Data Tips, Part 2

Gina Nichols

2020 Fall DataFEWsion Sharing Series

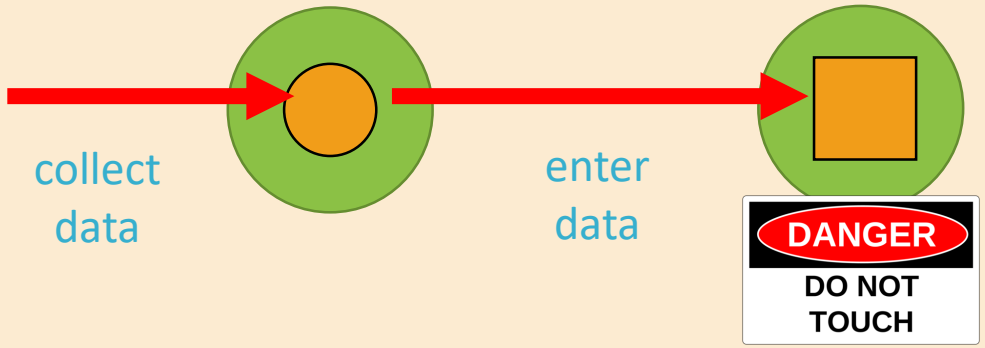
get some data
manuscript, glory, graduation

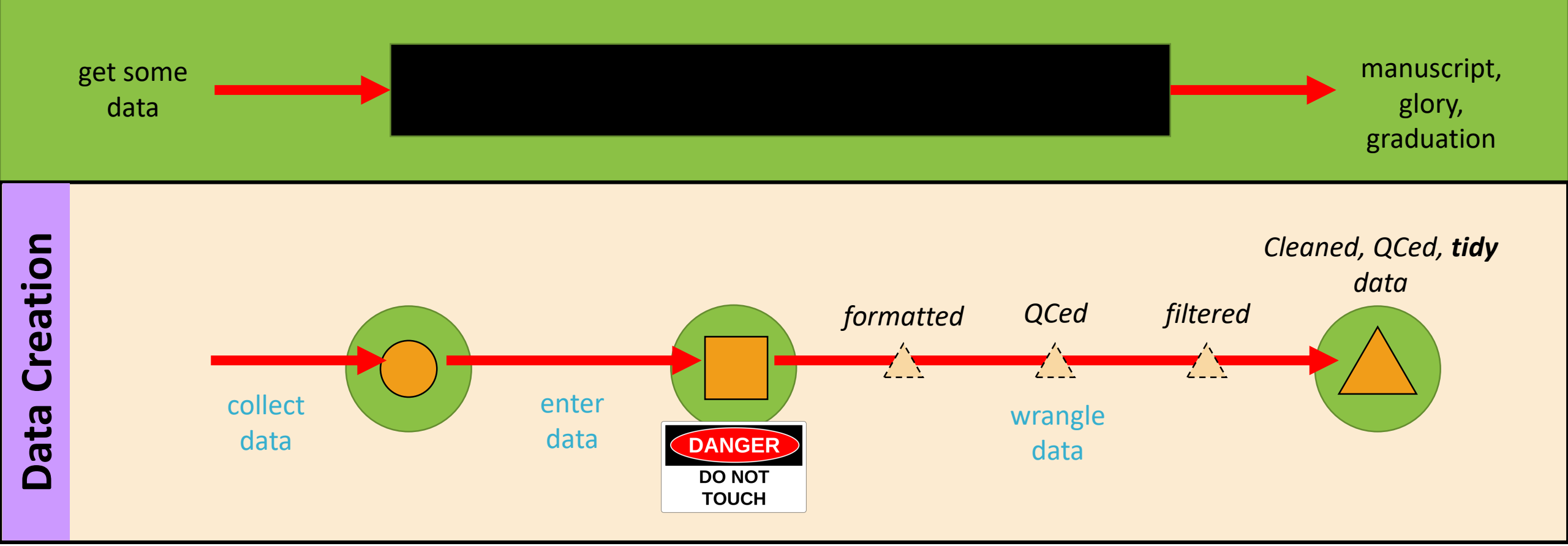
get some data
manuscript, glory, graduation
Data Creation
collect data

