

Introduction to R

Harvard Chan Bioinformatics Core

<https://tinyurl.com/hbc-Epi-RBootcamp>



Shannan Ho Sui
Director



John Hutchinson
Associate Director



Victor Barrera



Zhu Zhuo



Radhika Khetani
Training Director



Meeta Mistry



Jihe Liu



Will Gammerdinger



Maria Simoneau



James Billingsley



Sergey Naumenko



Peter Kraft
Faculty Advisor



Shannan Ho Sui
Director



John Hutchinson
Associate Director



Victor Barrera



Zhu Zhuo



Radhika Khetani
Training Director



Meeta Mistry



Jihe Liu



Will Gammerdinger



Maria Simoneau



James Billingsley



Sergey Naumenko



Peter Kraft
Faculty Advisor

Introductions!



Shannan Ho Sui
Director



John Hutchinson
Associate Director



Victor Barrera



Zhu Zhuo



Radhika Khetani
Training Director



Meeta Mistry



Jihe Liu



Will Gammerdinger



Maria Simoneau



James Billingsley

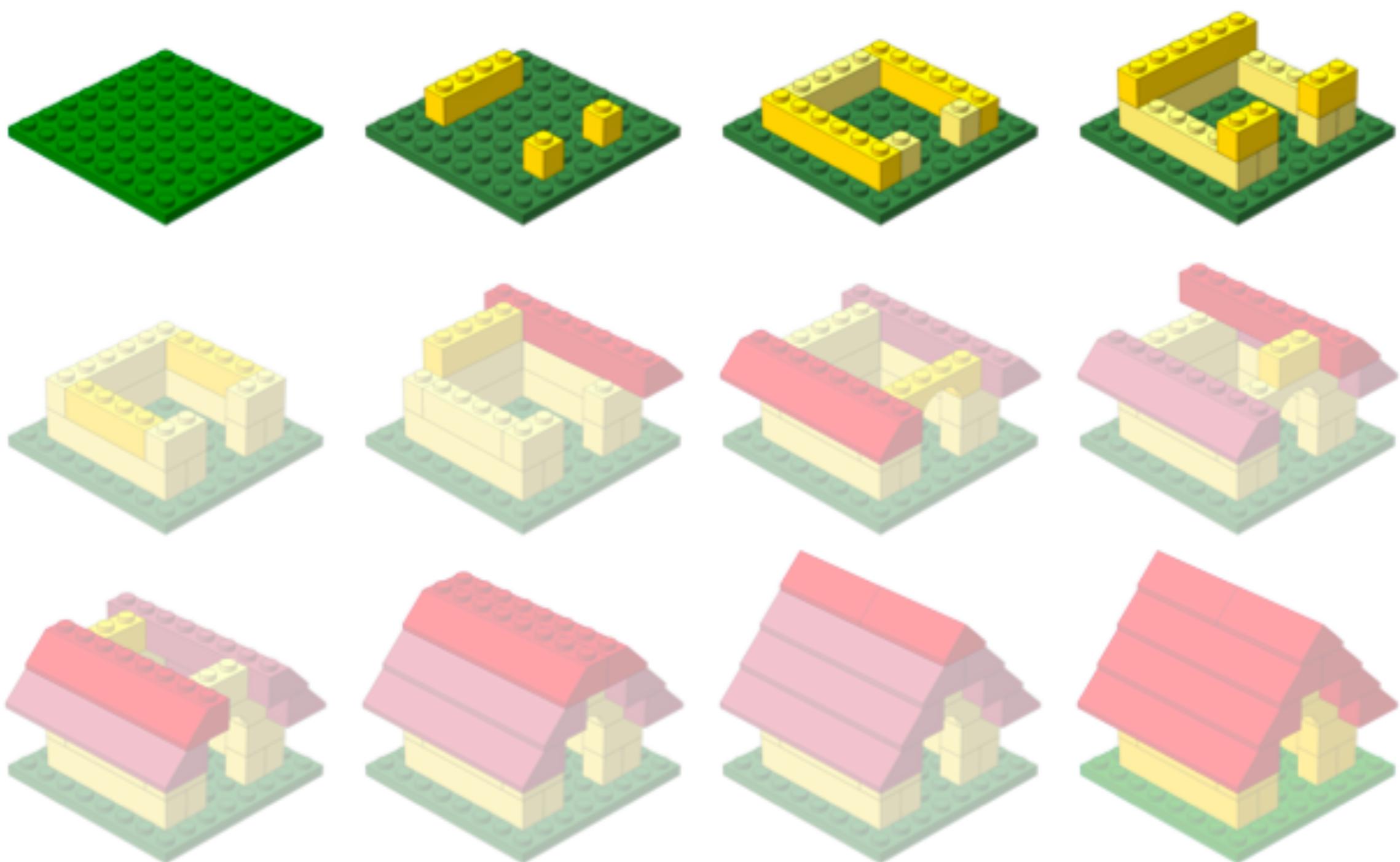


Sergey Naumenko



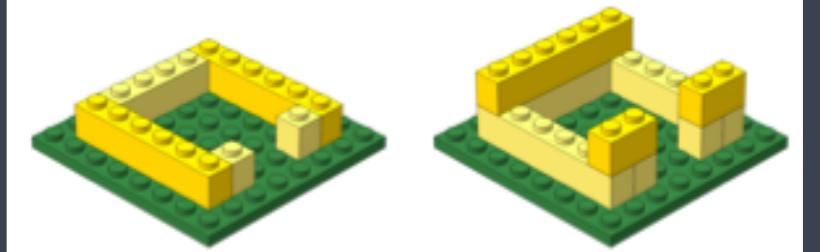
Peter Kraft
Faculty Advisor

Workshop Scope...



Learning R

Workshop Scope



- ✓ Comfortably use RStudio (a graphical interface for R)
- ✓ Fluently interact with R using RStudio
- ✓ Become familiar with R syntax
- ✓ Understand data structures in R
- ✓ Inspect and manipulate data structures
- ✓ Install packages and use functions in R

CRAN

(Comprehensive R Archive Network)



Available CRAN Packages By Name

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#)

