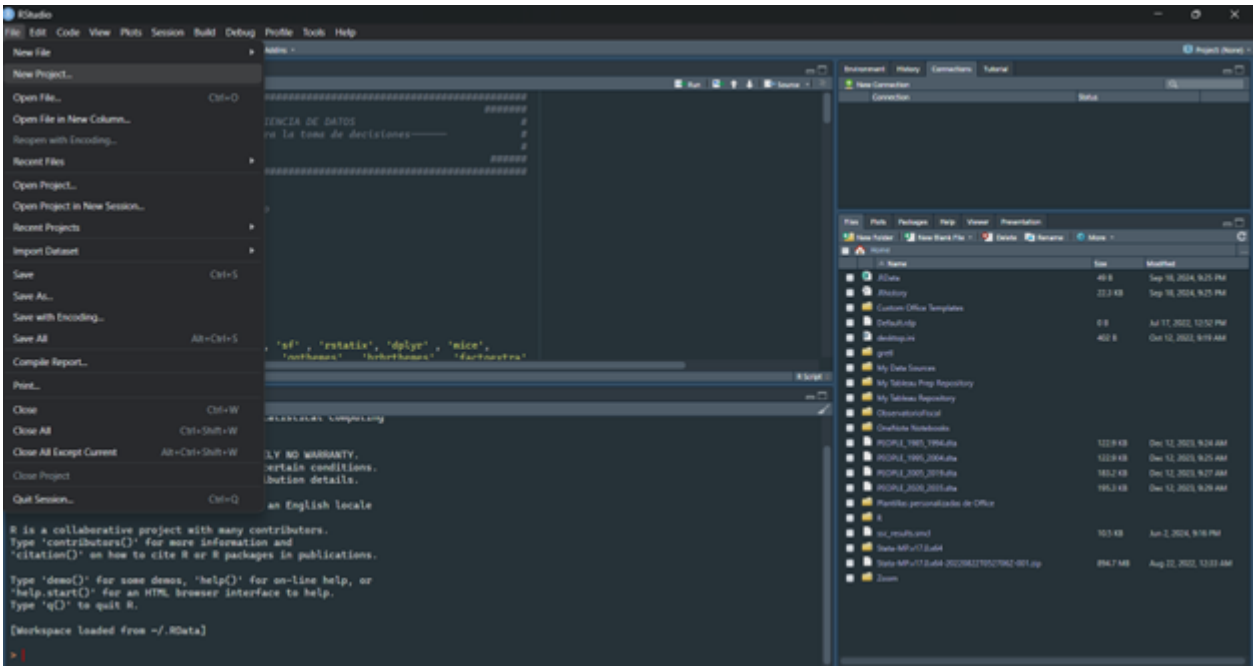


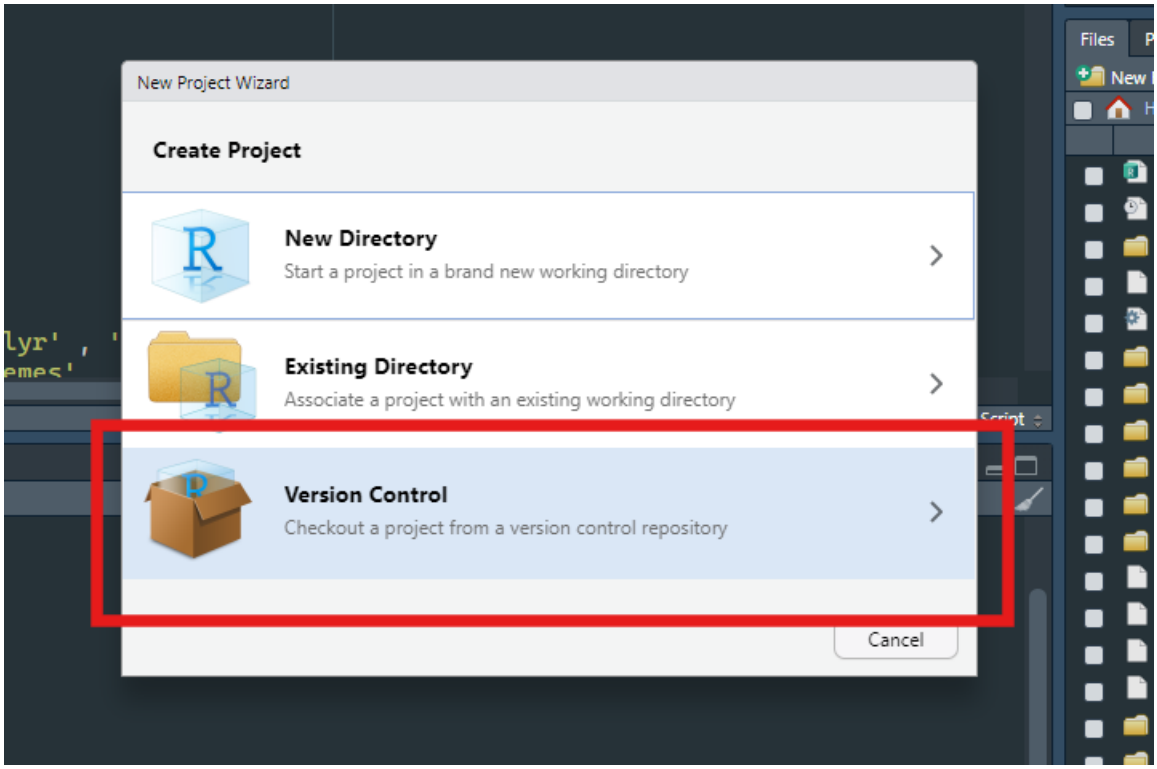


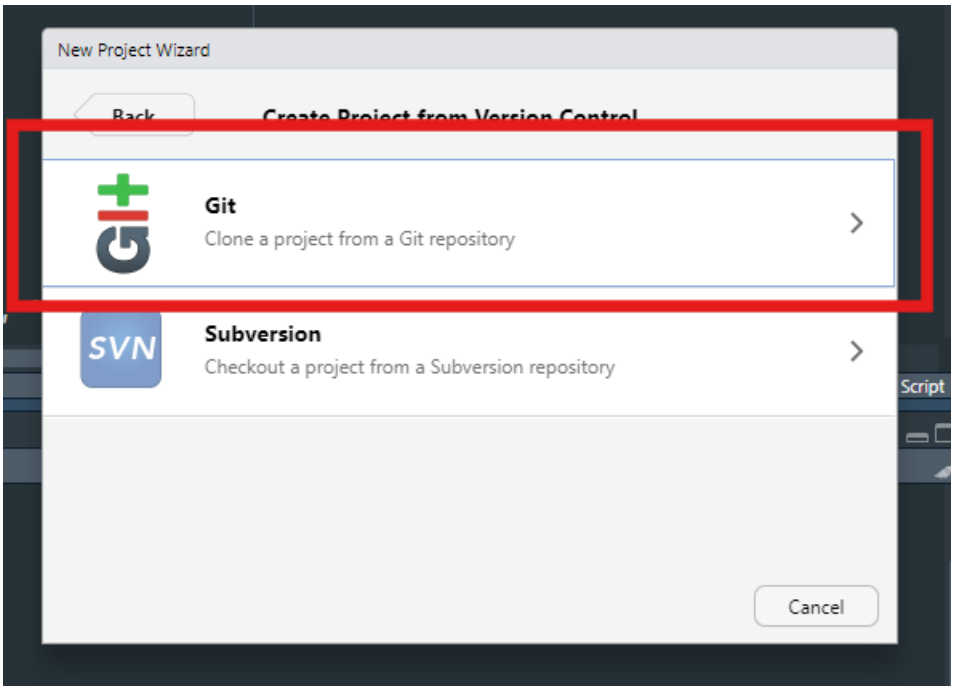
Elijemos el lenguaje que estemos utilizando

