Instructions for

Collaborating on GitHub Tutorial

Breakout 1: Connecting with your partner

- Say hello and introduce yourself
- Connect either on Slack (direct messaging), Zoom chat (make sure to select your partner, not send messages to everyone), phone, whatever you prefer
- Decide who will be Partner 1 and who will be Partner 2
- Exchange GitHub user names
- Practice sharing screen (choose to share Desktop so you can share as you jump between Rstudio and your web browser)

Breakout 2: Create repo, clone it, and try making changes

- Work through sections 9.2 – 9.7 here: http://ohi-science.org/data-science-training/collaborating.html

Breakout 3: Create and resolve merge conflicts

- Work through sections 9.8 – 9.9 here: http://ohi-science.org/data-science-training/collaborating.html

Breakout 4: Set up your collaborative website

- Work through sections 9.11 – 9.12 here: http://ohi-science.org/data-science-training/collaborating.html

Breakout 5: Explore the NYC flights data collaboratively

- Partner 2 add an image to your index.Rmd (talk Partner 1 through what you're doing or share your screen):
  - Search on google to find an image you like
  - On the image, right click and choose "Copy image address"
  - Add it to the .Rmd file with
  - Pull – stage – commit – push
  - Check that the image shows up on your website
    Remember address format for website:
    - my github repo: https://github.com/jules32/collab-research
    - my website url: https://jules32.github.io/collab-research/
    - note that the url starts with my **username.github.io**

Breakout 5: Explore the NYC flights data collaboratively

- With your partner, do the following tasks. Each of you should work on one task at a time. Since we're working closely on the same document, talk to each other and have one person create a heading and a R chunk, and then sync; the other person can then create a heading and R chunk and sync, and then you can both work safely.

- Remember to make your commit messages useful!

- As you work, you may get merge conflicts. Resolve them

Breakout 5: Explore the NYC flights data collaboratively

- Find all flights that:
  - Had an arrival delay of two or more hours
  - Flew to Houston (IAH or HOU)
  - Were operated by United, American, or Delta
  - Departed in summer (July, August, and September)
  - Arrived more than two hours late, but didn't leave late
  - Were delayed by at least an hour, but made up over 30 minutes in flight
  - Departed between midnight and 6am (inclusive)

- Another useful dplyr filtering helper is between(). What does it do? Can you use it to simplify the code needed to answer the previous challenges?

Breakout 5: Explore the NYC flights data collaboratively

- If you have more time:
  - Add some plots: What interesting patterns can we explore in this dataset? Examples:
    - What is the relationship between the departure delays and arrival delays (do planes tend to catch up while underway)?
    - How to arrival delays compare between different airlines? Or destinations? Or days (or any combination of those)