A3	Accurate, Adaptable, and Accessible Error Metrics for Predictive Models
abbyyR	Access to Abbyy Optical Character Recognition (OCR) API
abc	Tools for Approximate Bayesian Computation (ABC)
ABCanalysis	Computed ABC Analysis
abc.data	Data Only: Tools for Approximate Bayesian Computation (ABC)
abcdeFBA	ABCDE_FBA: A-Biologist-Can-Do-Everything of Flux Balance Analysis with this package
ABCOptim	Implementation of Artificial Bee Colony (ABC) Optimization
ABCp2	Approximate Bayesian Computational Model for Estimating P2
abcrf	Approximate Bayesian Computation via Random Forests

*CRAN
Mirrors
What's new?
Task Views
Search

About R
R Homepage
The R Journal*

- The main repository for R packages
- Easy to install

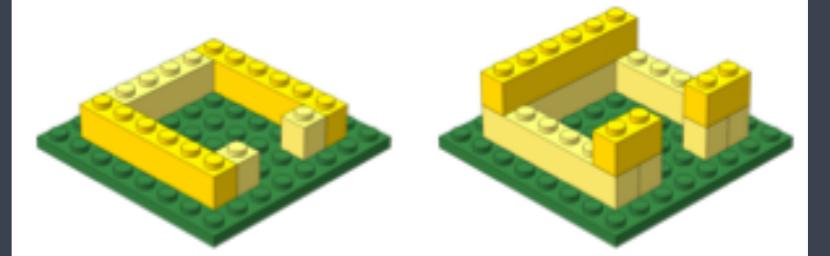
<https://cran.r-project.org/>



- An alternative package repository; “..provides tools for the analysis and comprehension of *high-throughput genomic data*.”
- Includes (but is not limited to) tools for:
 - performing statistical analysis
 - accessing public datasets
- Open source and open development
- Free

www.bioconductor.org

Workshop Scope



- Comfortably use RStudio (a graphical interface for R)
 - Fluently interact with R using RStudio
 - Become familiar with R syntax
 - Understand data structures in R
 - Inspect and manipulate data structures
 - Install packages and use functions in R
- ✓ Visualize data using *ggplot2*
- ✓ Utilize pipes, tibbles and functions from the Tidyverse package suite