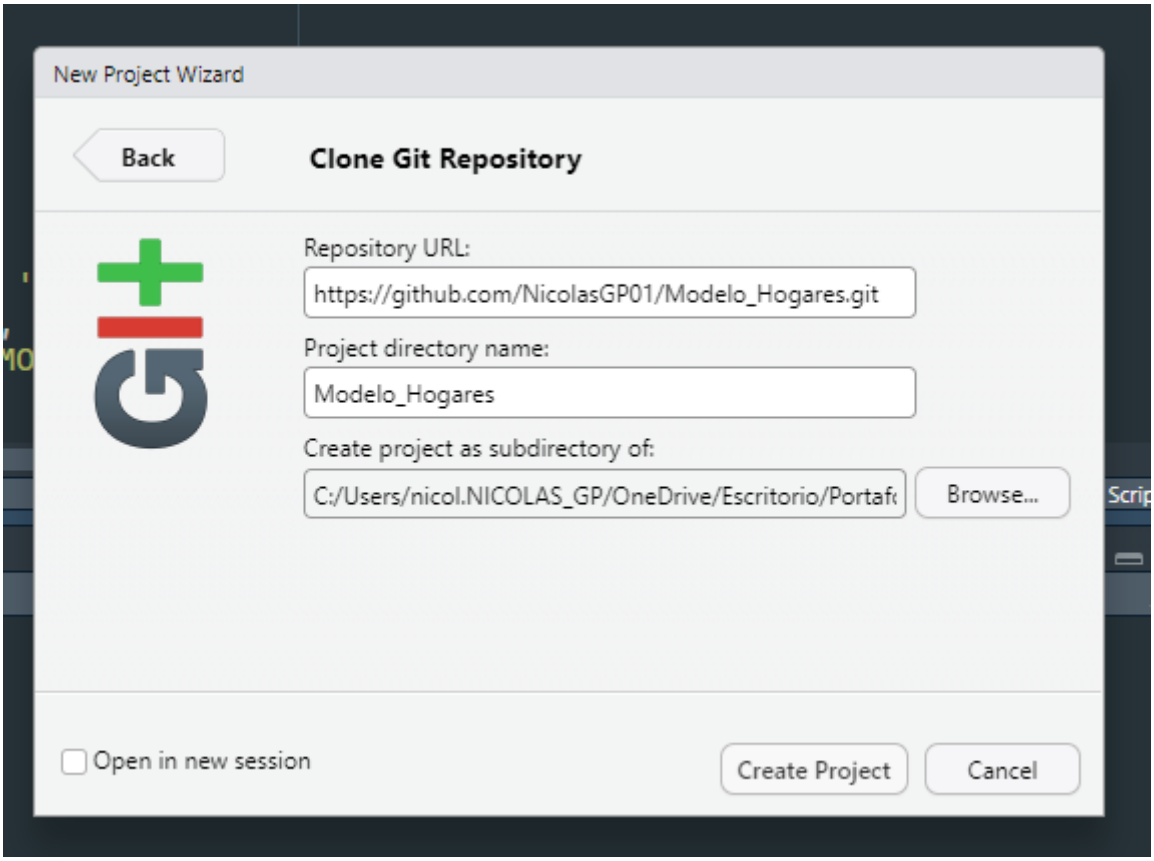

Create a new repository
A repository contains all project files, including the revision history. Already have a project repository elsewhere?
Import a repository.
.gitignore template
Filter...
None
AL
Actionscript
Ada
Agda
Android
AppEngine
AppceleratorTitanium
ArchLinuxPackages
Autotools
.gitignore template: None
Choose which files not to track from a list of templates. Learn more about ignoring files.
Choose a license
License: None
A license tells others what they can and can't do with your code. Learn more about licenses.


NicolasGP01 / Modelo_Hogares
Code
Issues
Pull requests
Actions
Projects
Security
Insights
Settings
Modelo_Hogares
Private
Unwatch 1
Starred 1
main
1 Branch
0 Tags
Go to file
Add file
Code
About
Repositorio que comprende la información de un modelo para los hogares
Local
Codespaces
Clone
HTTPS
SSH
GitHub CLI
Clone using the web URL.
Open with GitHub Desktop
Download ZIP
NicolasGP01 Initial commit
README.md
README
Modelo_Hogares
Repositorio que comprende la información de un modelo para los
Readme
Activity
1 star
1 watching
0 forks
Releases
No releases published
Create a new release
Packages
No packages published
Publish your first package

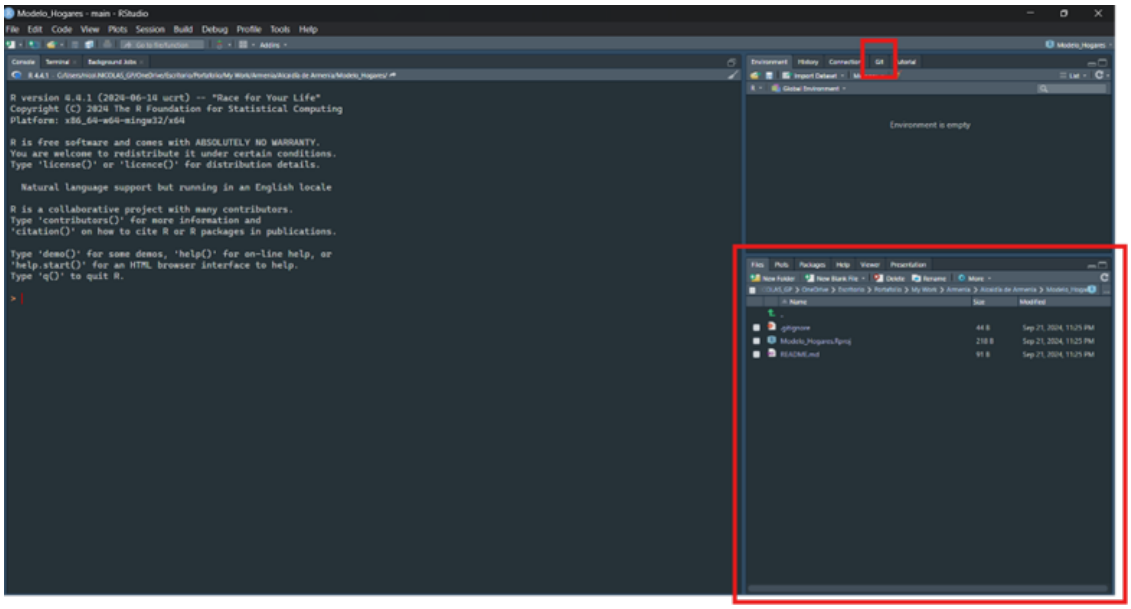


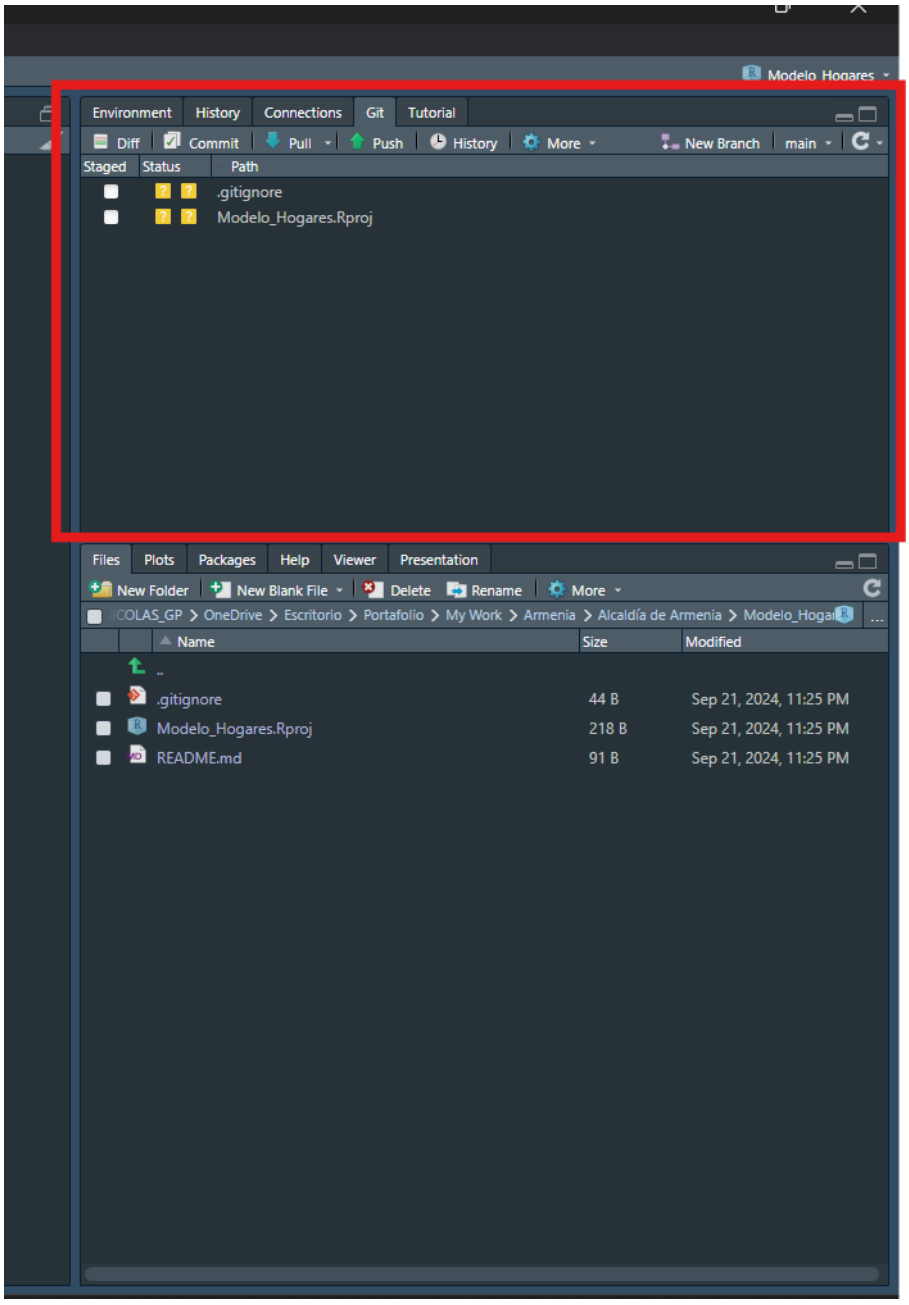

New Project Wizard
Create Project
New Directory
Start a project in a brand new working directory
Existing Directory
Associate a project with an existing working directory
Version Control
Checkout a project from a version control repository
Cancel
Files

