

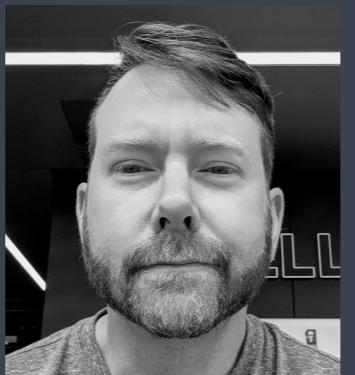
Introduction to R

Harvard Chan Bioinformatics Core

<https://tinyurl.com/bph-intro-r>



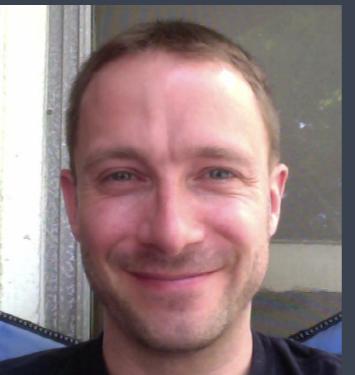
Shannan Ho Sui
Director



John Hutchinson
Associate Director



Victor Barrera



Rory Kirchner



Zhu Zhuo



Preetida Bhetariya



Meeta Mistry



Mary Piper
Assoc. Training Director



Jihe Liu



Radhika Khetani
Training Director



Maria Simoneau
Project coordinator



James Billingsley



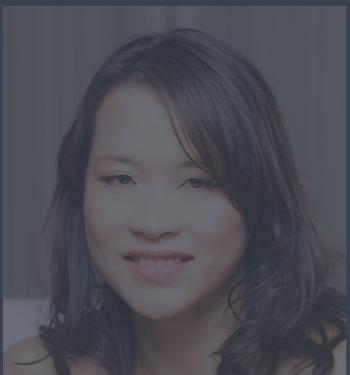
Sergey Naumenko



Joon Yoon



Peter Kraft
Faculty Advisor



Shannan Ho Sui
Director



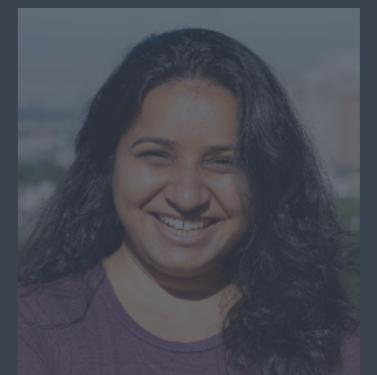
John Hutchinson
Associate Director



Victor Barrera



Zhu Zhuo



Preetida Bhetariya



Meeta Mistry



Mary Piper
Assoc. Training Director



Jihe Liu



Radhika Khetani