Logistics

Course webpage

<https://tinyurl.com/hbc-Epi-RBootcamp>

Course schedule online

Workshop Schedule

Day 1

Time	Topic	Instructor
13:00 - 13:30	Workshop Introduction	Meeta
13:30 - 14:15	Introduction to R and RStudio	Jihe
14:15 - 14:20	Break	
14:20 - 15:00	R Syntax and Data Structure	Meeta

Day 2

Time	Topic	Instructor
13:00 - 13:45	Functions and Arguments	Jihe

Course materials online



Introduction to R

[View on GitHub](#)

Approximate time: 70 min

Learning Objectives

- Employ variables in R.
- Describe the various data types used in R.
- Construct data structures to store data.

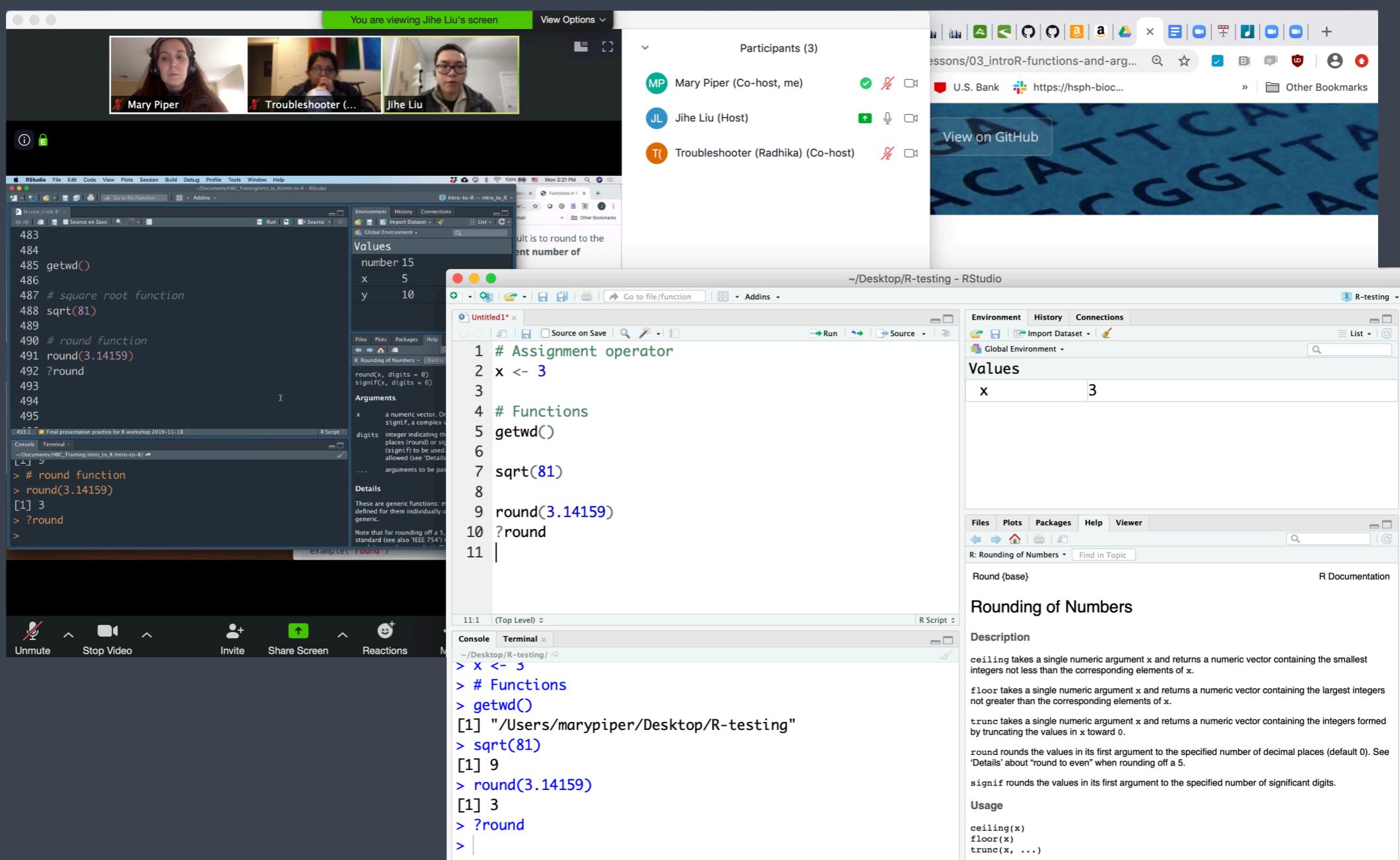
The R syntax

Now that we know how to talk with R via the script editor or the console, we want to use R for something more than adding numbers. To do this, we need to know more about the R syntax.