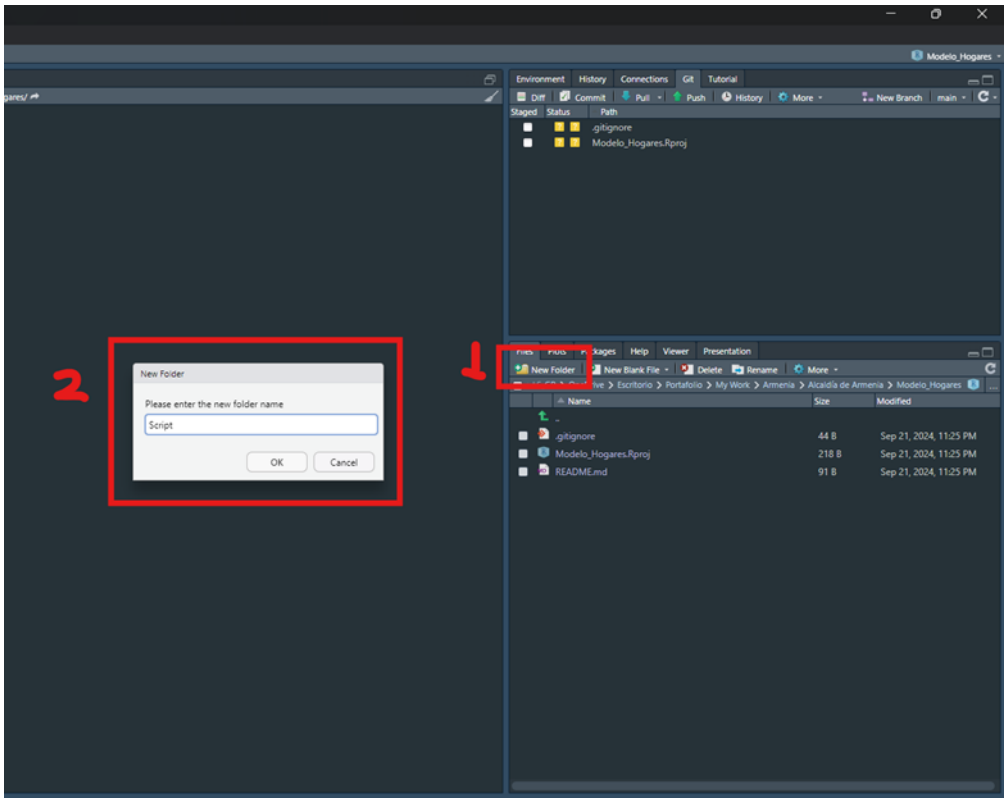


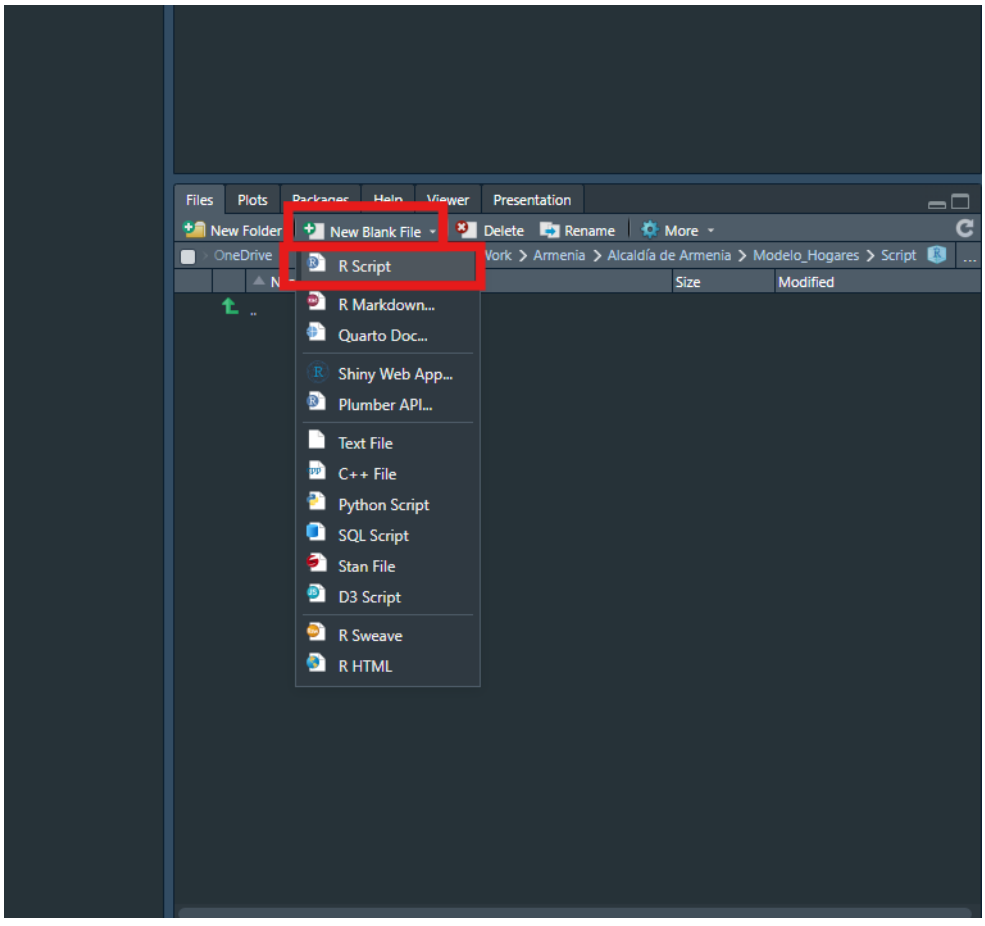


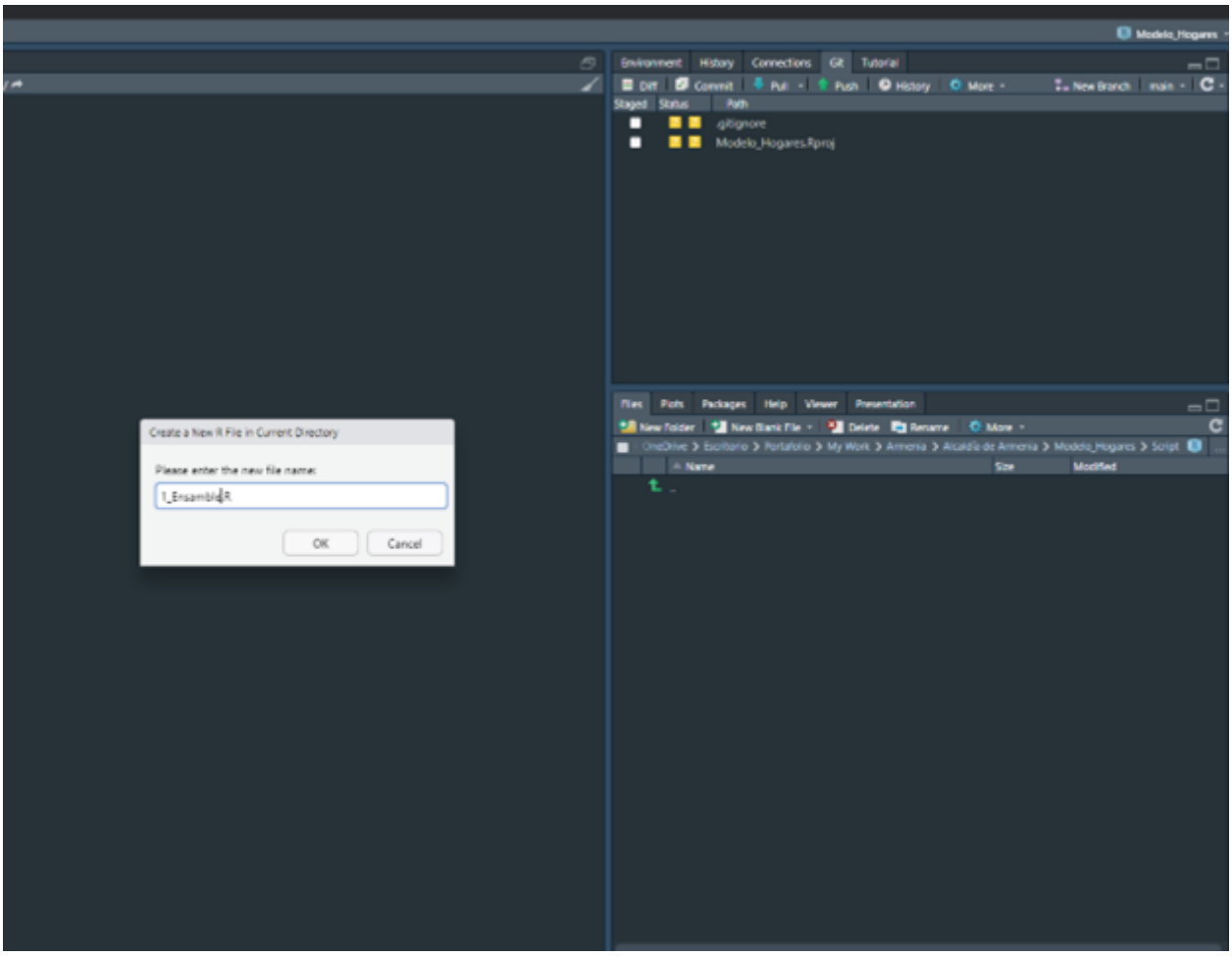
Inicialmente nos puede pedir autenticación