Training Director



Maria Simoneau



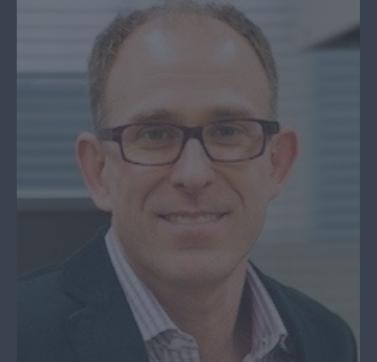
James Billingsley



Sergey Naumenko



Joon Yoon



Peter Kraft
Faculty Advisor

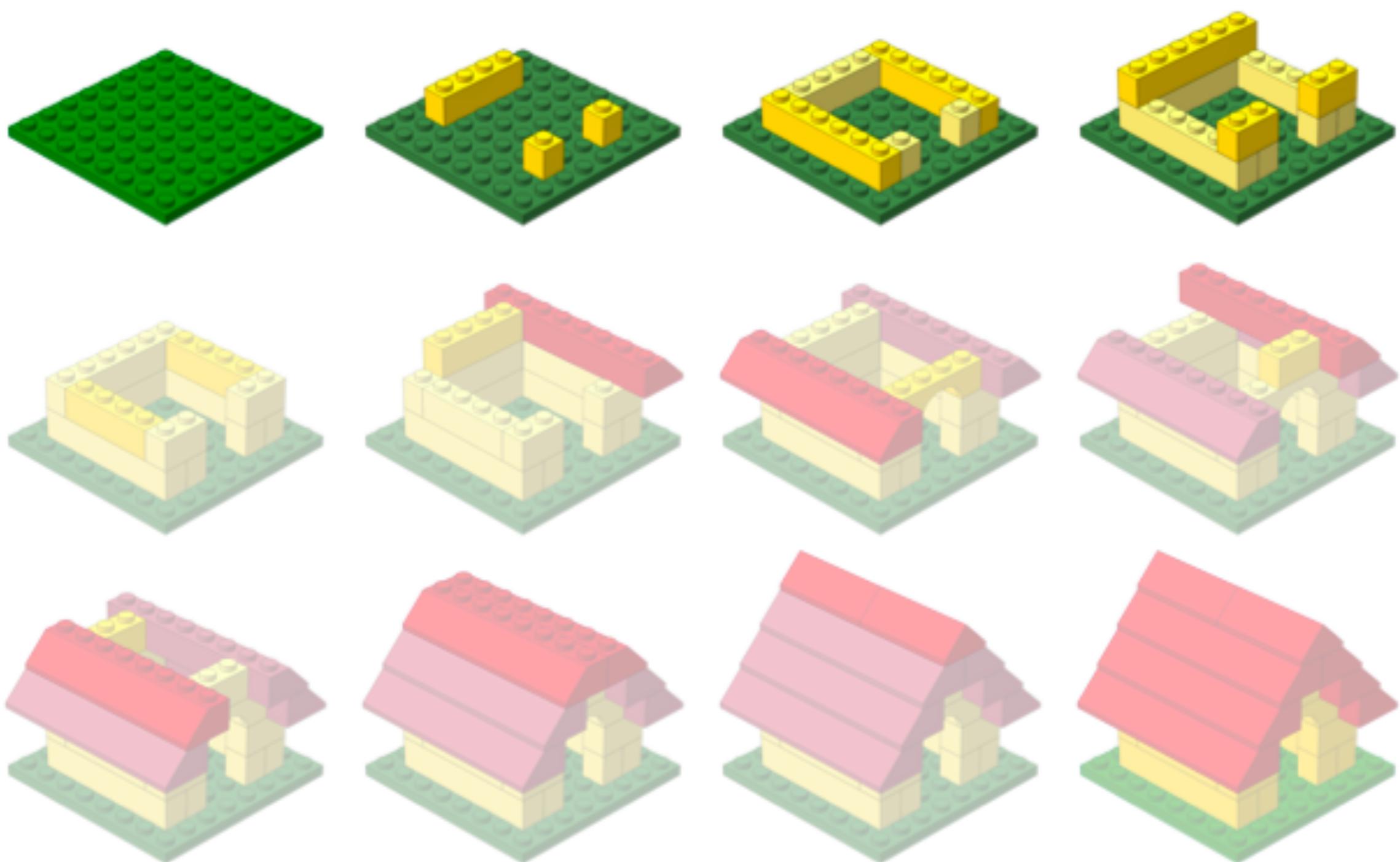
Training

- **Basic Data Skills**
 - Introduction to **command line** (Unix) and **high-performance computing**
 - Introduction to **R**
- **Advanced Analyses of NGS Data**
 - Bulk RNA-seq
 - Single-cell RNA-seq
 - ChIP-seq
- Monthly, short workshops on various bioinformatics topics

Consulting

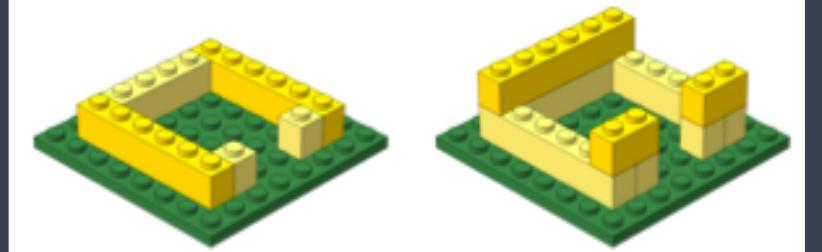
- **Transcriptomics:** RNA-seq, small RNA-seq, scRNA-Seq
- **Epigenetics:** ChIP-seq, genome-wide methylation, ATAC-Seq
- **DNA Variation:** WGS, resequencing, exome-seq and CNV studies
- **Functional enrichment** analysis
- **Exp. design help & grant support**