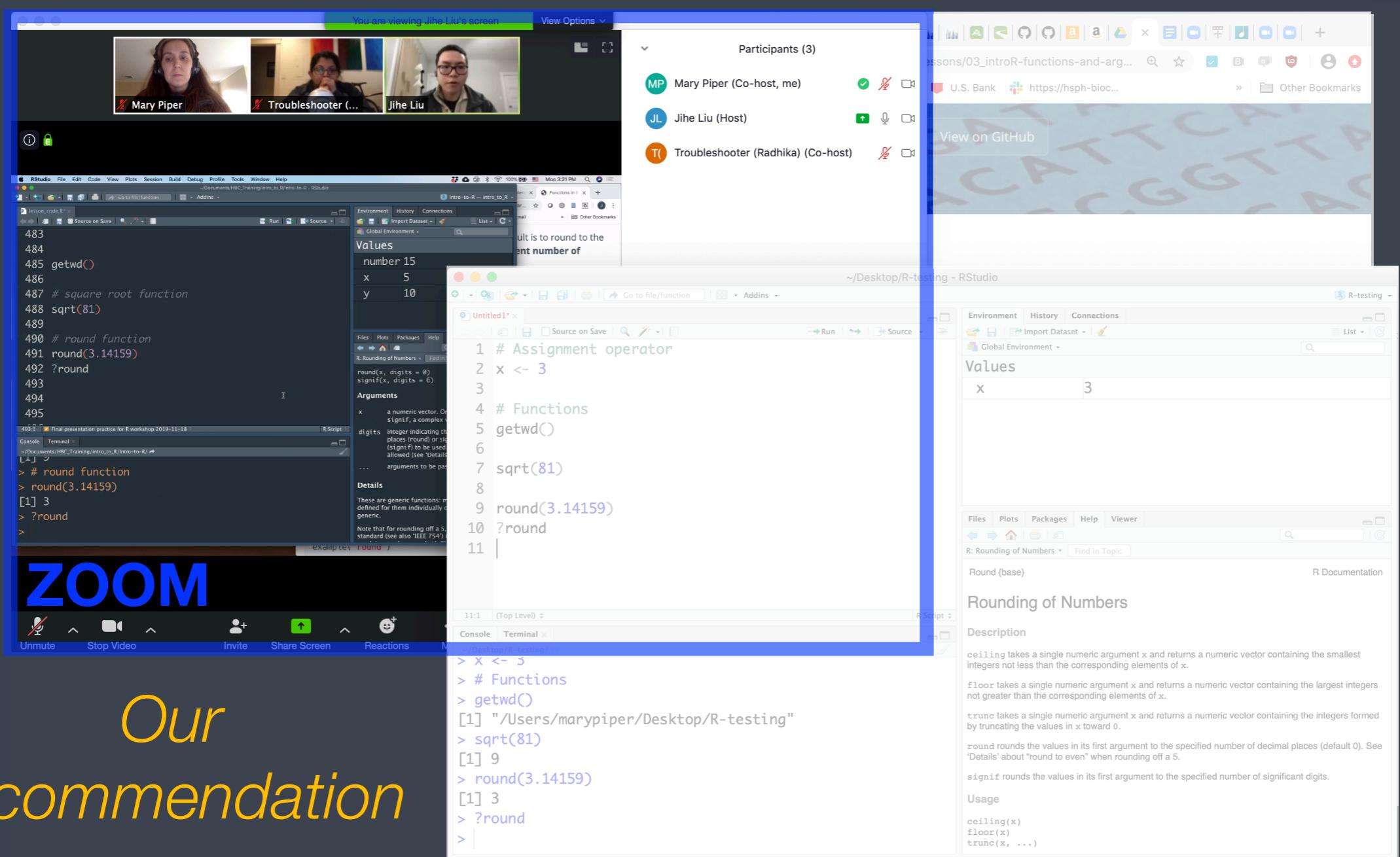
Below is an example script highlighting the many different “parts of speech” for R (syntax):

- the **comments** `#` and how they are used to document function and its content
- **variables and functions**
- the **assignment operator** `<-`

Single screen & 3 windows?