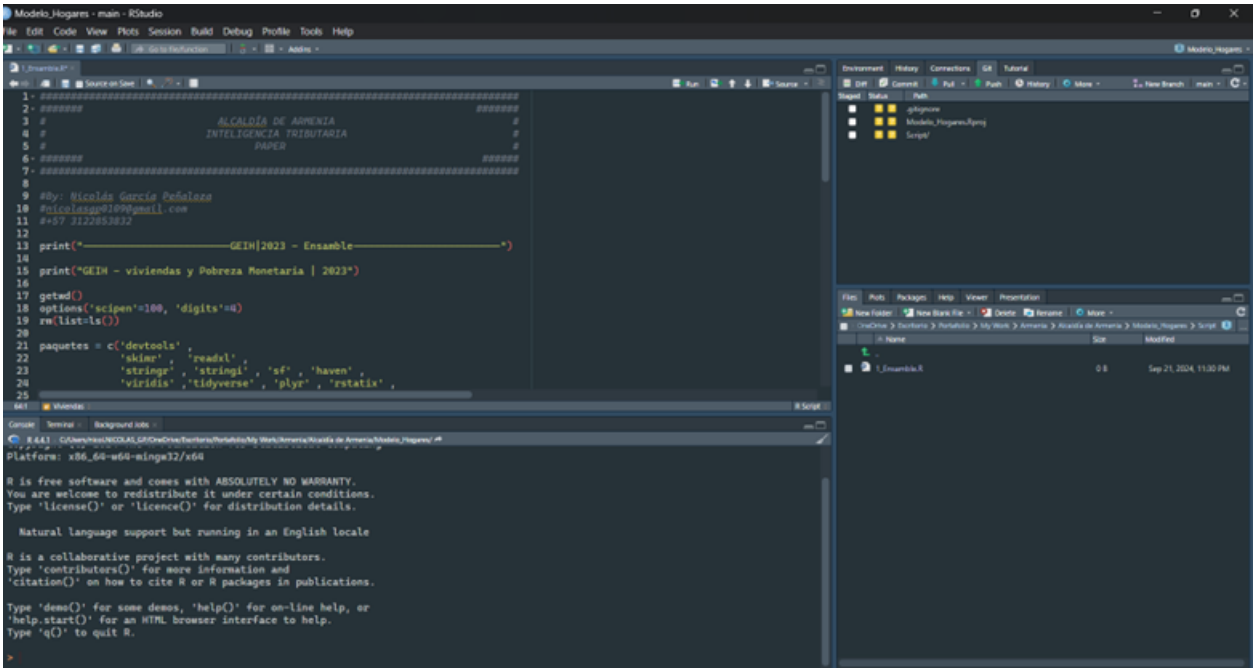


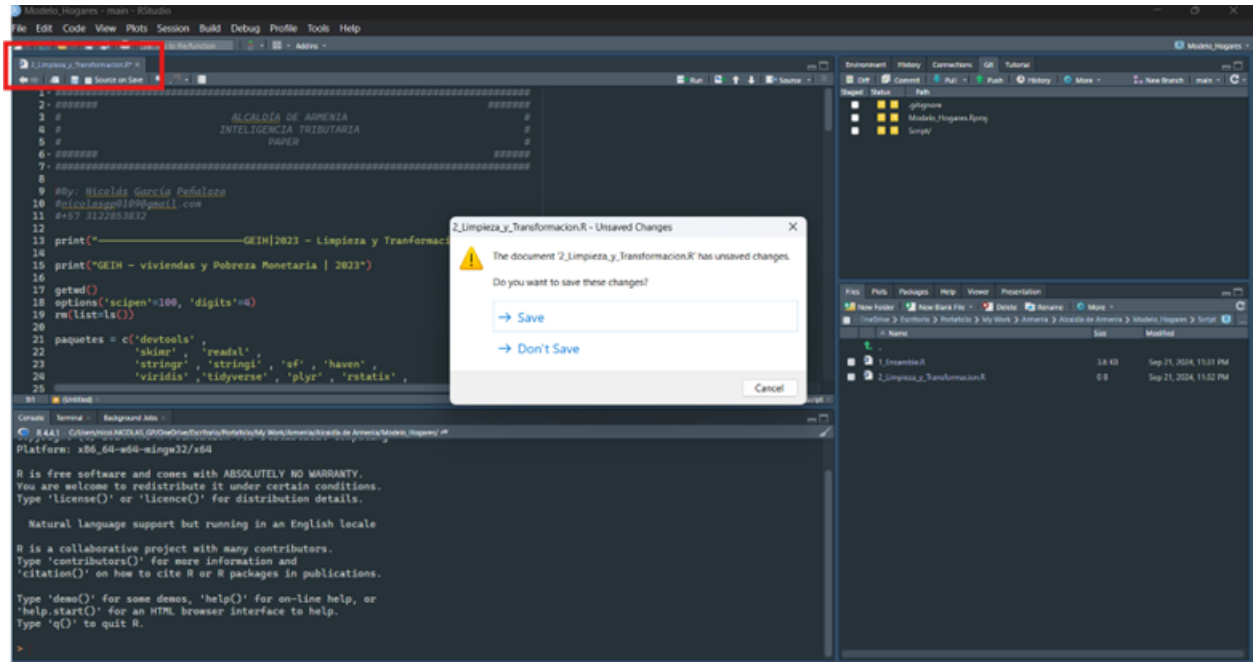

Modelo_Hogares
Environment
History
Connections
Git
Tutorial
Diff
Commit
Pull
Push
History
More
New Branch
main
Staged
Status
Path
.gitignore
Modelo_Hogares.Rproj
Files
Plots
Packages
Help
Viewer
Presentation
New Folder
New Blank File
Delete
Rename
More
COLAS_GP > OneDrive > Escritorio > Portafolio > My Work > Armenia > Alcaldía de Armenia > Modelo_Hogar
Name
Size
Modified
..
.gitignore
44 B
Sep 21, 2024, 11:25 PM
Modelo_Hogares.Rproj
218 B
Sep 21, 2024, 11:25 PM
README.md
91 B
Sep 21, 2024, 11:25 PM

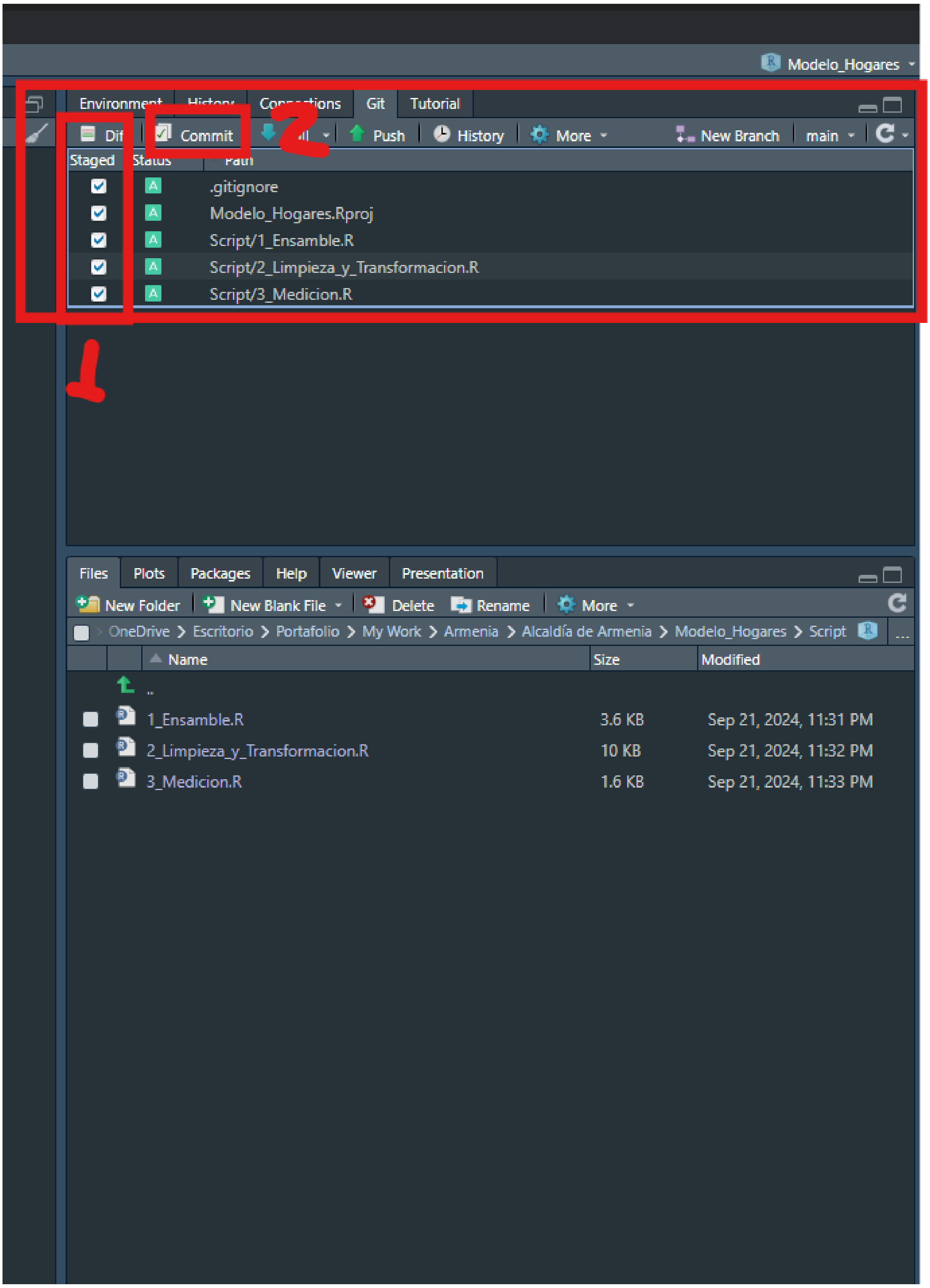

2
New Folder
Please enter the new folder name
Script
OK
Cancel
New Folder
New Blank File
Delete
Rename
More
Name
Size
Modified
.gitignore
Modelo_Hogares.Rproj
README.md