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

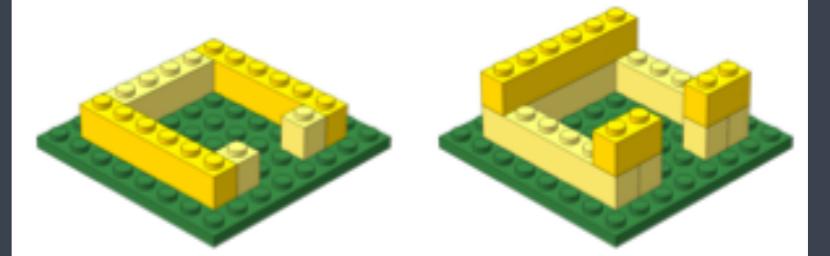
<https://cran.r-project.org/>



- 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



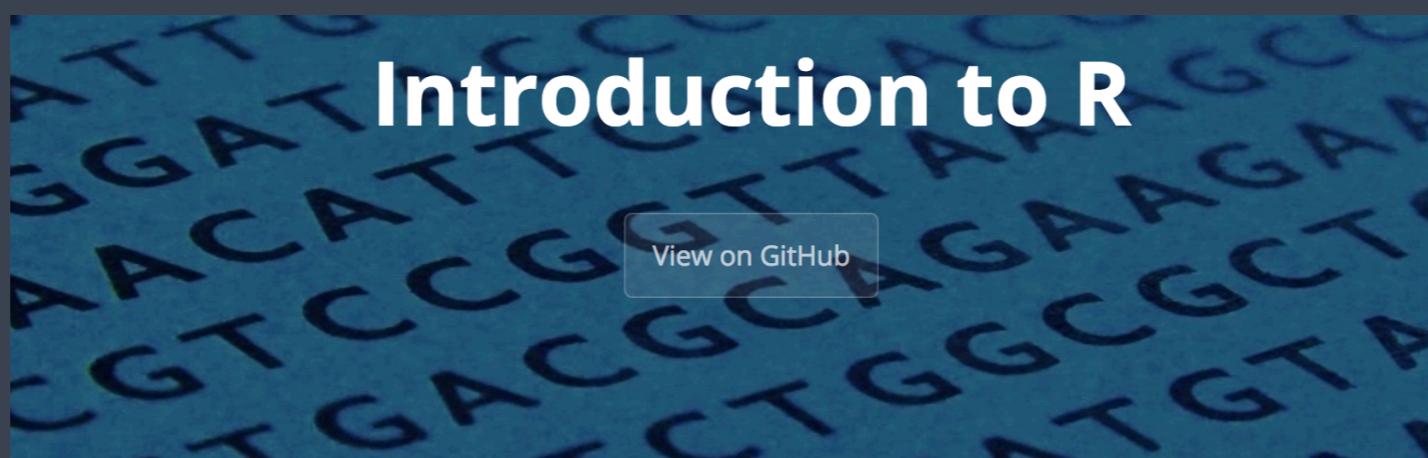
- Comfortably use RStudio (a graphical interface for R)
 - Fluently interact with R using RStudio
 - Become familiar with R syntax
 - Understand data structures in R
 - Inspect and manipulate data structures
 - Install packages and use functions in R
- ✓ Visualize data using *ggplot2*
- ✓ Utilize pipes,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

Logistics

Course webpage

<https://tinyurl.com/bph-intro-r>

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.

