Setting up git for (solo) data science workflow

Ronald (Ryy) Glenn Thomas

8/28/23

Table of contents

1 Introduction

Version Control for biostatistics is the challenge.

Lets take it one step at a time.

Scenario 1: rgt47 has been working on a data analysis for some ADNI data. Its moderately complex and uses lots of packages. He's ready to have his team join the analysis process. What are the first steps to do that? Start by adding git user rgt4748 to the team….

2 Methods

create a branch:

create new branch for testing

git checkout -b test; git pull origin master; git push origin test; git checkout master; git pull origin test

- #merging branch back into master
-
- git checkout test
- git pull origin master
- git checkout master
- git merge test
- git push origin master
-
-
- #to delete branch
- git branch -d test
- git push origin –delete test

Draft ….

GIT for nitwits

git init

git add fname

git status #see what happens on commit git commit -am “commit message”

git push

git branch work

git checkout work

… make changes … git add * git commit -m “something”

git checkout master

git merge work

git branch -d work

git log #see all commits

git checkout HASH #Restore old branch

Consider editing ./.git/config

View file in master branch. git show master:a101.Rmd | mvim -

Copy file from other branch (master) git checkout master uw.png

Troubleshooting git pull –allow-unrelated-histories