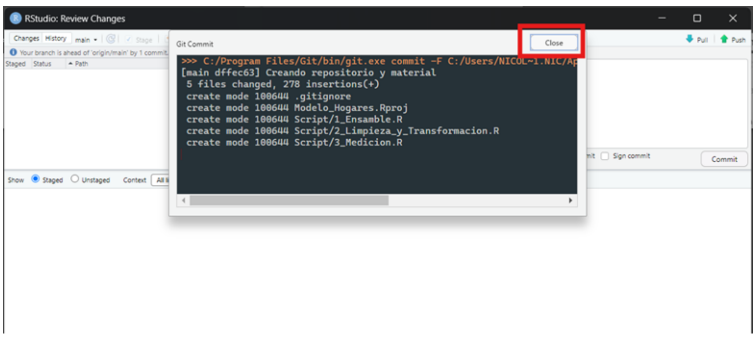

Files
Plots
Viewer
Presentation
New Folder
New Blank File
Delete
Rename
More
OneDrive
Work > Armenia > Alcaldía de Armenia > Modelo_Hogares > Script
Size
Modified
R Script
R Markdown...
Quarto Doc...
Shiny Web App...
Plumber API...
Text File
C++ File
Python Script
SQL Script
Stan File
D3 Script
R Sweave
R HTML

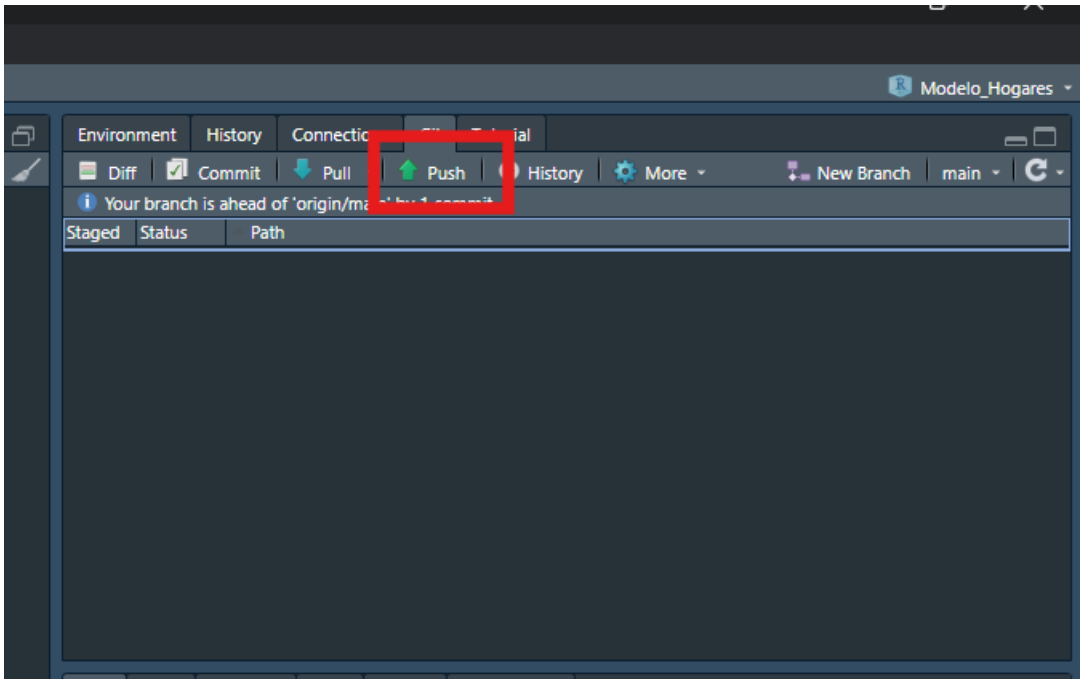

Modelo_Hogares
Environment
History
Connections
Git
Tutorial
Diff
Commit
Pull
Push
History
More
New Branch
main
Staged
Status
Path
.gitignore
Modelo_Hogares.Rproj
Files
Plots
Packages
Help
Viewer
Presentation
New Folder
New Blank File
Delete
Rename
More
OneDrive > Escritorio > Portafolio > My Work > Armenia > Alcaldía de Armenia > Modelo_Hogares > Script
Name
Size
Modified
Create a New R File in Current Directory
Please enter the new file name:
1_Ensamble.R
OK
Cancel

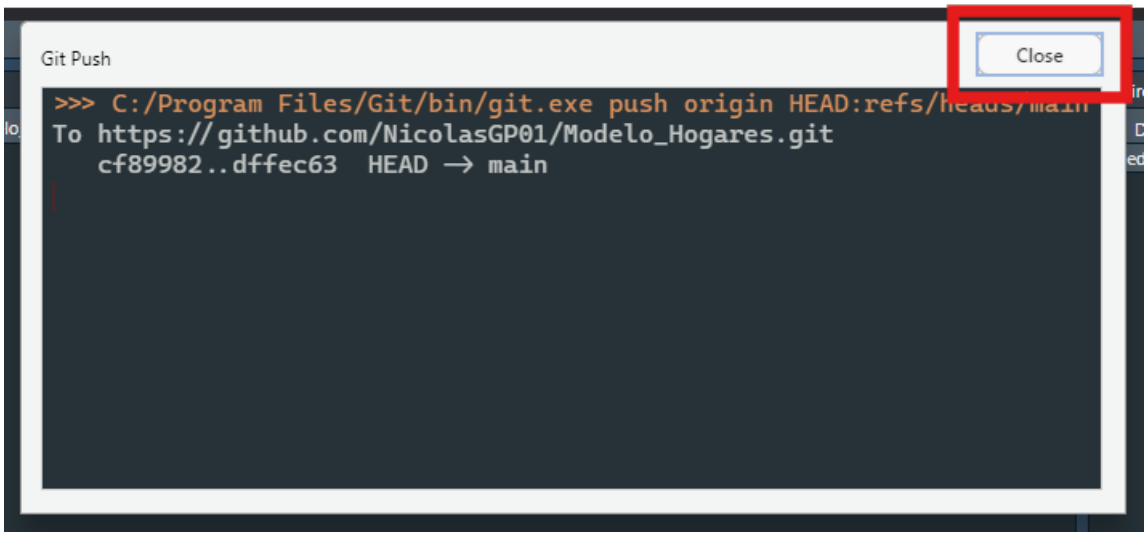

Modelo_Hogares - main - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
ALCALDÍA DE ARMENIA
INTELIGENCIA TRIBUTARIA
PAPER
#By: Nicolás García Peñaloza
#nicolasgp0109@gmail.com
#+57 3122853832
print("----------GEIH|2023 - Ensamble----------")
print("GEIH - viviendas y Pobreza Monetaria | 2023")
getwd()
options('scipen'=100, 'digits'=4)
rm(list=ls())
paquetes = c('devtools' ,
'skimr' , 'readxl' ,
'stringr' , 'stringi' , 'sf' , 'haven' ,
'viridis' ,'tidyverse' , 'plyr' , 'rstatix' ,
Console
Terminal
Background Jobs
Platform: x86_64-w64-mingw32/x64
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
Natural language support but running in an English locale
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

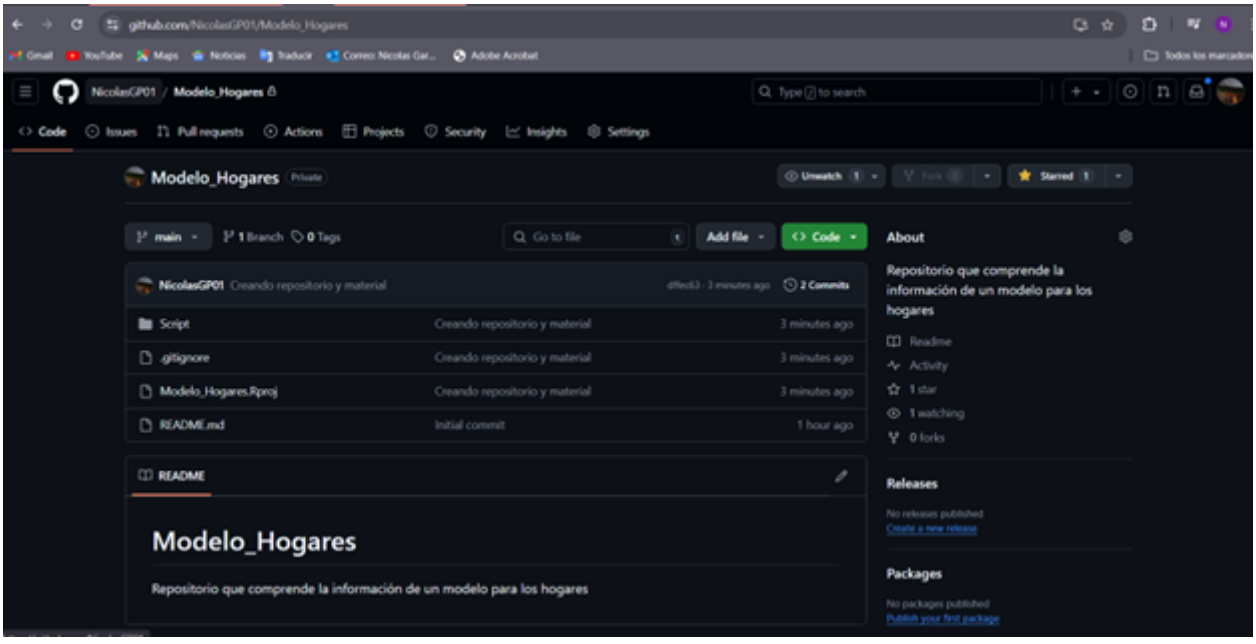

2_Limpieza_y_Transformacion.R - Unsaved Changes
The document '2_Limpieza_y_Transformacion.R' has unsaved changes.
Do you want to save these changes?
Save
Don't Save
Cancel