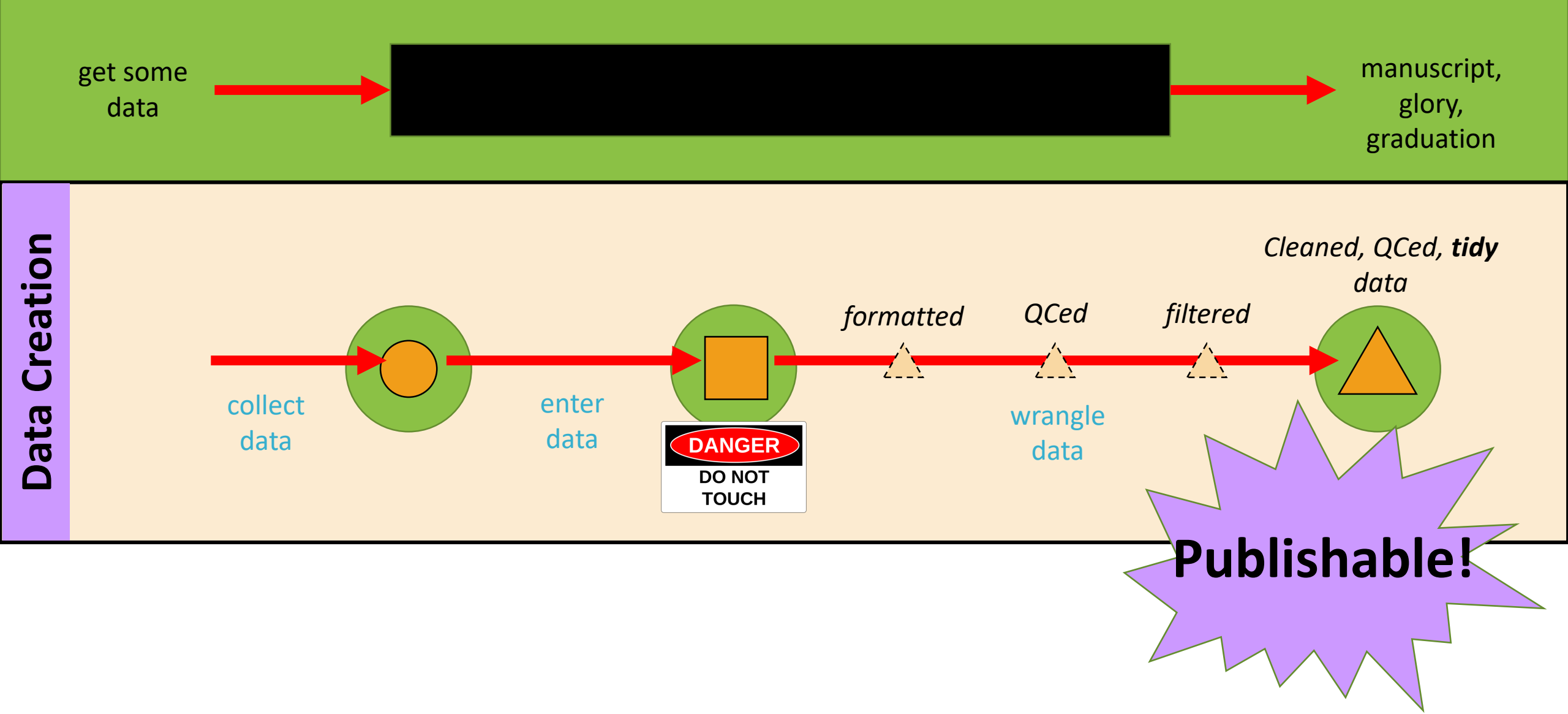
get some data

manuscript, glory, graduation

Data Creation




get some data
manuscript, glory, graduation
Data Creation
collect data
enter data
DANGER
DO NOT TOUCH
formatted
QCed
filtered
wrangle data
Cleaned, QCed, tidy data


get some data
manuscript, glory, graduation
Data Creation
collect data
enter data
DANGER
DO NOT TOUCH
formatted
QCed
filtered
wrangle data
Cleaned, QCed, tidy data
Publishable!

Publishable!
nstr.iastate.libguides.com/datashare
Toggl
Climate Outlooks |...
COVID map
Restore pages?
Chrome didn't shut down correctly.
Restore
Open startup pages
ary
CTIONS
BORROWING
HELP & SERVICES
OFF-CAMPUS ACCESS
LIBRARY GUIDES
U's Open Research Data Repository
Search this Guide
Search
Share is an open access repository for sharing and publishing research data created by Iowa State University scholars and
FAQs
DataShare: Getting Started
DataShare: Step-by-step
Submission Form
File Formats
Data Documentation
File Name Recommendations
Data licensing options
Linking ORCiD
Terms of Use
Contact
THE OPEN DATA REPOSITORY OF
IOWA STATE UNIVERSITY
DataShare
We're working from home!
Please contact us by email:
datashare@iastate.edu
For more help visit the University Library's
Data Management Plan Guide
Share your research data with confidence
Iowa State University's DataShare is an open access repository for sharing, publishing, and archiving research data created by Iowa State University scholars and researchers.

Publishable!

Effect of cover crops on weed biomass and density in the US…

Dataset posted on 02.04.2020

Virginia Nichols

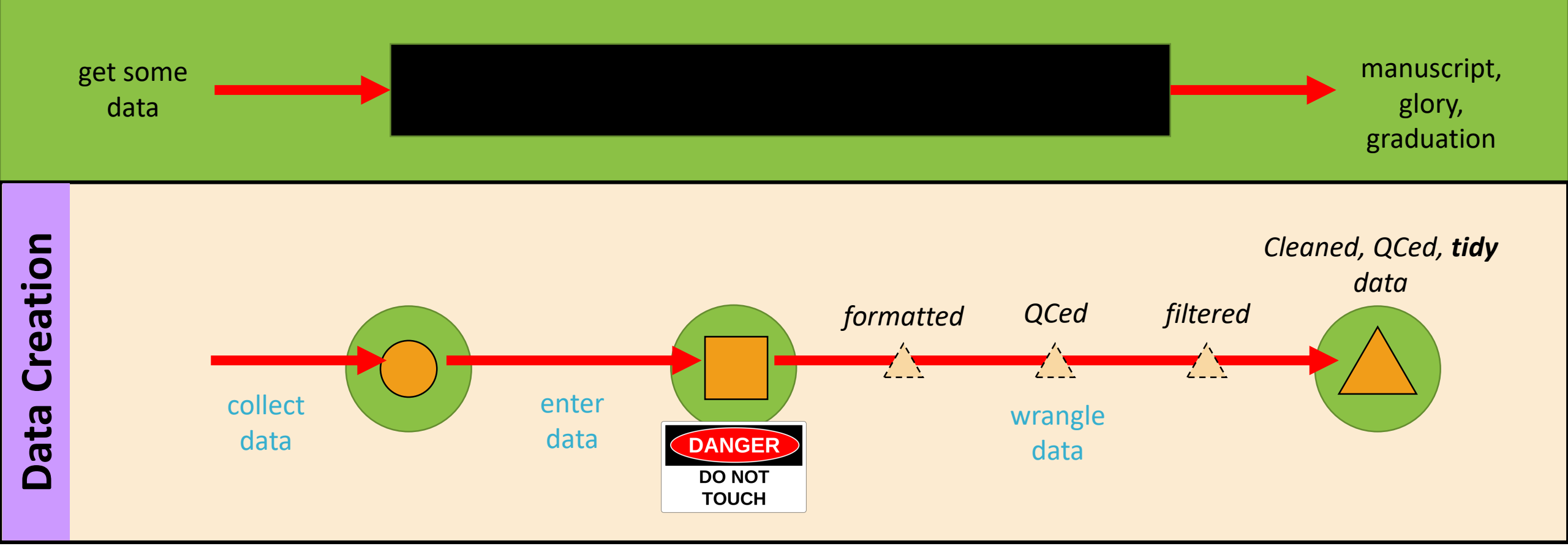
Effect of crop rotation on weed biomass and density literature…