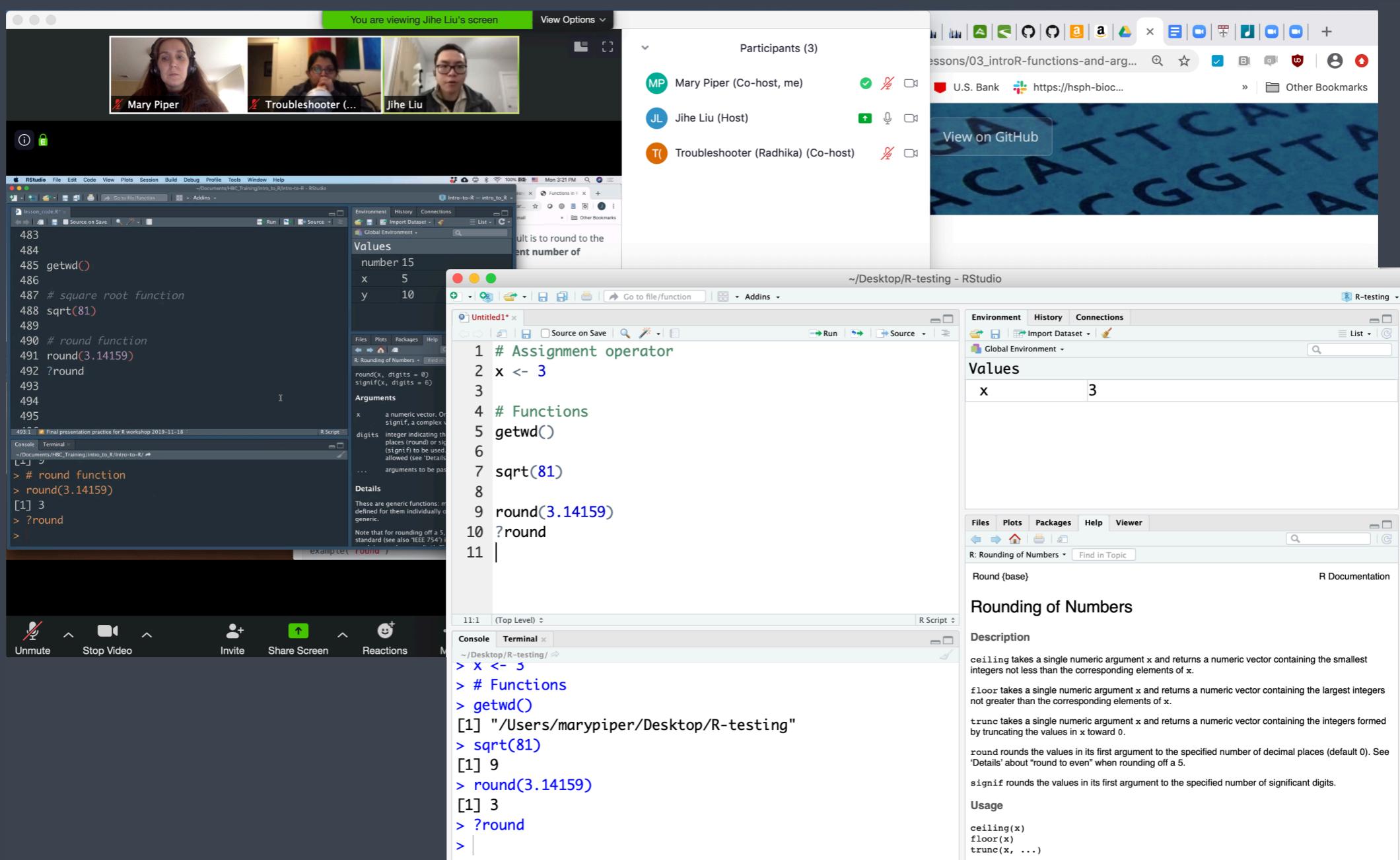
The R syntax

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.

