

# GitHub

## Big Data y Machine Learning para Economía Aplicada

Ignacio Sarmiento-Barbieri

Universidad de los Andes

# Intro to Git(Hub)

## Motivation

### ► Git

- Git is a distributed version control system
- It is optimized for the things that economists and data scientists spend a lot of time working on (e.g. code).

### ► GitHub

- It's important to realize that Git and GitHub are distinct things.
- GitHub is an online hosting platform built on top of the Git system. (Like Bitbucket and GitLab.)
- Just like we don't *need* Rstudio to run R code, we don't *need* GitHub to use Git

# Collaboration time

- ▶ This is what Andrés and I will do
  - ▶ Andrés: Fork the Problem Set folder and invite me, Ignacio, to join as a collaborator on the PS GitHub repo (See the Settings tab of your repo.)
  - ▶ Andrés will add a script to the folder.
  - ▶ Ignacio: Clone Andrés repo to my local machine. Make some edits to the script. Stage, commit and push these changes.
  - ▶ Andrés: Make your own changes to the script on your local machine. Stage, commit and then try to push them

# Collaboration time

- ▶ Andrés is going to encounter a 'merge conflict' error.

# Collaboration time

- ▶ Andrés is going to encounter a 'merge conflict' error.

```
# script
Some text here.
<<<<<<< HEAD
Text added by Partner 2.
=====
Text added by Partner 1.
>>>>>>> 814e09178910383c128045ce67a58c9c1df3f558.
More text here.
```

- ▶ Git is protecting Andres by refusing the merge. It wants to make sure that you don't accidentally overwrite all of your changes by pulling my version of the script.

# Collaboration time

- ▶ Fixing these conflicts is a simple matter of (manually) editing the script file.
  - ▶ Delete the lines of the text that you don't want.
  - ▶ Then, delete the special Git merge conflict symbols.
- ▶ Once that's done, you should be able to stage, commit, pull and finally push your changes to the GitHub repo without any errors.
  - ▶ Andrés gets to decide what to keep because they fixed the merge conflict.
  - ▶ The full commit history is preserved, I can always recover their changes if desired.
  - ▶ Another solution is using branches

# Ramas o Branches

# Branches





# Branches