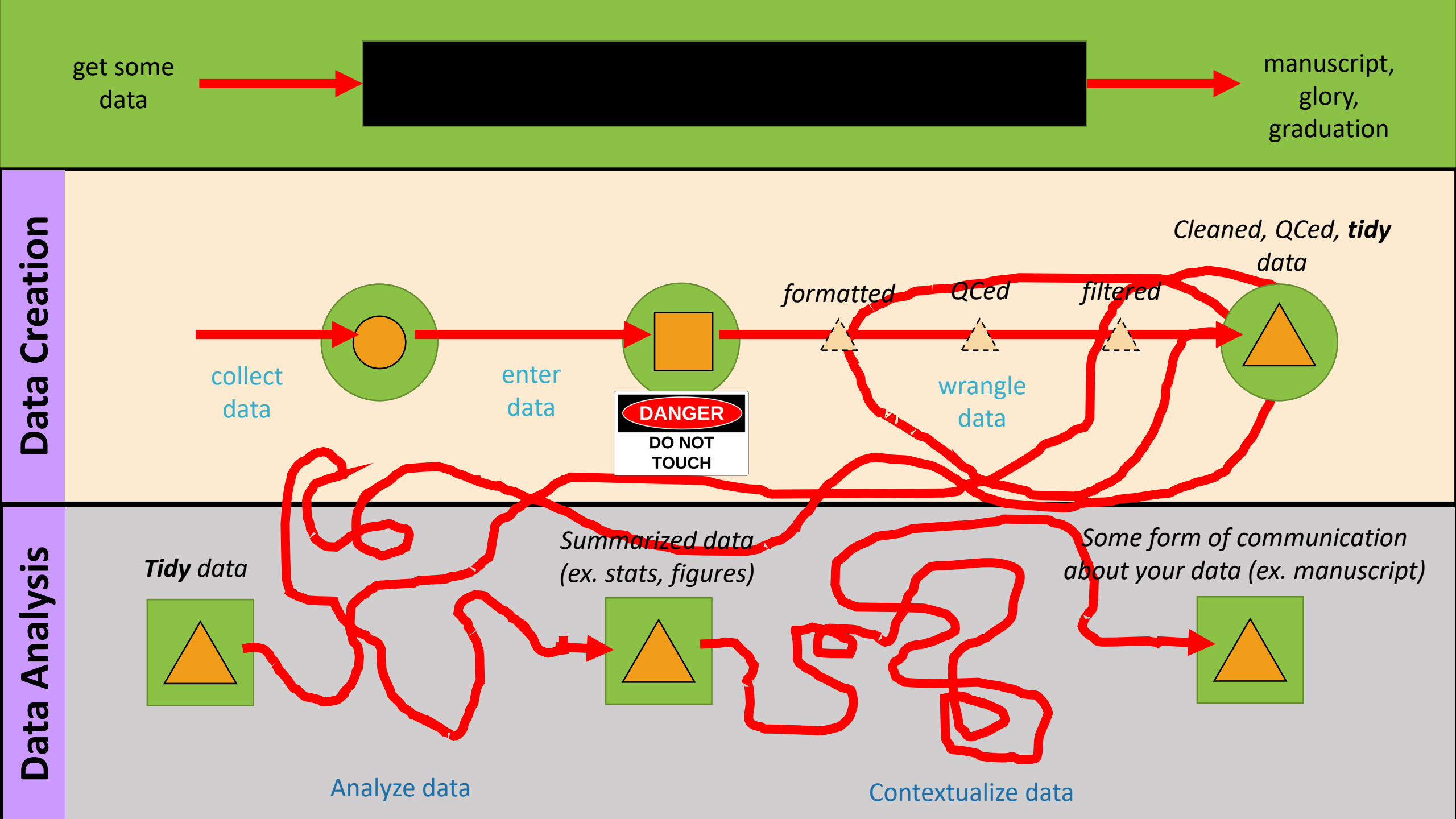
Dataset posted on 09.07.2019

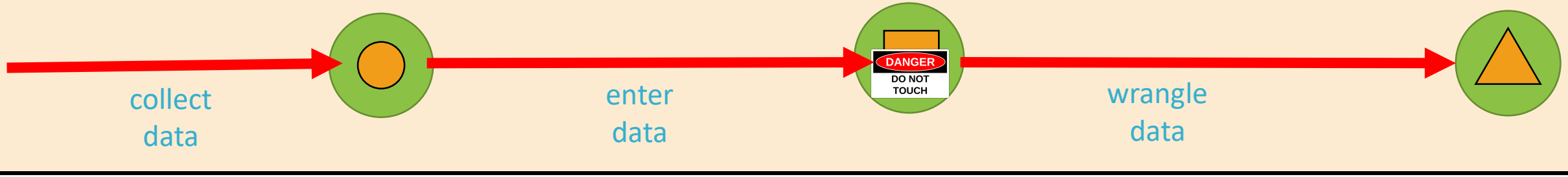
Virginia Nichols

Nichols, Virginia; Weisberger, David; Liebman, Matt (2019): Effect of crop rotation on weed biomass and density literature summary. Iowa State University. Dataset. https://doi.org/10.25380/iastate.7771010.v2

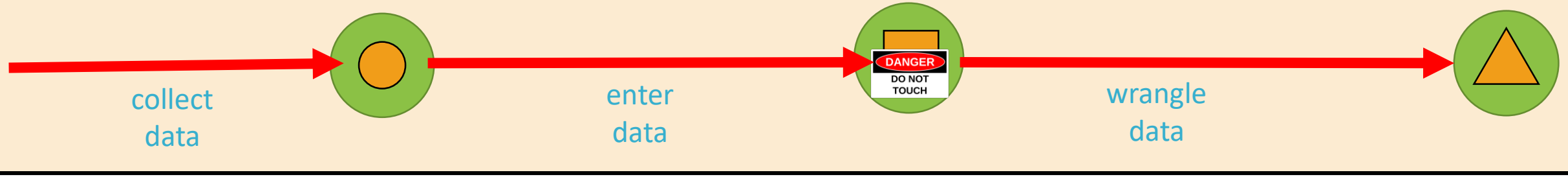
Nichols, Virginia; Basche, Andrea D.; Weisberger, David (2020): Effect of cover crops on weed biomass and density in the US Midwest Corn Belt meta-analysis dataset. Iowa State University. Dataset. https://doi.org/10.25380/iastate.11933214.v1