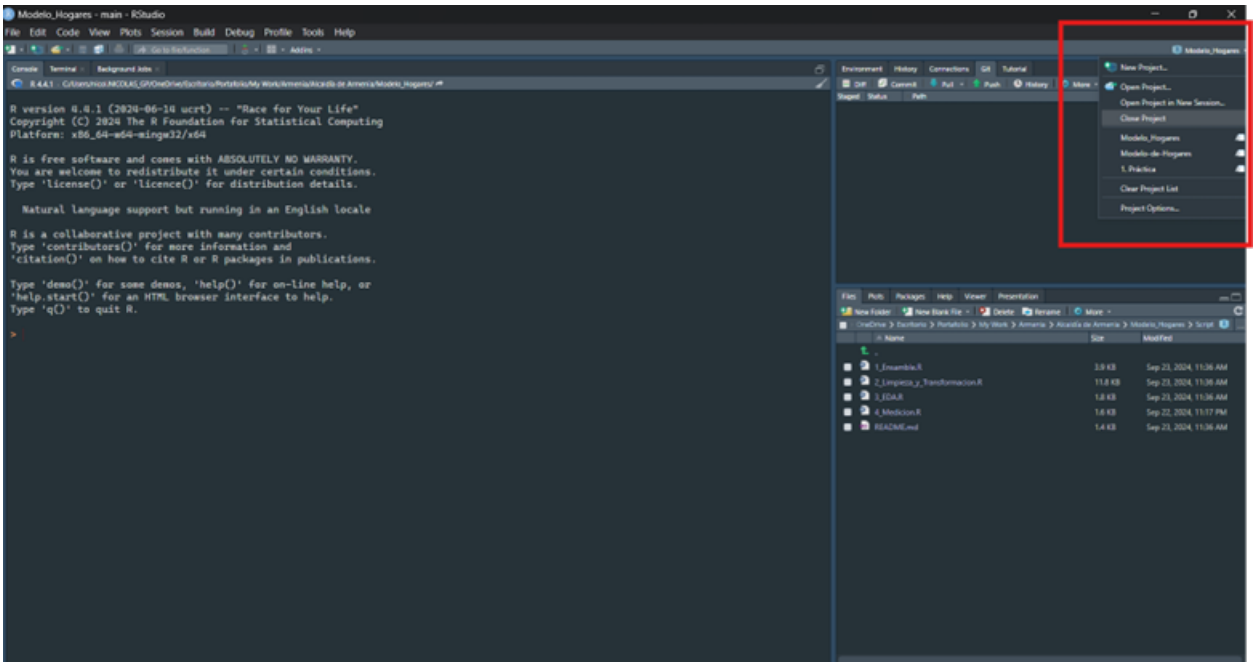

Modelo_Hogares
Environment
Connections
Git
Tutorial
Commit
Push
History
More
New Branch
main
Staged
Status
.gitignore
Modelo_Hogares.Rproj
Script/1_Ensamble.R
Script/2_Limpieza_y_Transformacion.R
Script/3_Medicion.R
1
2
Files
Plots
Packages
Help
Viewer
Presentation
New Folder
New Blank File
Delete
Rename
More
OneDrive > Escritorio > Portafolio > My Work > Armenia > Alcaldía de Armenia > Modelo_Hogares > Script
Name
Size
Modified
1_Ensamble.R
3.6 KB
Sep 21, 2024, 11:31 PM
2_Limpieza_y_Transformacion.R
10 KB
Sep 21, 2024, 11:32 PM
3_Medicion.R
1.6 KB
Sep 21, 2024, 11:33 PM


RStudio: Review Changes
Git Commit
Close
>>> C:/Program Files/Git/bin/git.exe commit -F C:/Users/NICOL~1.NIC/Ap
[main dffec63] Creando repositorio y material
5 files changed, 278 insertions(+)
create mode 100644 .gitignore
create mode 100644 Modelo_Hogares.Rproj
create mode 100644 Script/1_Ensamble.R
create mode 100644 Script/2_Limpieza_y_Transformacion.R
create mode 100644 Script/3_Medicion.R


Modelo_Hogares
Environment
History
Diff
Commit
Pull
Push
History
More
New Branch
main
Staged
Status
Path


Git Push
Close
>>> C:/Program Files/Git/bin/git.exe push origin HEAD:refs/heads/main
To https://github.com/NicolasGP01/Modelo_Hogares.git
cf89982..dffec63 HEAD → main

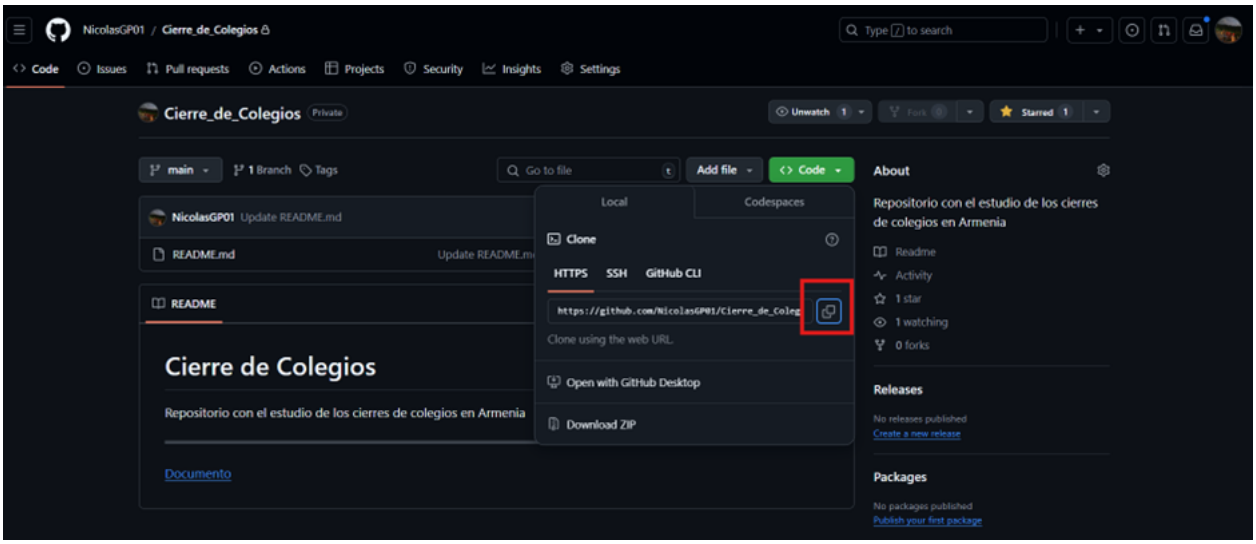

Modelo_Hogares
Code
Issues
Pull requests
Actions
Projects
Security
Insights
Settings
main
Add file
Code
About
Repositorio que comprende la información de un modelo para los hogares
README
Releases
Packages
Modelo_Hogares
Repositorio que comprende la información de un modelo para los hogares

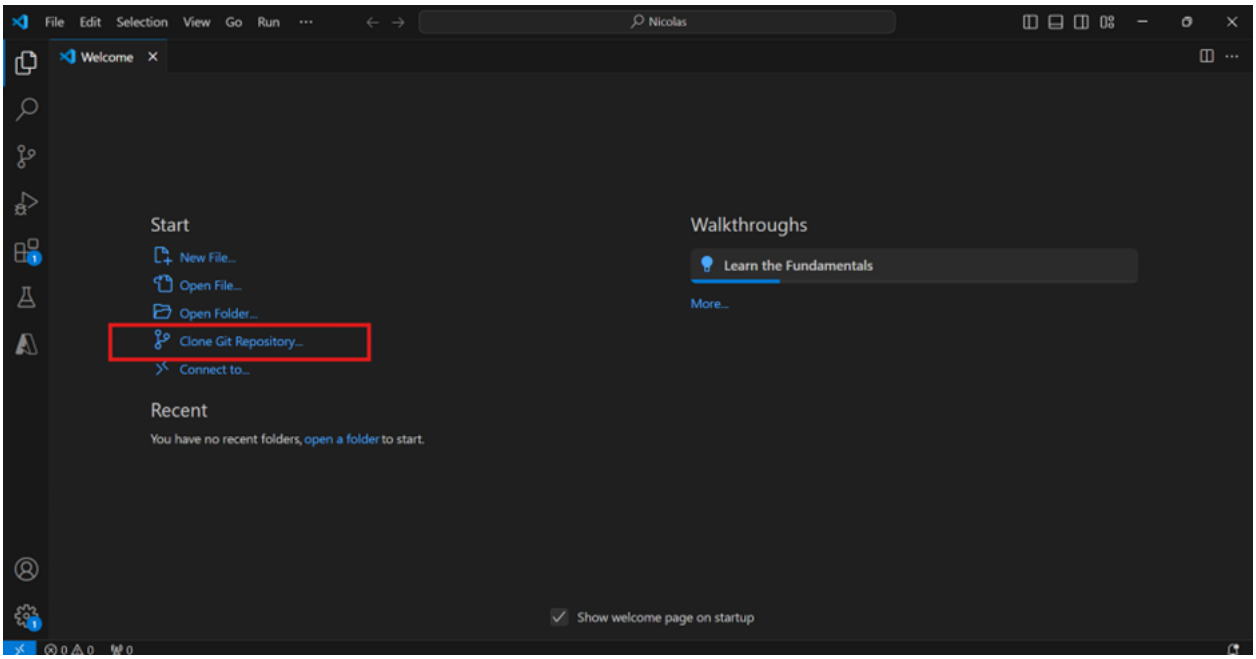


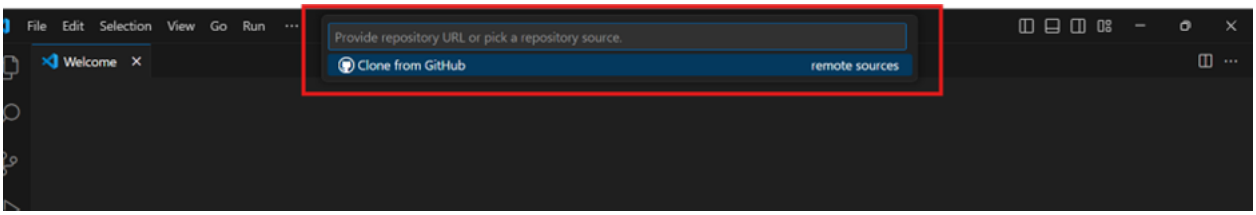
3. Python

En este caso desarrollaremos nuestro código de **Python** en **Visual Studio Code**.