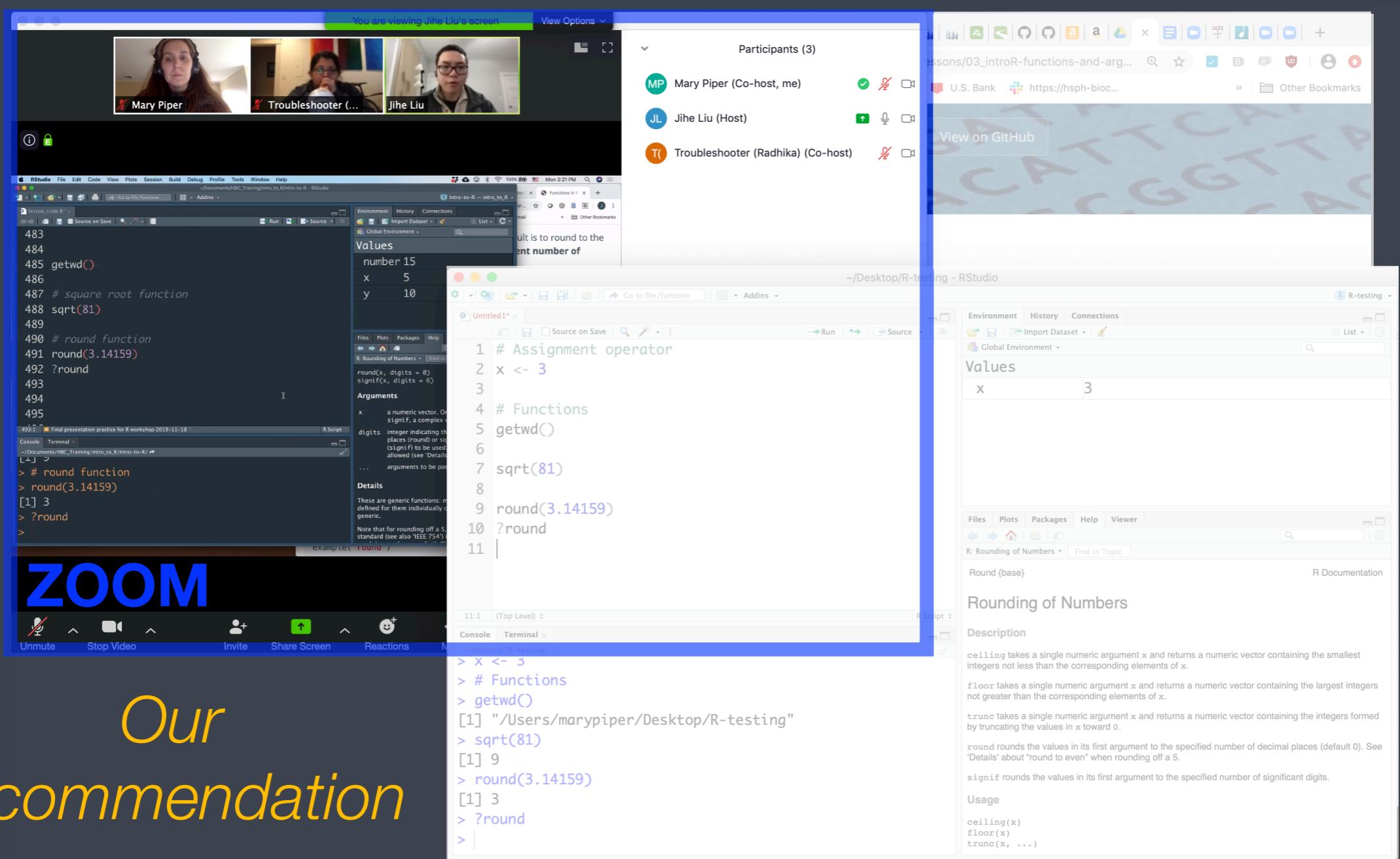
Below is an example script highlighting the many different “parts of speech” for R (syntax):

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

Single screen & 3 windows?