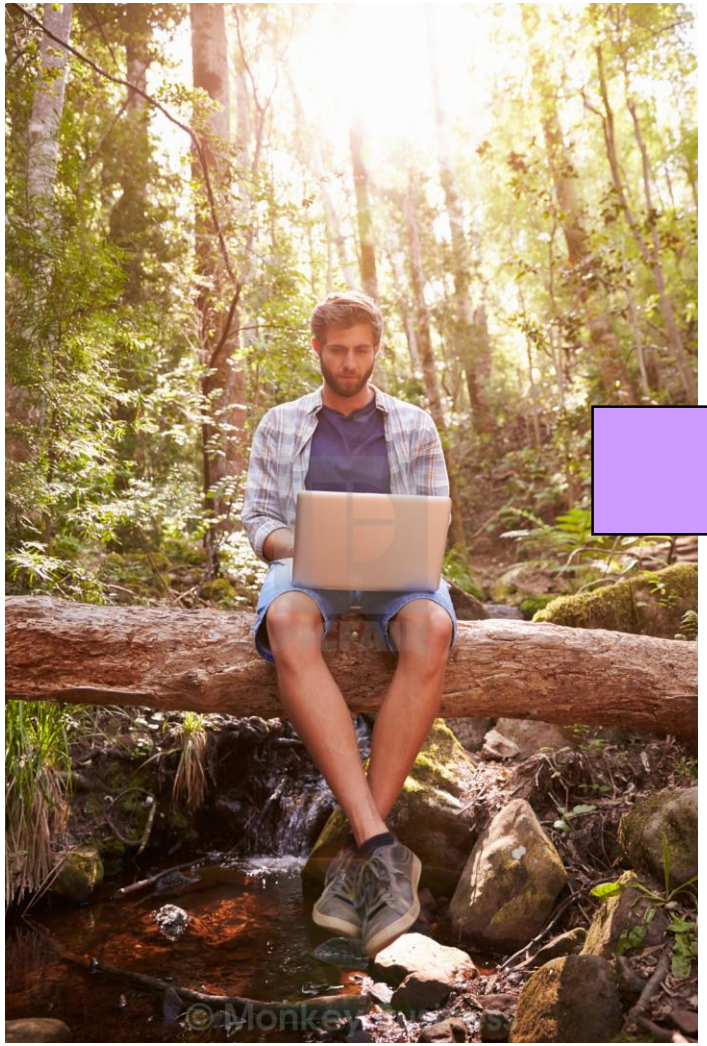

get some data
manuscript, glory, graduation
Data Creation
collect data
enter data
DANGER
DO NOT TOUCH
formatted
QCed
filtered
wrangle data
Cleaned, QCed, tidy data


get some data
manuscript, glory, graduation
Data Creation
Cleaned, QCed, tidy data
formatted
QCed
filtered
collect data
enter data
DANGER
DO NOT TOUCH
wrangle data
Data Analysis
Tidy data
Summarized data (ex. stats, figures)
Some form of communication about your data (ex. manuscript)
Analyze data
Contextualize data

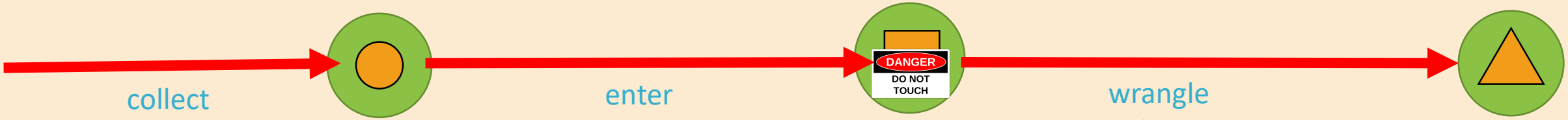


sighting_id	observer_id	location_id	species_id	date	visual_recognition	audio_recognition
1	492	1	AMRO	2016-05-11 19:01:00	yes	no
2	492	1	EABL	2016-05-11 19:53:00	yes	no
3	213	2	DOWO	2016-05-16 16:27:00	yes	yes
4	492	3	AMGO	2016-05-17 06:55:00	yes	no
5	213	2	CLSW	2016-06-03 21:43:00	yes	no
6	213	2	GHOW	2016-06-03 22:24:00	no	yes
7	492	1	EABL	2016-06-06 18:28:00	yes	no
8	213	1	AMRO	2016-06-17 14:43:00	yes	yes
9	213	1	GBHE	2016-06-17 15:07:00	yes	no
10	213	1	PUFI	2016-06-17 15:45:00	yes	no
11	213	1	EABL	2016-06-17 16:00:00	yes	no
12	492	3	AMGO	2016-06-18 07:42:00	yes	no





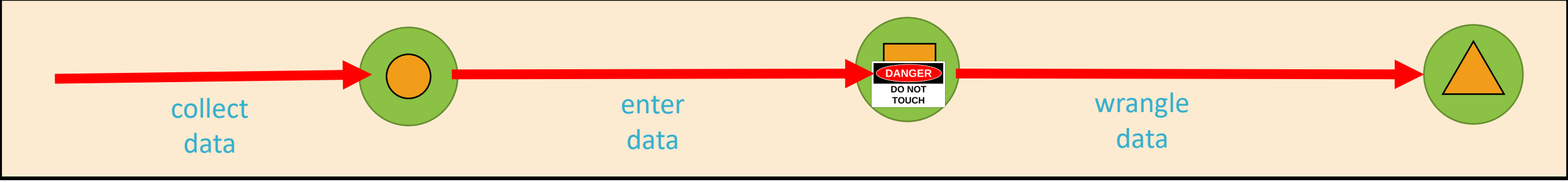
sighting_id	observer_id	location_id	species_id	date	visual_recognition	audio_recognition
1	492	1	AMRO	2016-05-11 19:01:00	yes	no
2	492	1	EABL	2016-05-11 19:53:00	yes	no
3	213	2	DOWO	2016-05-16 16:27:00	yes	yes
4	492	3	AMGO	2016-05-17 06:55:00	yes	no
5	213	2	CLSW	2016-06-03 21:43:00	yes	no
6	213	2	GHOW	2016-06-03 22:24:00	no	yes
7	492	1	EABL	2016-06-06 18:28:00	yes	no
8	213	1	AMRO	2016-06-17 14:43:00	yes	yes
9	213	1	GBHE	2016-06-17 15:07:00	yes	no
10	213	1	PUFI	2016-06-17 15:45:00	yes	no
11	213	1	EABL	2016-06-17 16:00:00	yes	no
12	492	3	AMGO	2016-06-18 07:42:00	yes	no



sighting_id	observer_id	location_id	species_id	date	visual_recognition	audio_recognition
1	492	1	AMRO	2016-05-11 19:01:00	yes	no
2	492	1	EABL	2016-05-11 19:53:00	yes	no
3	213	2	DOWO	2016-05-16 16:27:00	yes	yes
4	492	3	AMGO	2016-05-17 06:55:00	yes	no
5	213	2	CLSW	2016-06-03 21:43:00	yes	no
6	213	2	GHOW	2016-06-03 22:24:00	no	yes
7	492	1	EABL	2016-06-06 18:28:00	yes	no
8	213	1	AMRO	2016-06-17 14:43:00	yes	yes
9	213	1	GBHE	2016-06-17 15:07:00	yes	no
10	213	1	PUFI	2016-06-17 15:45:00	yes	no
11	213	1	EABL	2016-06-17 16:00:00	yes	no
12	492	3	AMGO	2016-06-18 07:42:00	yes	no

