

WORKSHOP MATERIALS:

github.com/jrogol/OverDRIVE2020



Advancement

THINKING REPRODUCIBLY THE WORKSHOP

James Rogol

APRA OverDRIVE/

23 March 2020



Advancement

re·pro·duc·i·bil·i·ty, *noun*



- Repeatable
- Identical
- Shareable

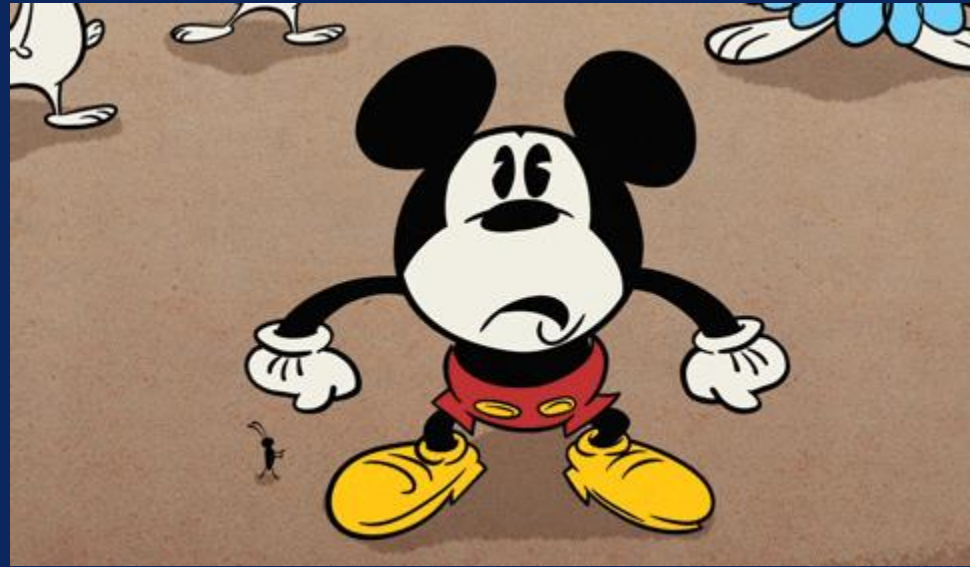
AGENDA

- Form Groups (5')



AGENDA

- Form Groups (5')
- Brainstorm Ideas (10')



AGENDA

- Form Groups (5')
- Brainstorm Ideas (10')
- Brainstorm Solutions (15')



AGENDA

- Form Groups (5')
- Brainstorm Ideas (10')
- Brainstorm Solutions (15')
- Share Results (10')
- Exercises in R

