

# Module 1: Data Wrangling with R

Jeffrey Woo

MSDS Program, University of Virginia

# Welcome

- You can download this set of slides. Find with the materials for the live session in Module 1.
- Remind me to record the live session!
- Recommended: put yourself on mute unless you want to speak.
- There is a “raise hand” button for you. Click on “Reactions” in the panel at the bottom.

# Agenda

- Q&A
- Some comments on Module 1.
- Small group discussion of guided question set
- Large group discussion of guided question set plus other questions that pop up

# Q&A

- Any questions?
- Logistical or from module 1?

# Base R vs Tidyverse

Data wrangling operations can be performed using either approach. Use whichever approach (or even a combination) that you are comfortable with. You should be aware of both approaches.

# Base R vs Tidyverse

I am not aware of any study done to see if one approach is easier to learn than another. A lot of anecdotal testimonies are out on the world wide web and my conclusion is: people find the one they learn first easier.

# Base R vs Tidyverse

tidyverse has received a lot of active promotion through RStudio, which is a commercial entity with full-time employees. R is still an open-source project where contributors work in their free time.

## Small Group Discussion

- Materials can be found under Module 1 Live session.
- Have the guided question set and corresponding data set open.
- Have R open.
- Recommended: have easy access to your notes, textbook, as well as the tutorial.
- You can see who your group members are. As well as some roles you will have in your small group. Roles will rotate each session.
- You will be evaluating your group members based on how prepared they are and their contribution to the small group discussion.



# Large Group Discussion

## Comments about Data Wrangling

- Don't feel you have to memorize these functions.
- Refer to old notes / examples.
- Use documentation.
- Use google.
- I have definitely not covered everything, but hopefully you eventually get comfortable learning on your own.

## What is due...

- Homework 1, due Sep 5 (Mondays after the live session)
- Submit via assignments
- Please check the guidelines for homeworks.
- Get prepared for Module 2 and next Tuesday's live session.
- Continue reviewing Module A.

**WARNING:** Collab does not work well when you open multiple tabs. If you upload your HW with multiple tabs open, your submission is likely to screw up! You can always check if the submission went through correctly, and you can resubmit if needed.

## Getting help...

- Office hours on Zoom: Mondays and Thursdays (GTA), 7:30 to 8:30pm.
- Discussion forums.
- Email is to be used for more personal questions.
- We can also set up time to meet on Zoom if a discussion is needed.