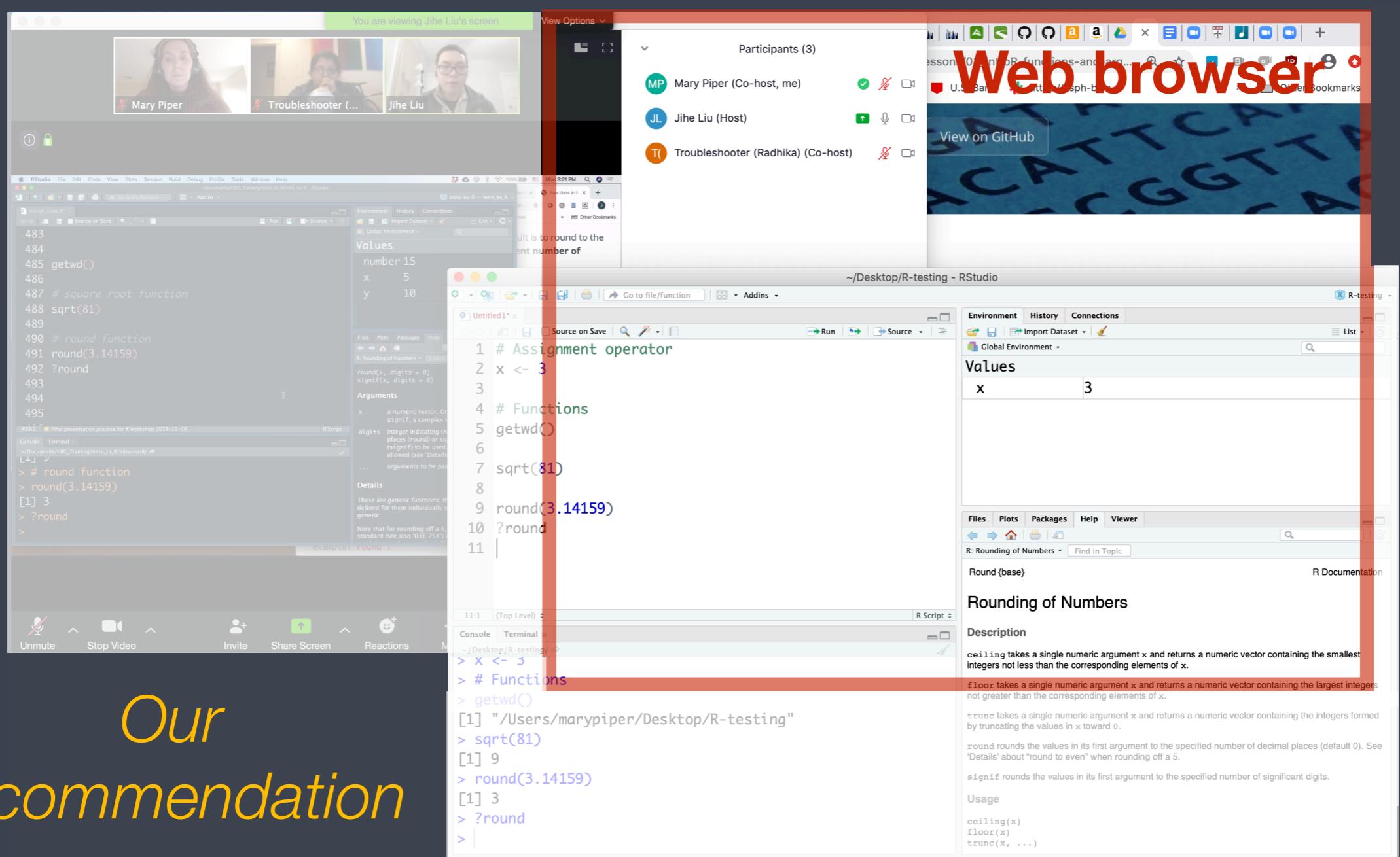


Single screen & 3 windows?