name	date	time	species	visual recognition	audio recognition

name	date	time	species	visual recognition	audio recognition

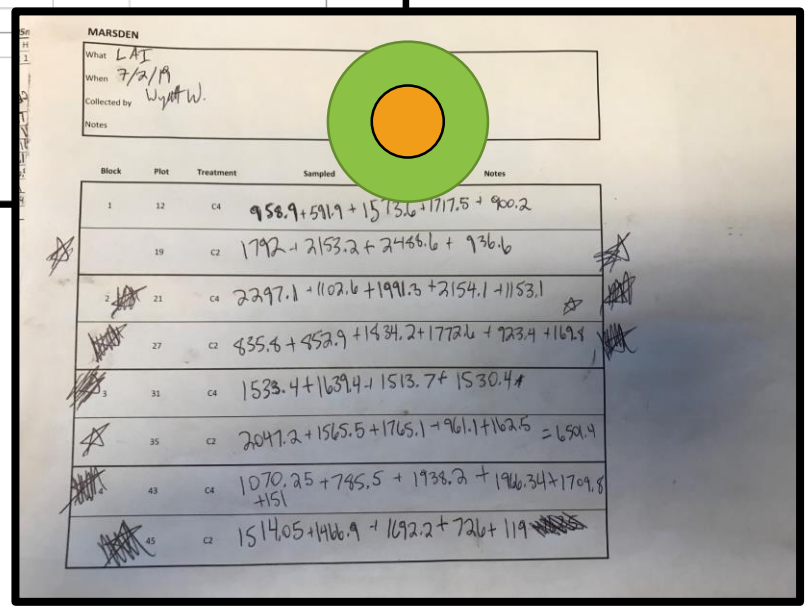


Page ___ of ___

Observer name:
Location of observations:
Date of observations:
Obs. notes:

Electronically entered by:
Date electronically entered:
Notes:

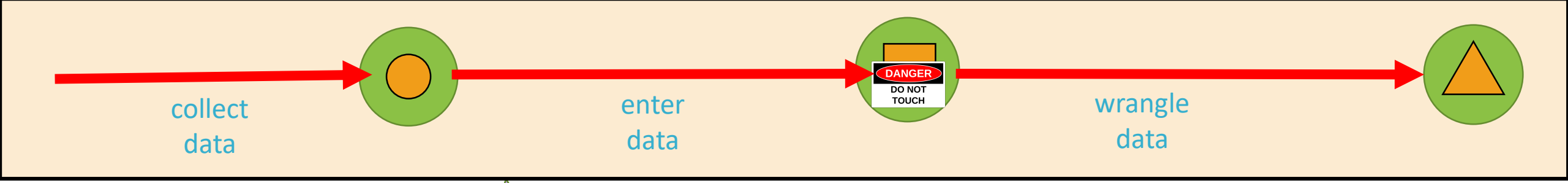
obs_nu	time	species	visual recognition	audio recognition	notes
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