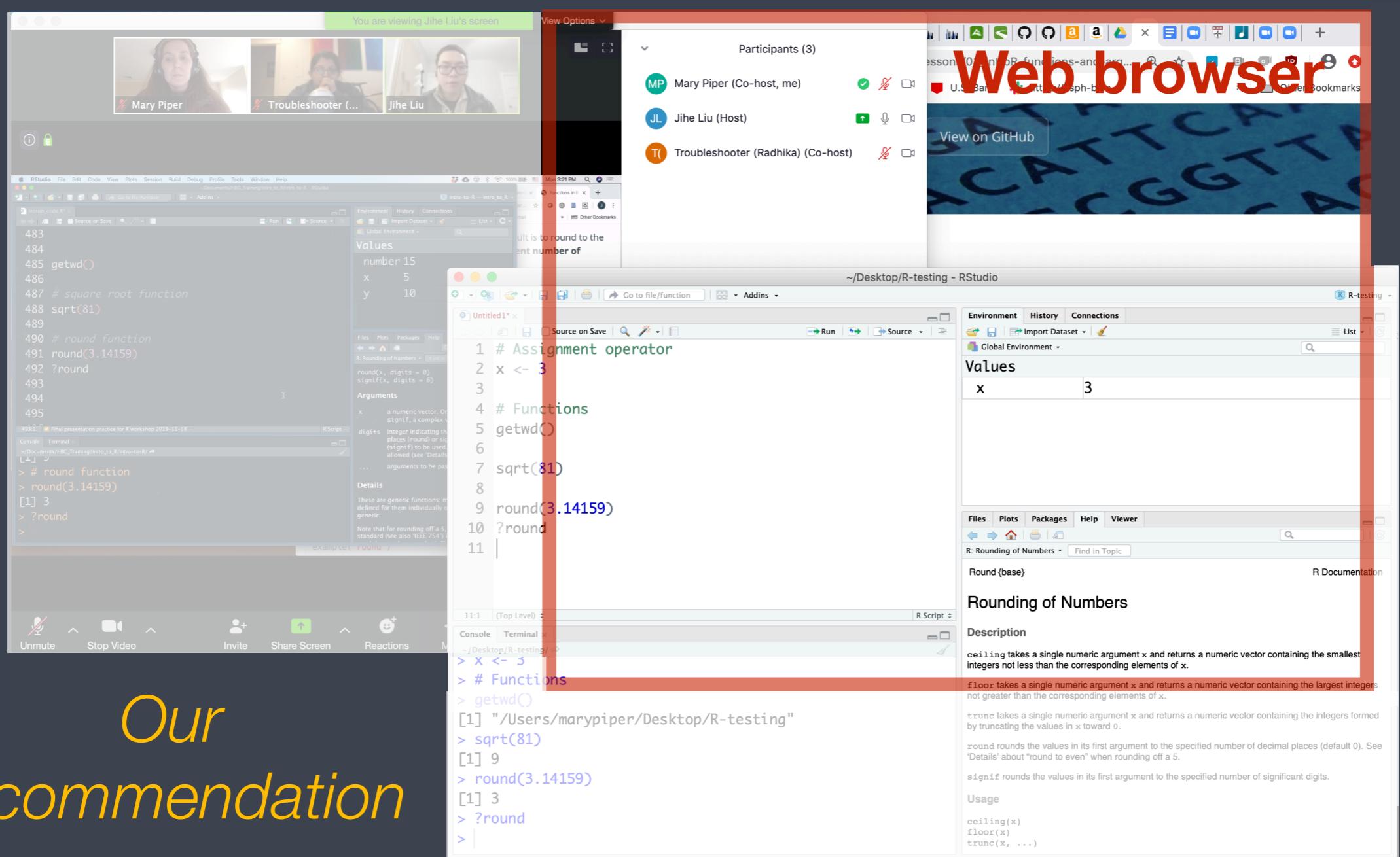


Single screen & 3 windows?