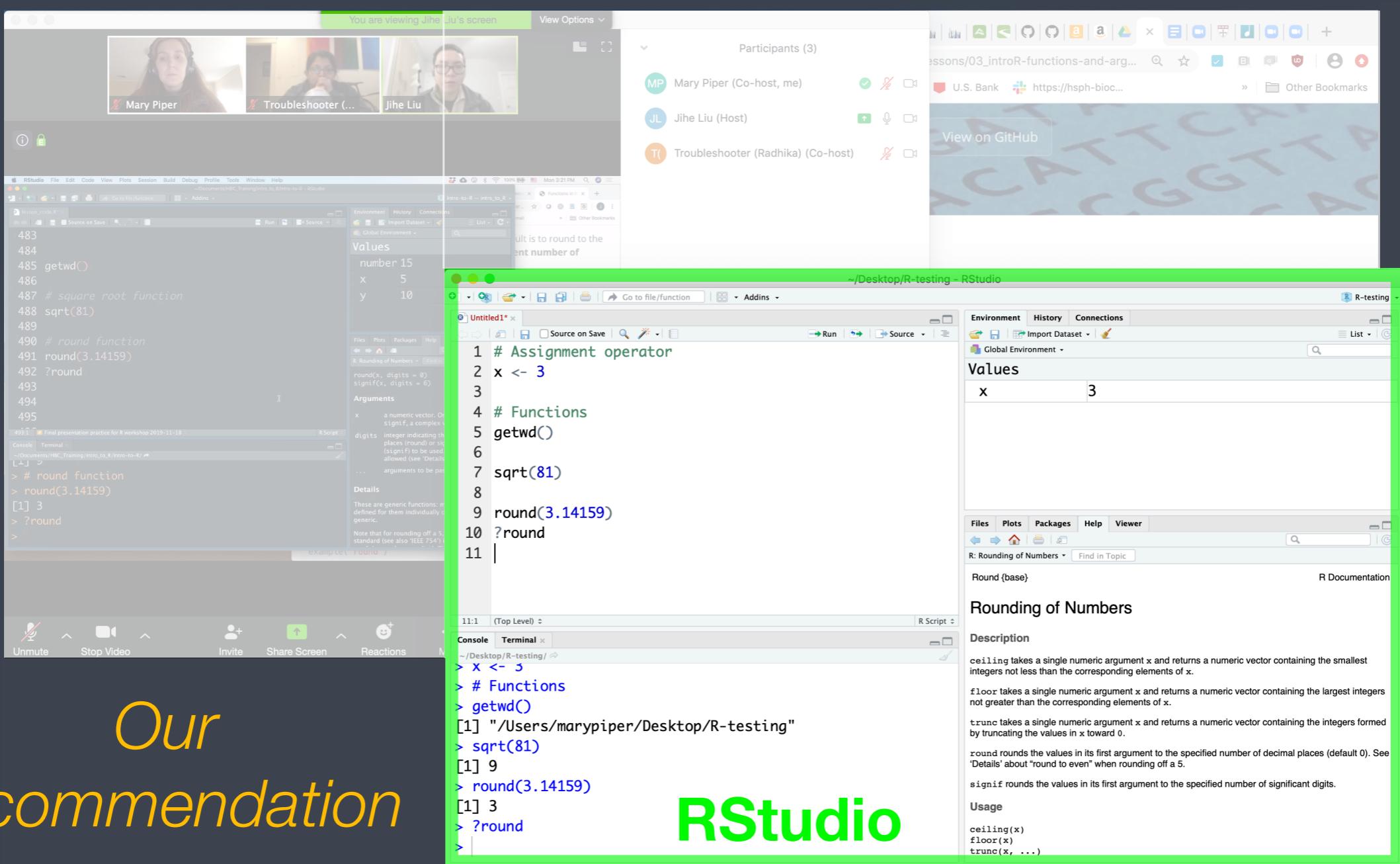
*Our
recommendation*

Single screen & 3 windows?