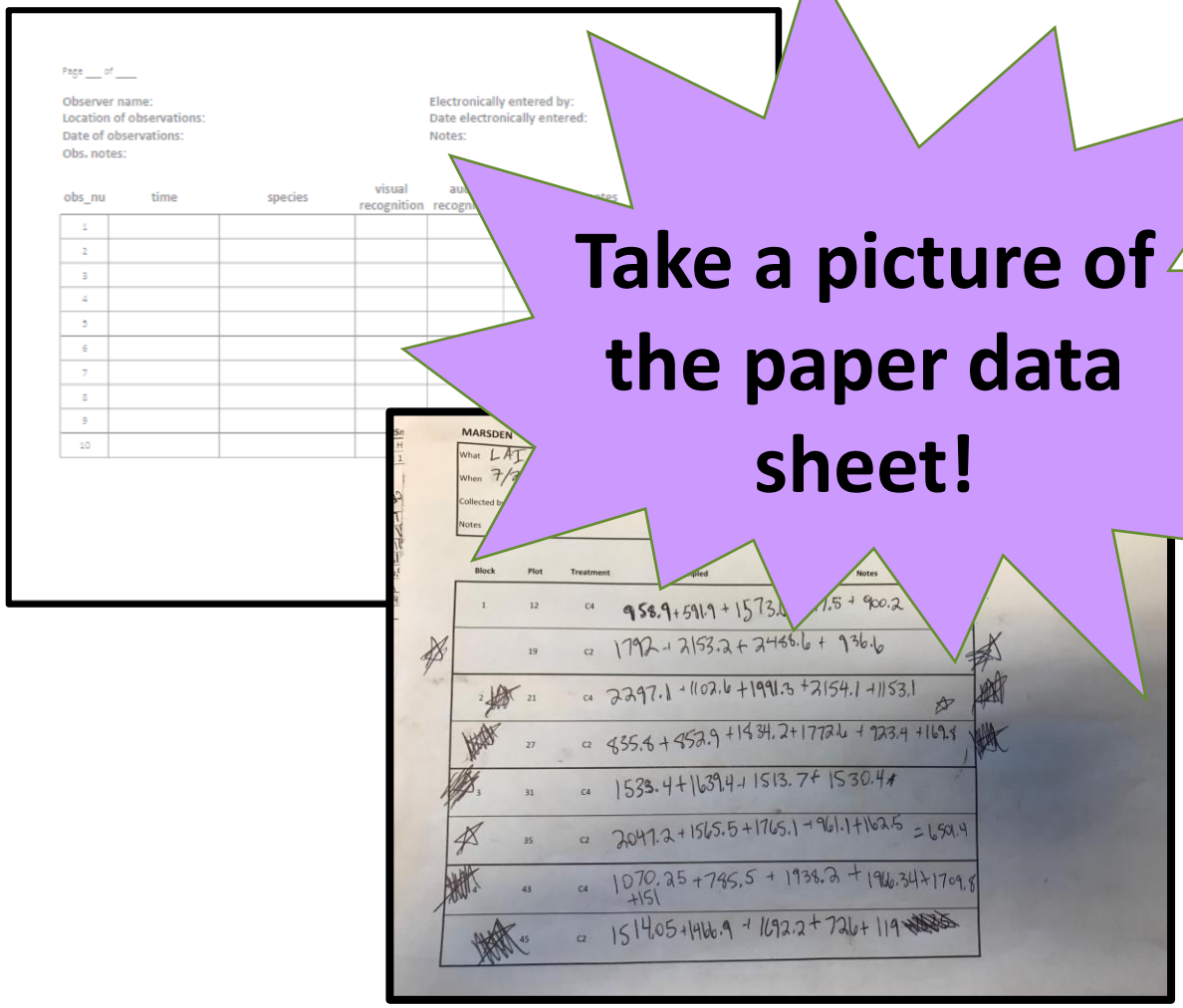


MARSDEN

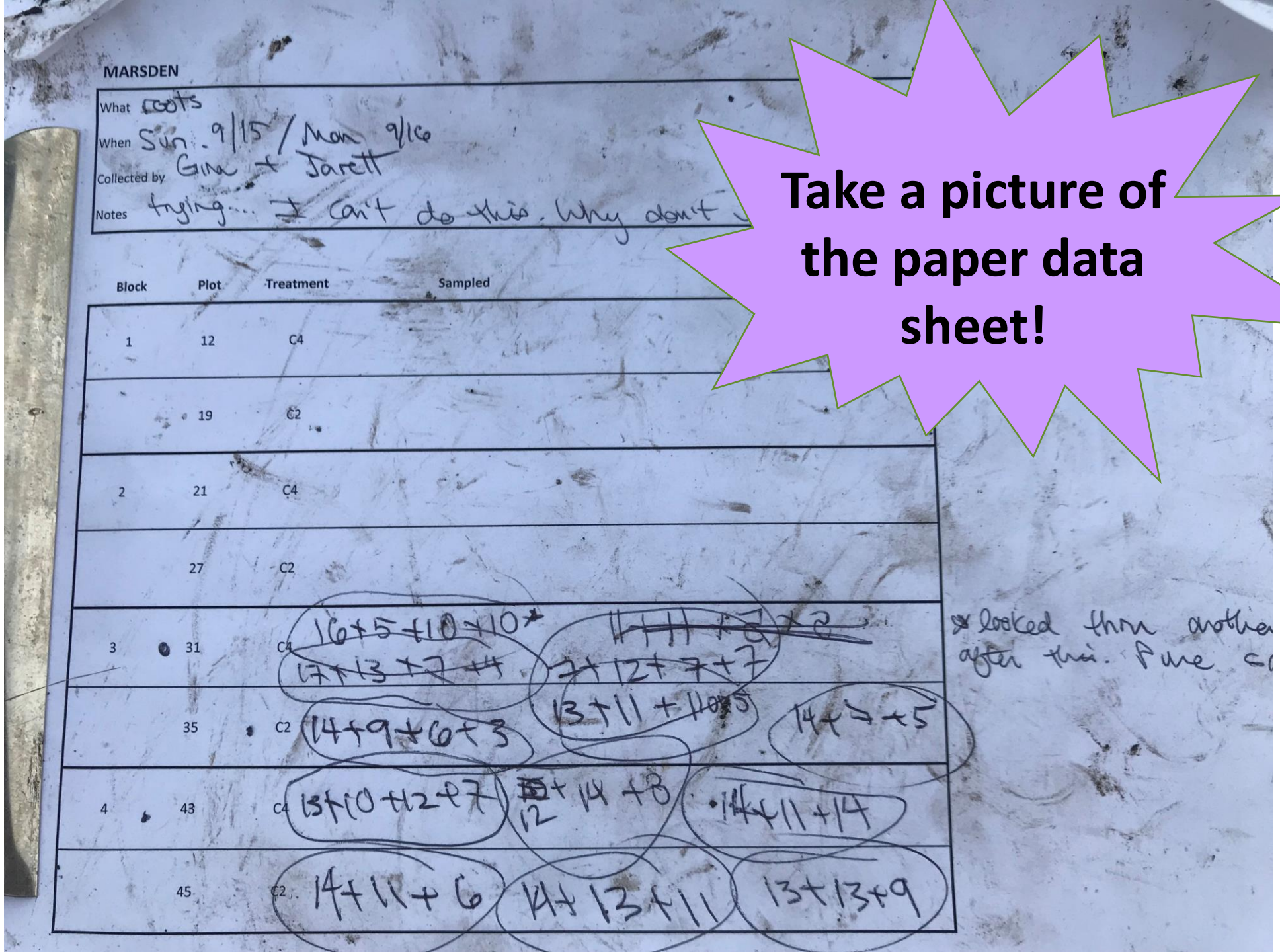
What LAI
When 7/2/19
Collected by Wyatt W.
Notes

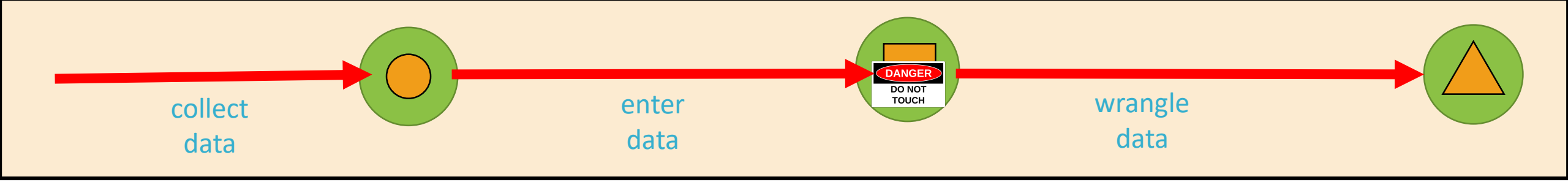
Block	Plot	Treatment	Sampled	Notes
1	12	C4	958.9+591.9+1573.6+1717.5+900.2	
	19	C2	1792+2153.2+2466.6+936.6	
2	21	C4	2797.1+1102.6+1991.3+2154.1+1153.1	
	27	C2	835.8+852.9+1434.2+1772.6+923.4+169.8	
3	31	C4	1533.4+1639.4+1513.7+1530.4	
	35	C2	2047.2+1565.5+1765.1+961.1+162.5=6501.4	
	43	C4	1070.25+785.5+1938.2+1966.34+1709.8+151	
	45	C2	1514.05+1466.9+1692.2+726+119	