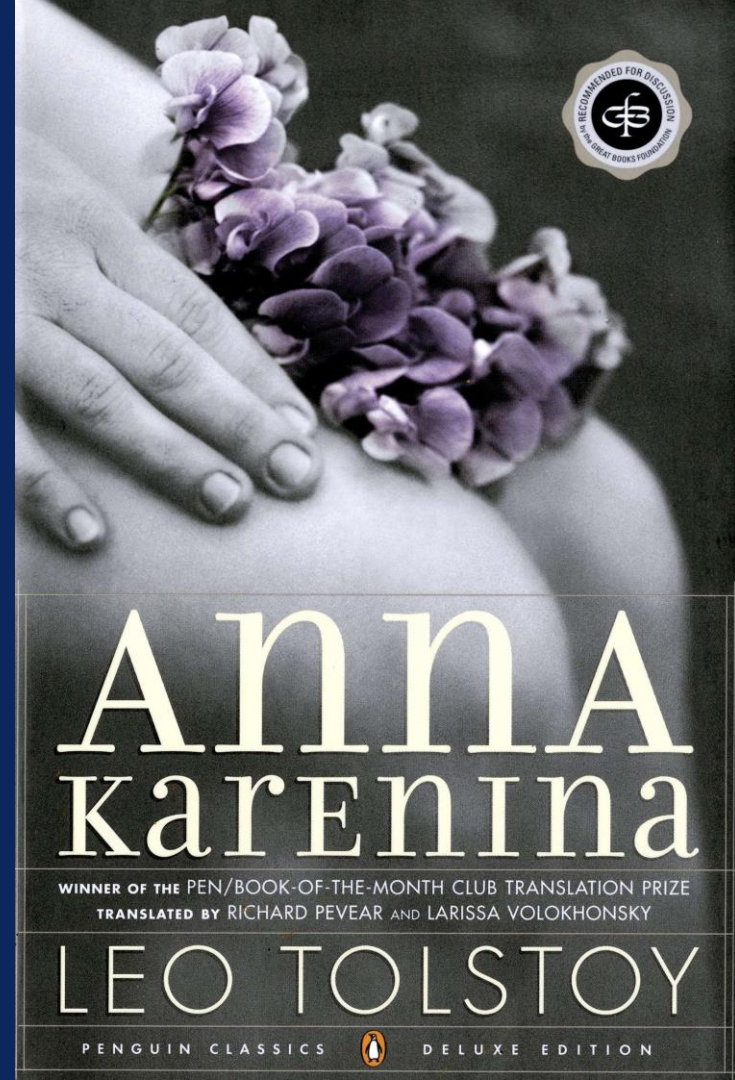


FORM GROUPS

.....

All **reproducible projects** are alike;
Each **non-reproducible** one **is**
not in its own way.

.....



***PAST
EXPERIENCE?***

EXAMPLES

- Connecting multiple data sources
 - APIs, Flat Files, Databases
- Reports suddenly breaking
- Reverse engineering someone else's work
- Non-standardized coding

THINKING REPRODUCIBLY

- Slow Down
- Know Your Data
 - Inflection Points
 - Connections
- Design
- Abstract
- Document



YOUR TURN!

RESULTS?

EXERCISES



OverDRIVE2020-repo.Rproj

DATA ACCESS

```
10 # Set a working directory and load a .csv file
11 setwd("C:/Users/jr2ef/Documents/ProjectDirectory/Output/")
12
13 data <- read_csv("ratings.csv")
14 # This works just fine on my machine - what about yours?
15
16 # myScript.R is in the /ProjectDirectory folder. Why won't this work?
17 source("myScript.R")
18
19
20 # File Structure -----
21
22
23 # When the data is downloaded from the database, it's an "Output"
24 myData <- read_csv("Output/rawData.csv")
25
26 # But someone else might place the same file in a "Data" folder.
27 yourData <- read_csv("Data/rawData.csv")
```

.....

“If the first line of your #rstats script is

```
setwd("C:\\Users\\jenny\\path\\that\\only\\I\\have")
```

I will come into your lab and

SET YOUR COMPUTER ON FIRE.”

.....

Jenny Bryan, RStudio

(Former) Adjunct Professor, University of British Columbia

CONTACT



rogol@virginia.edu



[/in/jrogol](https://www.linkedin.com/in/jrogol)



[/jrogol](https://twitter.com/jrogol)



[/jrogol](https://github.com/jrogol)

REFERENCES

Brent Waller. "Benny's cousins are in town."

Penguins Running, Mickey Thinking, Sherlock Pondering and Moss Programming from giphy.

Leo Tolstoy. *Anna Karenina*. Translated by Richard Pevear and Larissa Volokhonsky. 2004.

Icons made by Freepik, Darius Dan, Good Ware, iconixar, Becris, Smashicons, ProSymbols Gregor Cresnar and Pixel Perfect from www.flaticon.com, licensed CC 3.0 BY.

Jenny Bryan. "Ode to the here package."

Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani. *An Introduction to Statistical Learning (with Applications in R)*. Springer Press

Emily Robinson and Jacqueline Nolis. Build a Career in Data Science. April 2020.