RStudio

rstudio::conf(2022) WORKSHOP

What They Forgot to Teach You About R

SHANNON MCCLINTOCK PILEGGI

JENNY BRYAN

E. DAVID AJA

July 25-28th | WASHINGTON, D. C.

Register at rstd.io/conf

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

To view a copy of this license, visit http://creativecommons.org/licenses/by-sa/4.0/

Please complete the workshop survey!

rstd.io/conf-workshop-survey

Your feedback is crucial! Data from the survey informs curriculum and format decisions for future conf workshops and we really appreciate you taking the time to provide it.

Everyone is encouraged to open issues here:

rstd.io/wtf-2022-rsc
github.com/rstudio-conf-2022/wtf-rstats/issues

Record glitches, gotchas,
good sidebar discussions, etc.
to address now or later.

rstd.io/wtf-2022-rsc

Deep
Thoughts

Trevor A. Branch
@TrevorABranch

Follow

My rule of thumb: every analysis you do on a dataset will have to be redone 10–15 times before publication. Plan accordingly. #Rstats

expect to iterate

Reason to iterate #1:
Get it right! Keep it right!

New data?

New understanding of data?

Reason to iterate #2:
Refine and Extend

Make your code more

- ✔ Readable
- ✔ Efficient
- ✔ Resilient
- ✔ General

Deep
Thoughts

machine readable ***and*** human readable

plan to skim, iterate, tinker, remix, ...

yes, workflow is "a thing" and it's OK to care

embrace incrementalism

don't speculate & dither, run the experiment

if it hurts, do it more often

Workshop Instructors

Shannon Pileggi

@PipingHotData
@shannonpileggi

The Prostate Cancer Clinical Trials Consortium

Jenny Bryan

@JennyBryan
@jennybc

RStudio

David Aja

@peeltothepithy
@edavidaja

RStudio

Workshop Assistants

Andrie de Vries

@RevoAndrie

@andrie

RStudio

Amanda Gadrow

@ajmcoqui

@ajmcoqui

RStudio

Ryan Johnson

@ryjohnson09

@ryjohnson09

RStudio

Crystal Lewis

@Cghlewis

@Cghlewis

Freelance Data Management Consultant

Daniel Sjoberg

@statistishdan

@ddsjoberg

Memorial Sloan Kettering Cancer Center