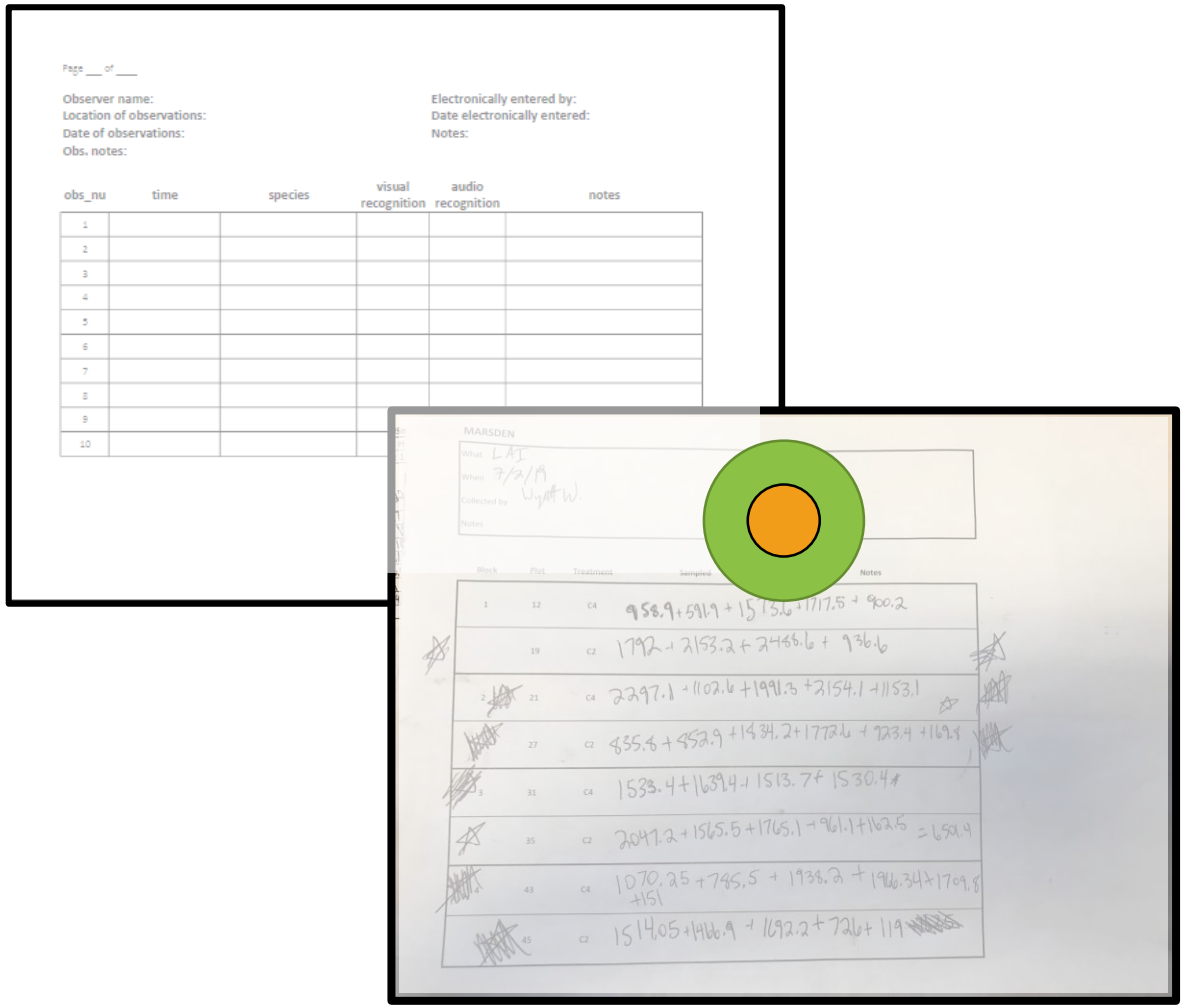




Take a picture of the paper data sheet!


MARSDEN
What roots
When Sun. 9/15 / Mon 9/16
Collected by Gina + Jarett
Notes trying... I can't do this. Why don't
Block Plot Treatment Sampled
1 12 C4
19 C2
2 21 C4
27 C2
3 31 C4
35 C2
4 43 C4
45 C2
Take a picture of the paper data sheet!





Page ___ of ___

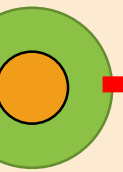
Observer name:
Location of observations:
Date of observations:
Obs. notes:

Electronically entered by:
Date electronically entered:
Notes:

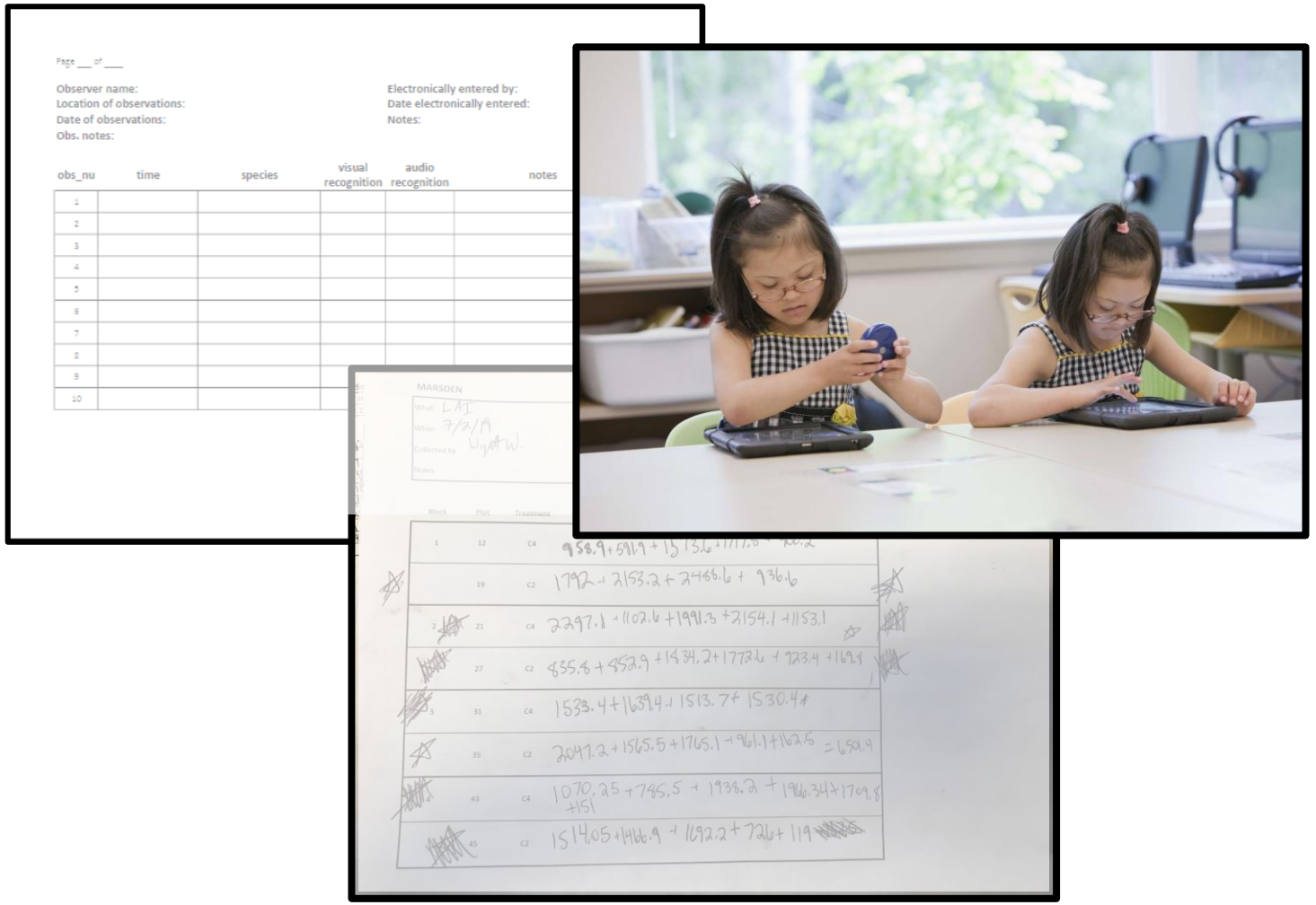
obs_nu	time	species	visual recognition	audio recognition	notes
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