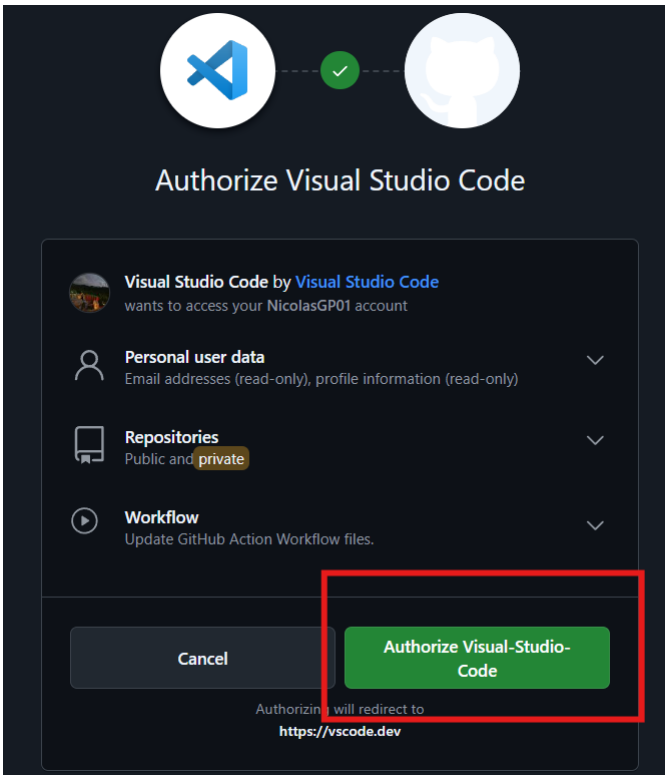
**Recuerde ya haber establecido el nombre y correo en GIT si apenas va empezar a utilizar este control de versiones con Python*



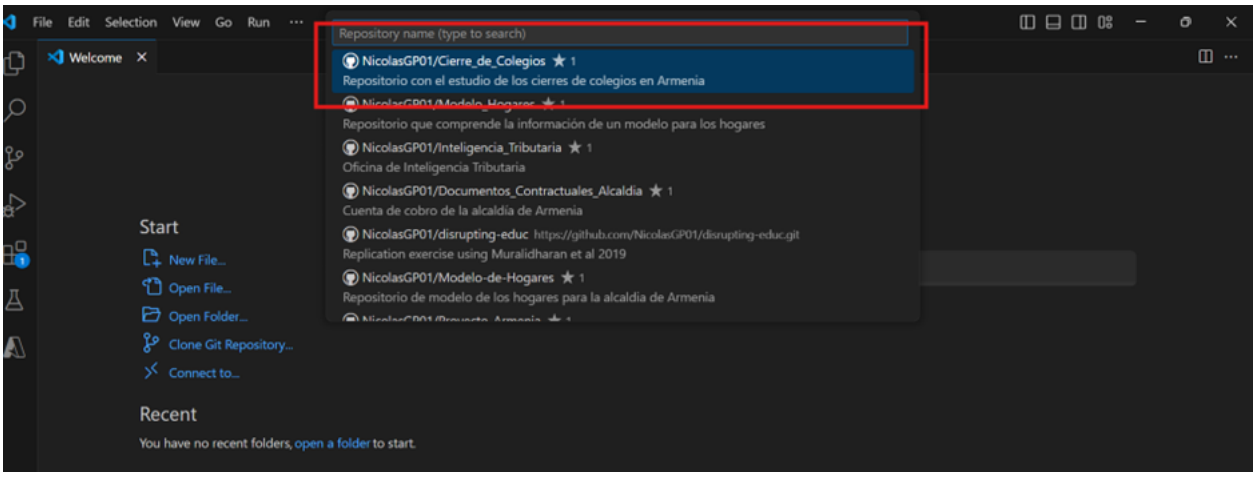


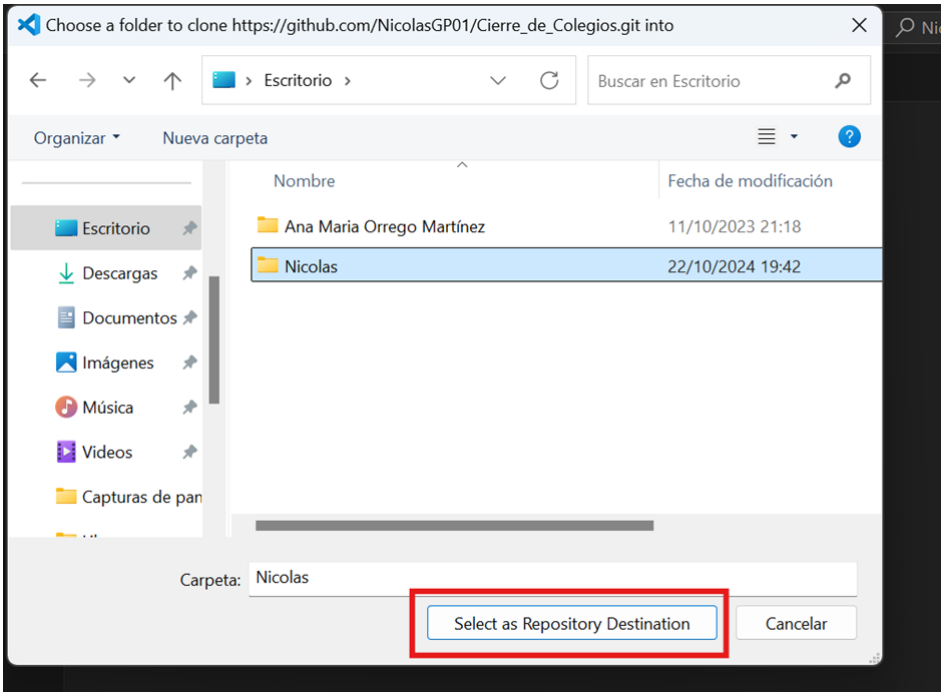


Le damos permitir

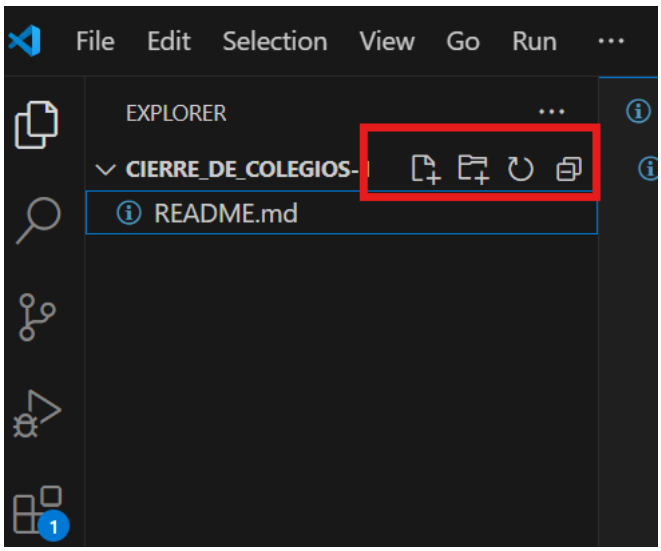


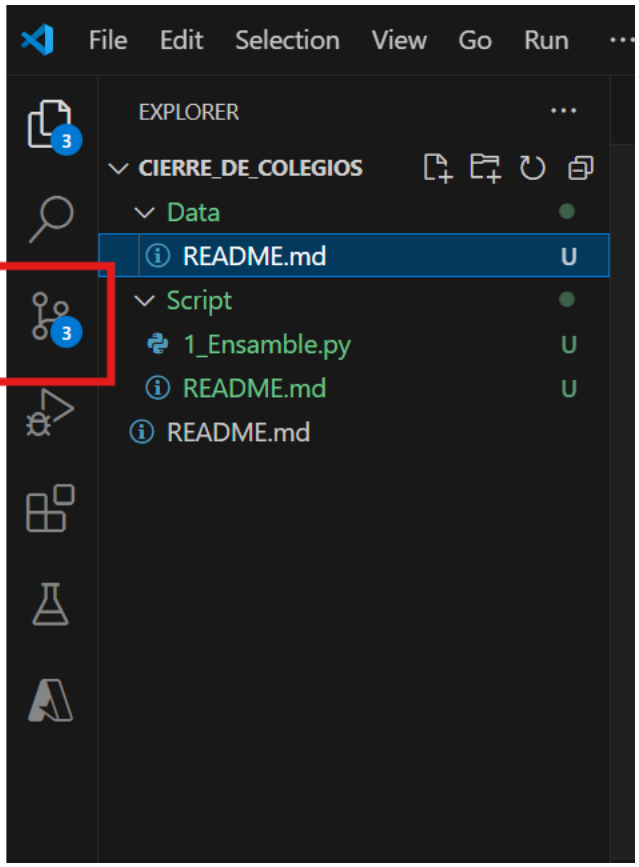
Elegimos el repositorio el cual vamos a trabajar