*Our
recommendation*

Single screen & 3 windows?



Our
recommendation

Single screen & 3 windows?

The screenshot shows a video conference interface with three windows:

- Video Feed:** Shows three participants: Mary Piper, Troubleshooter (Radhika), and Jihe Liu.
- RStudio Environment Window:** Contains an R script editor with the following code:

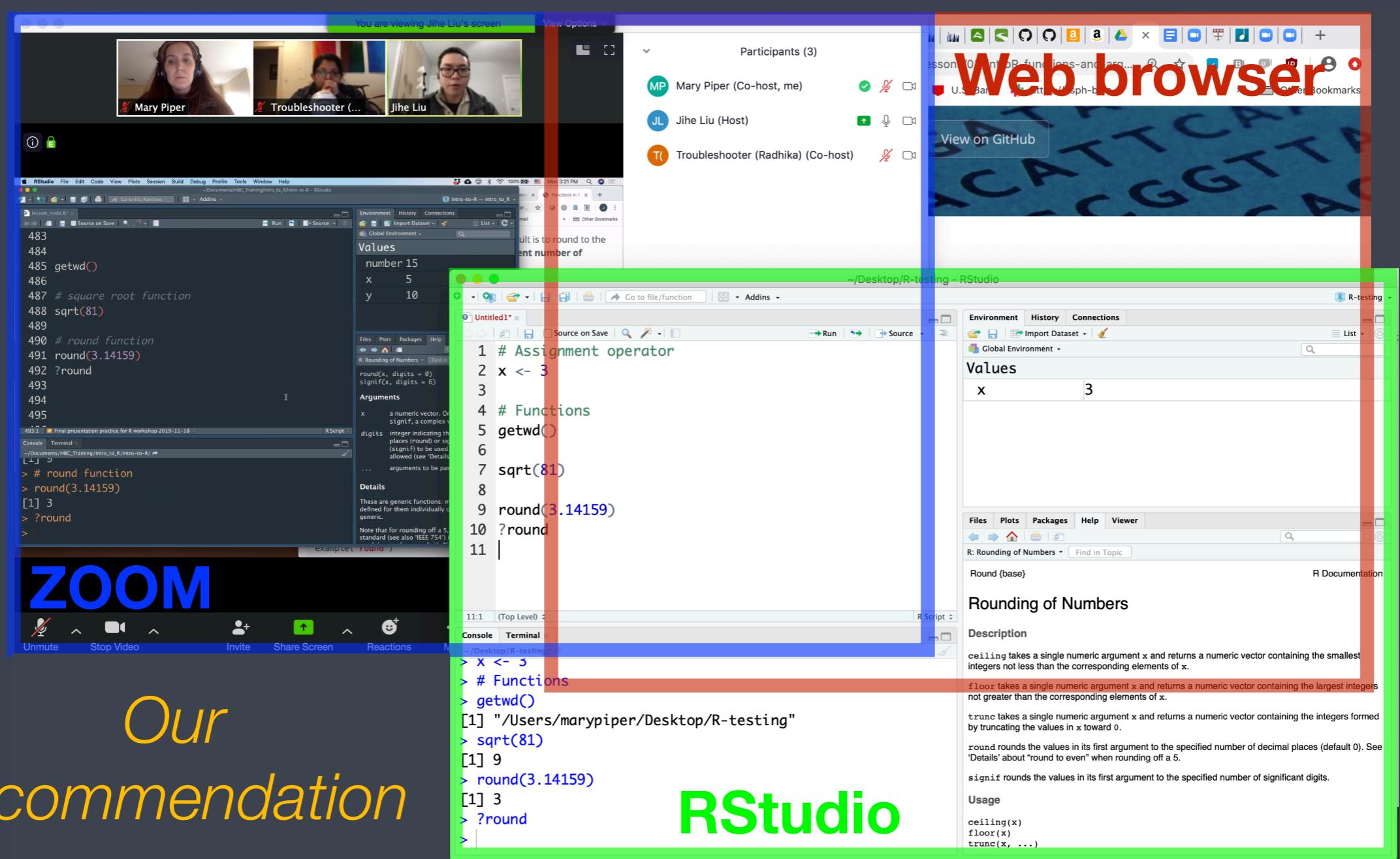
```
483  
484  
485 getwd()  
486  
487 # square root function  
488 sqrt(81)  
489  
490 # round function  
491 round(3.14159)  
492 ?round  
493  
494  
495
```

Console output:

```
> # round function  
> round(3.14159)  
[1] 3  
> ?round  
>
```
- Web Browser Window:** Displays a GitHub page for "intro-to-R-functions-and-args".

