Rotman

TD Management Data and Analytics Lab

Git is an active, open source, version control system

-developed by Linus Torvalds

GitHub is a hosting service for Git

-subsidiary of Microsoft since 2018

-200 million repositories

-83 million developers

Name	2015	2017	2018	2022
Git	69.3%	69.2%	87.2%	93.9%
Subversion	36.9%	9.1%	16.1%	5.2%
TFVC	12.2%	7.3%	10.9%	[ii]
Mercurial	7.9%	1.9%	3.6%	1.1%
CVS	4.2%	[ii]	[ii]	[ii]
Perforce	3.3%	[ii]	[ii]	[ii]
VSS	[ii]	0.6%	[ii]	[ii]
ClearCase	[ii]	0.4%	[ii]	[ii]
Zip file backups	[ii]	2.0%	7.9%	[ii]
Raw network sharing	[ii]	1.7%	7.9%	[ii]
Other	5.8%	3.0%	[ii]	[ii]
None	9.3%	4.8%	4.8%	4.3%

1. Code Sharing is Now Standard Practice at Journals
2. Different Filenames Break Dependencies
3. You Can Explore New Ideas Without Fear
4. You Can Fix Bugs (and other people can too!)

Rotman

Code Sharing is Now Standard Practice (and Good Practice!)

Management Science Policy for Data and Code Disclosure

Effective June 1, 2019

Download **Data and Code Disclosure Form**

A fundamental principle of the scientific method is replication: the validity of a research finding requires that it can be reproduced by other researchers. The intent of the Data and Code Disclosure policy is to assure the availability of the material necessary to replicate the research published in the journal.

A secondary benefit of this policy is to advance the research in the fields covered by the journal. Inevitably, the sharing of data and codes will be of value to the relevant research community, allowing them to leverage this prior work in their own pursuits. This sharing should increase the rate of scientific progress and impact.

So you will need <u>some</u> platform to distribute your code. Start now, not later.
This is isn't just "red-tape" to address replication. It increases your impact.

"Filename Method" Break Dependencies

"Filename Method" Break Dependencies

TD-MDAL

"Filename Method" Breaks Dependencies

Using Filename Method

calc_utilities.R
calc_utilities2.R
calc_utilities3.R
calc_utilities4.R
calc_utilities5.R
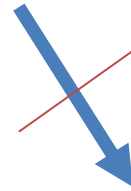
simulate_choices.R   calc_loglikelihood.R

Sometimes will give an error but can lead to "hidden" bugs
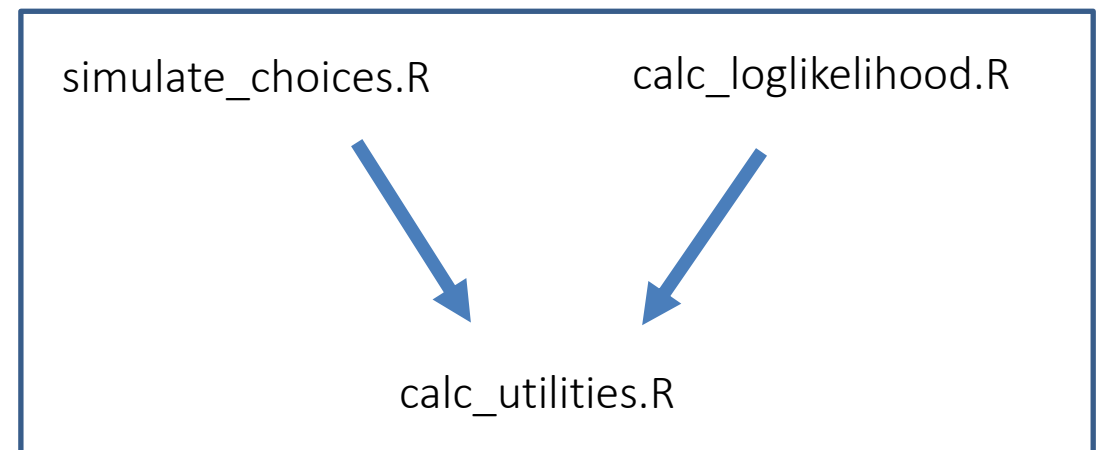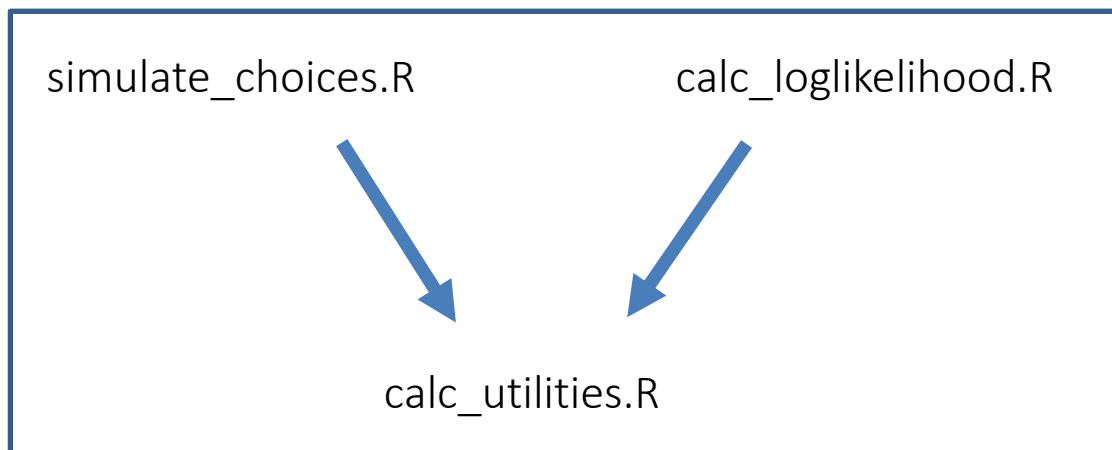
calc_utilities2.RR

Using Git

Branch 1

simulate_choices.R   calc_loglikelihood.R

calc_utilities.R

Branch 2

simulate_choices.R   calc_loglikelihood.R

calc_utilities.R

Rotman

You Can Explore New Ideas Without Fear

TD-MDAL

Rotman

You Can Fix Bugs (and others can too!)

Why Use GitHub? Or "Mistakes I Have Made"

1. Code Sharing is Now Standard Practice at Journals
2. Different Filenames Break Dependencies
3. You Can Explore New Ideas Without Fear
4. You Can Fix Bugs (and other people can too!)

Version control matters most at the END of a project
     - it is an investment, but one that will pay off

TD Management Data and Analytics Lab

Rotman

Dates:
Thursday, January 26, 2023 – Finance theme
Tuesday, March 21, 2023 – Mixed research theme

Highlight TD MDAL research grant recipients and their projects in the areas of data analytics, machine learning and artificial intelligence.