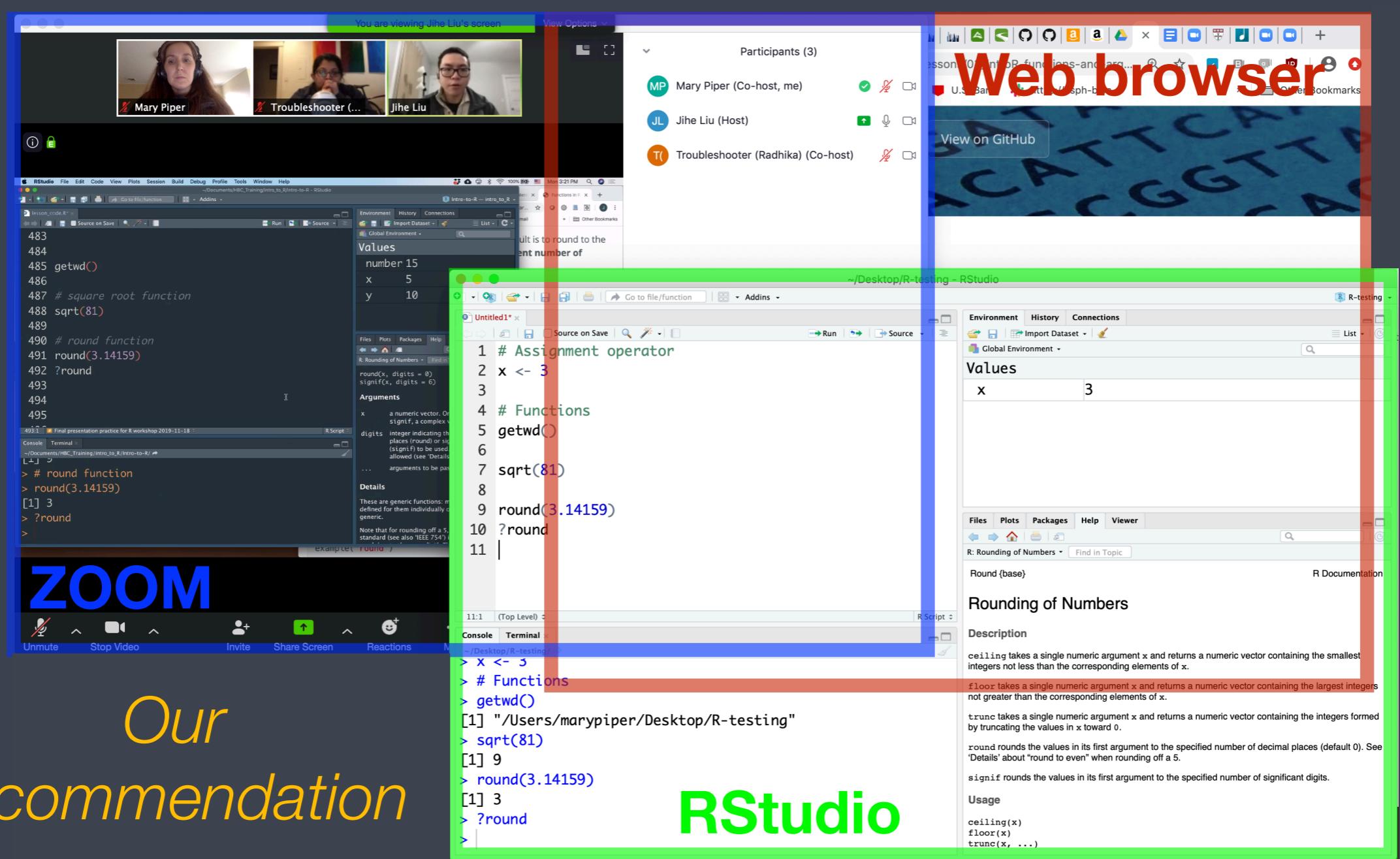


*Our
recommendation*

Single screen & 3 windows?



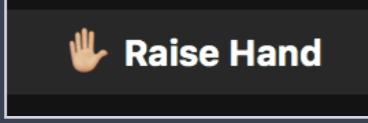
Single screen & 3 windows?

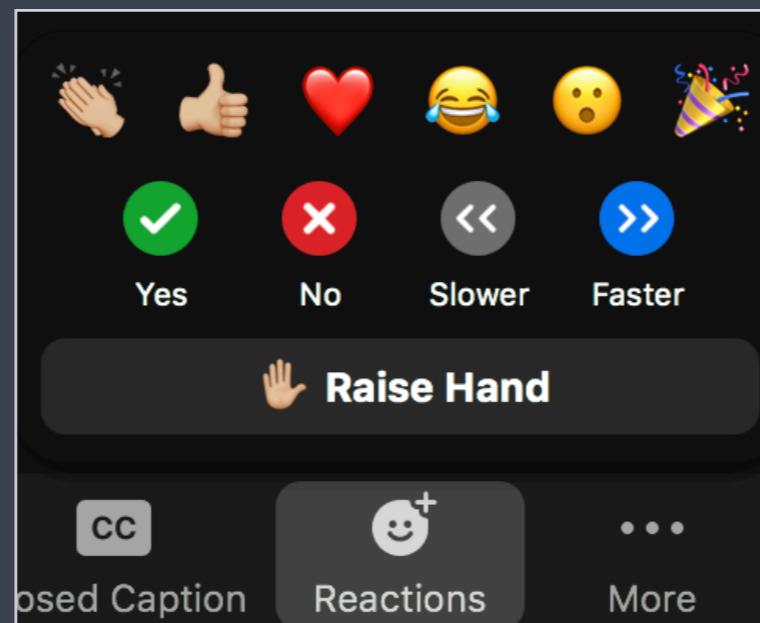


Odds and Ends

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

Odds and Ends

- ❖ 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**)
 - Type “X” in chat = "disagree", “I’m not ready” (equivalent to a **red post-it**)



Odds and Ends

- ❖ 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**)
 - Type “X” in chat = "disagree", "I'm not ready" (equivalent to a **red post-it**)
- ❖ Questions for the presenter?
 - Post the question in the Chat window OR
 - Unmute yourself and ask

Odds and Ends

- ❖ 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**)
 - Type "X" in chat = "disagree", "I'm not ready" (equivalent to a **red post-it**)
- ❖ Questions for the presenter?
 - Post the question in the Chat window OR
 - Unmute yourself and ask
- ❖ Technical difficulties with R or RStudio?
 - Start a private chat with the *instructor who is not presenting*.

Contact us!

HBC training team: hbctraining@hsph.harvard.edu

HBC consulting: bioinformatics@hsph.harvard.edu

Twitter

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