Our recommendation: RStudio

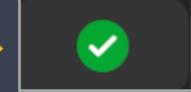
Single screen & 3 windows?

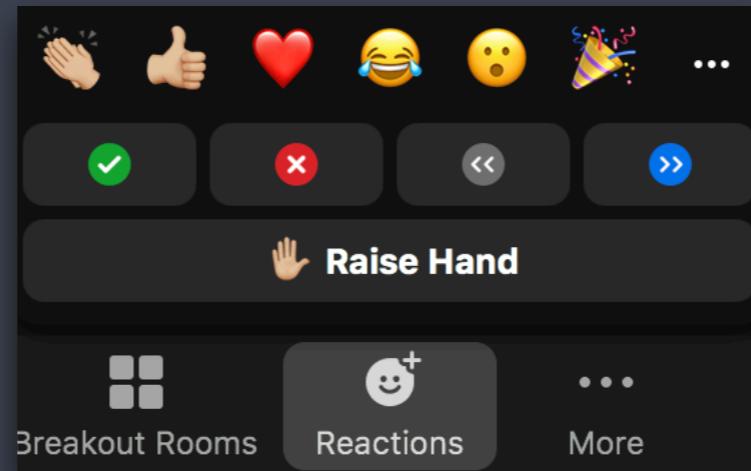


Odds and Ends

- ❖ Quit/minimize all applications that are not required for class

Odds and Ends (1/2)

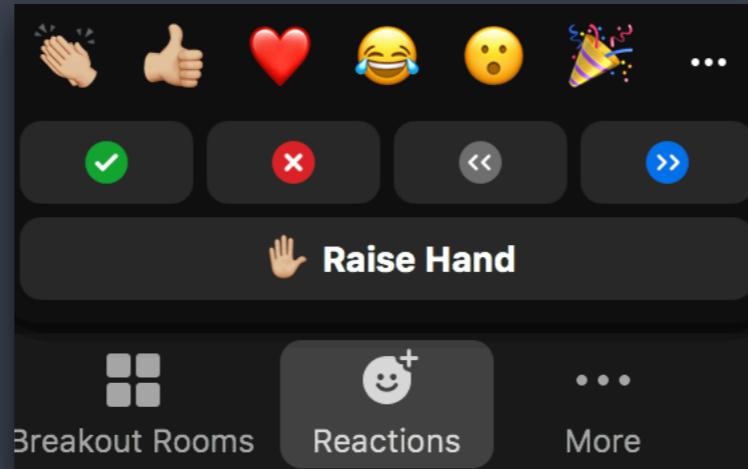
- ❖ Quit/minimize all applications that are not required for class
- ❖ Are you all set?
 - ▶  = "agree", "I'm all set" (equivalent to a **green post-it**)
 - ▶  = "disagree", "I need help" (equivalent to a **red post-it**)



Odds and Ends (2/2)

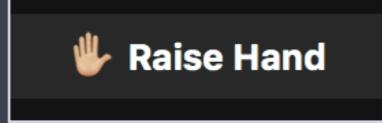
- ❖ Questions for the presenter?

- Post the question in the Chat window OR
-  when the presenter asks for questions
- Let the Moderator know



Odds and Ends (2/2)

- ❖ Questions for the presenter?

- Post the question in the Chat window OR
-  when the presenter asks for questions
- Let the Moderator know

- ❖ Technical difficulties with RStudio or R?

- Start a private chat with the Troubleshooter with a description of the problem.

Contact us!

HBC training team: hbctraining@hsph.harvard.edu

HBC consulting: bioinformatics@hsph.harvard.edu

Twitter

[@bioinfocore](https://twitter.com/bioinfocore)