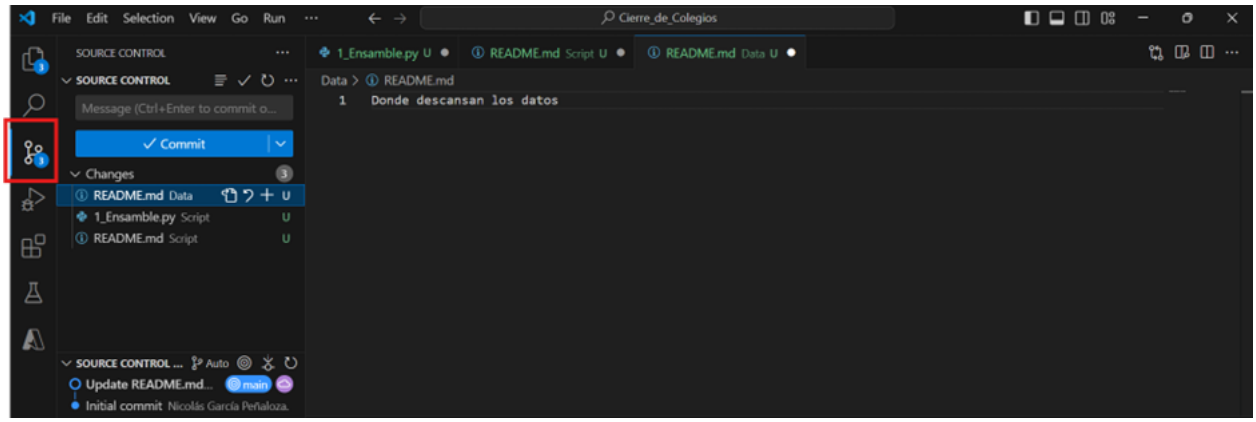


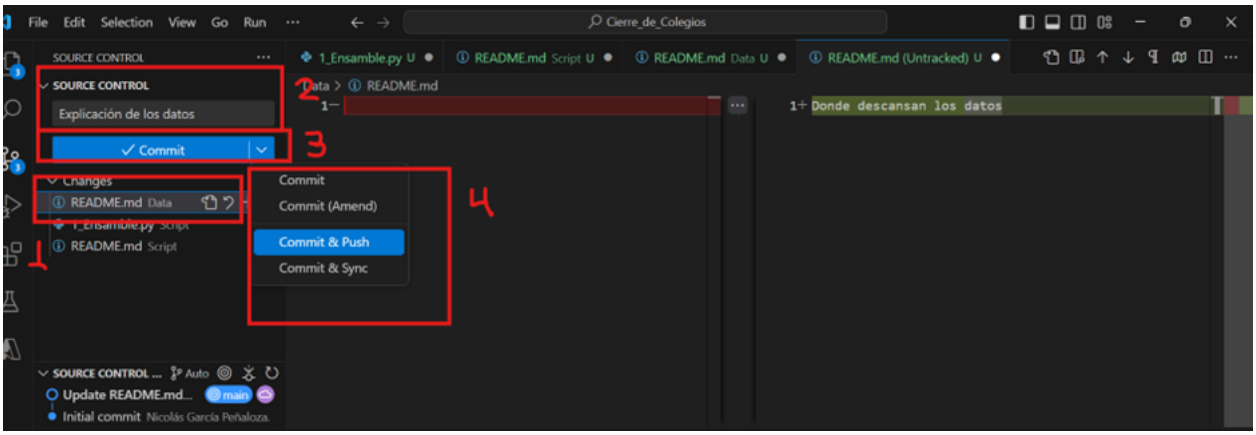
Le damos en verificar en la opción de su gusto Luego abrimos allí mismo el repositorio Podemos elegir abrir una carpeta o un script



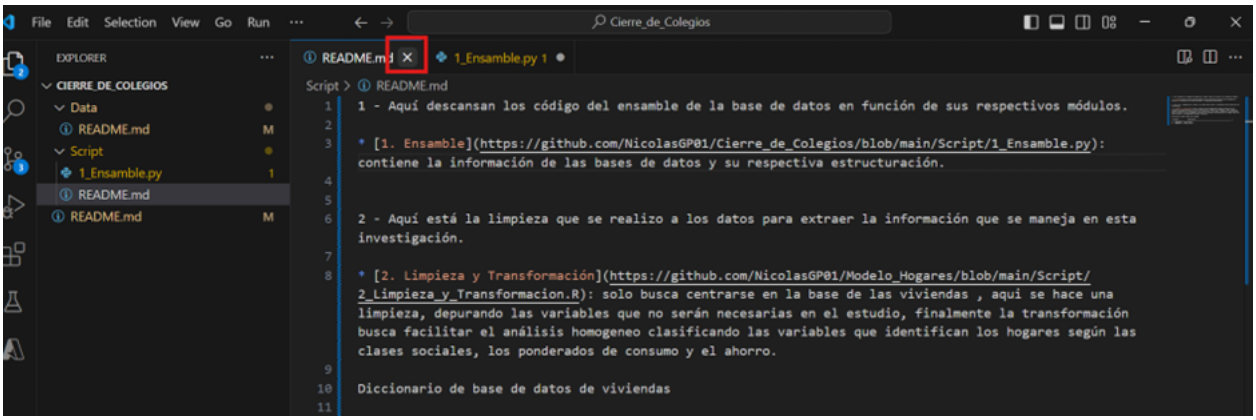


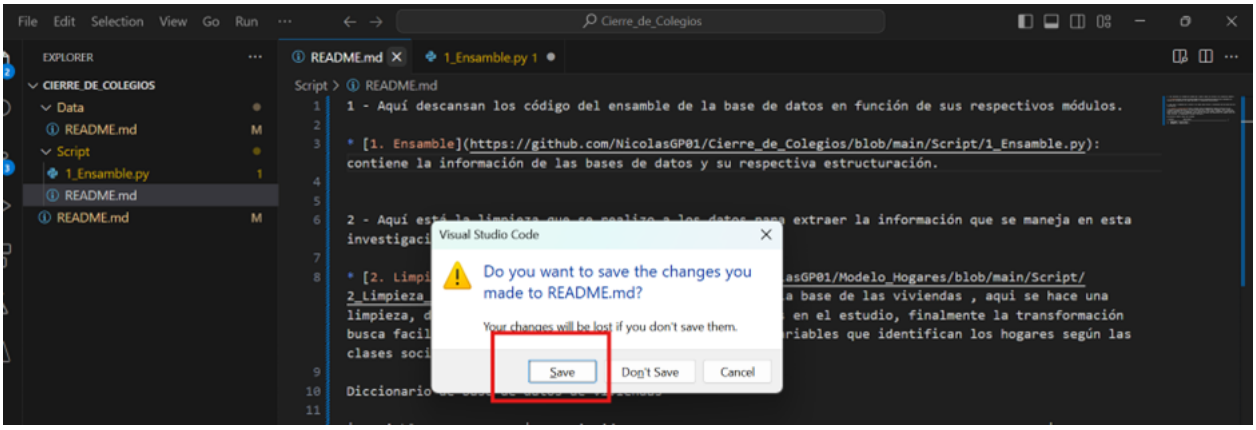
Enseguida de la U en verde del archivo debemos tener el símbolo + para que se puedan ejecutar los cambios





Una vez ya guardado y enviado para ejecutar los cambios debemos cerrar los archivos y guardarlos


File Edit Selection View Go Run
Cierre_de_Colegios
EXPLORER
README.md
1_Ensamble.py 1
CIERRE_DE_COLEGIOS
Data
README.md
Script
1_Ensamble.py
README.md
README.md
Script > README.md
1 - Aquí descansan los código del ensamble de la base de datos en función de sus respectivos módulos.
* [1. Ensamble](https://github.com/NicolasGP01/Cierre_de_Colegios/blob/main/Script/1_Ensamble.py):
contiene la información de las bases de datos y su respectiva estructuración.
2 - Aquí está la limpieza que se realizo a los datos para extraer la información que se maneja en esta
investigación.
* [2. Limpieza y Transformación](https://github.com/NicolasGP01/Modelo_Hogares/blob/main/Script/
2_Limpieza_y_Transformacion.R): solo busca centrarse en la base de las viviendas , aquí se hace una
limpieza, depurando las variables que no serán necesarias en el estudio, finalmente la transformación
busca facilitar el análisis homogeneo clasificando las variables que identifican los hogares según las
clases sociales, los ponderados de consumo y el ahorro.
Diccionario de base de datos de viviendas


File Edit Selection View Go Run
Cierre_de_Colegios
EXPLORER
README.md
1_Ensamble.py 1
CIERRE_DE_COLEGIOS
Data
README.md
Script
1_Ensamble.py
README.md
README.md
Script > README.md
1 - Aquí descansan los código del ensamble de la base de datos en función de sus respectivos módulos.
* [1. Ensamble](https://github.com/NicolasGP01/Cierre_de_Colegios/blob/main/Script/1_Ensamble.py):
contiene la información de las bases de datos y su respectiva estructuración.
Visual Studio Code
Do you want to save the changes you made to README.md?
Your changes will be lost if you don't save them.
Save
Don't Save
Cancel