collect data

enter data



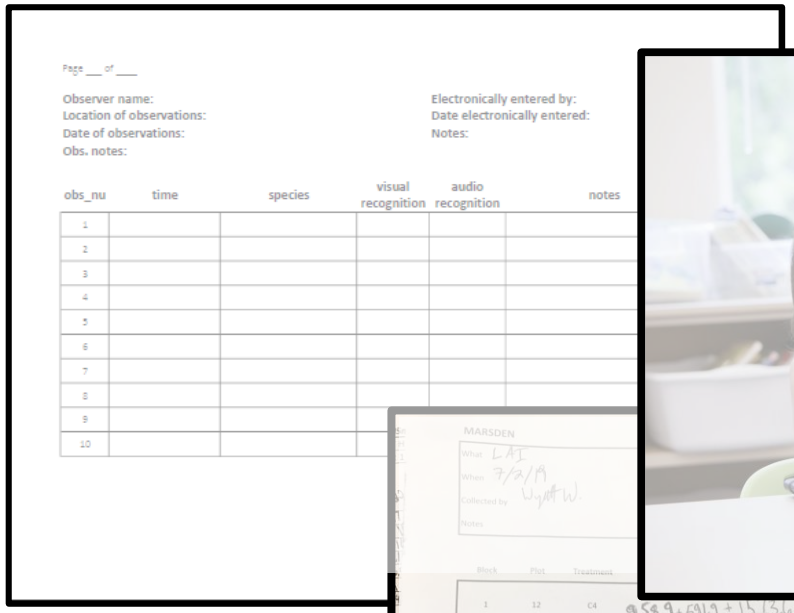
wrangle data



collect data

enter data

wrangle data

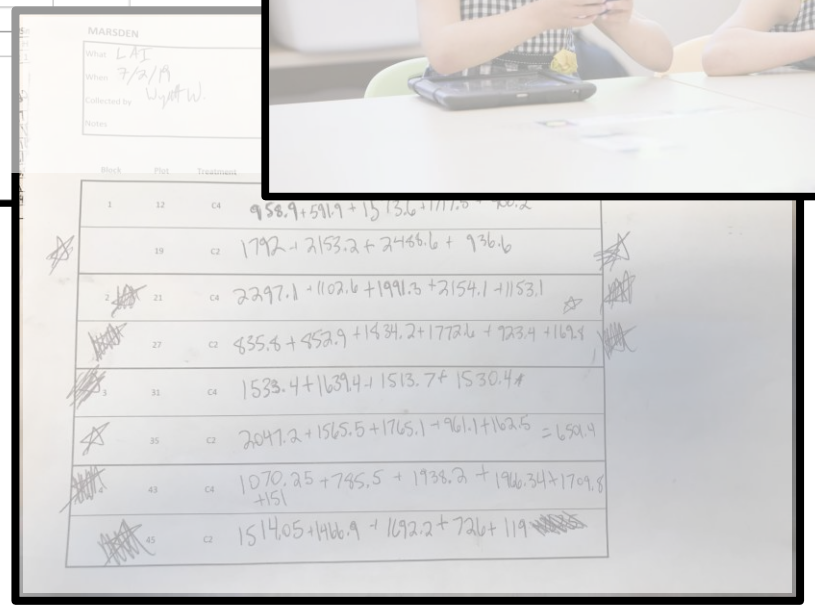


Page ___ of ___

Observer name:
Location of observations:
Date of observations:
Obs. notes:

Electronically entered by:
Date electronically entered:
Notes:

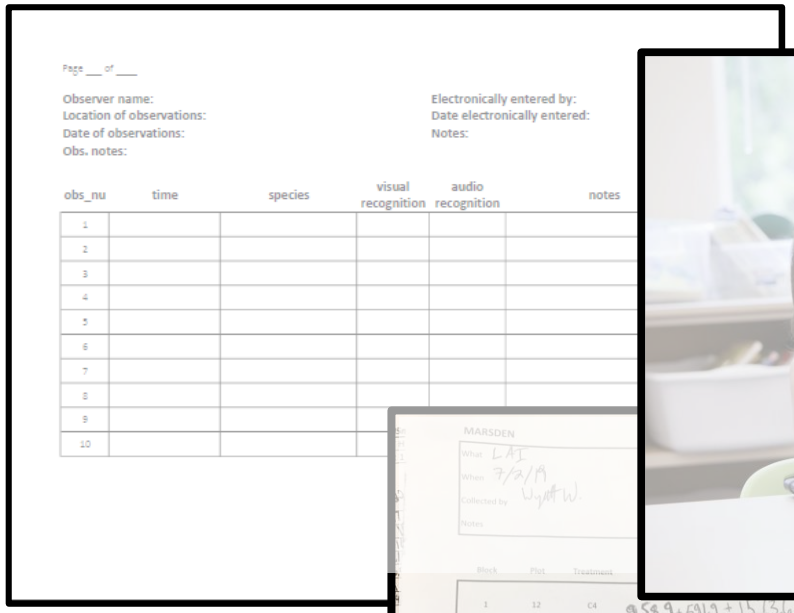
obs_nu	time	species	visual recognition	audio recognition	notes
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

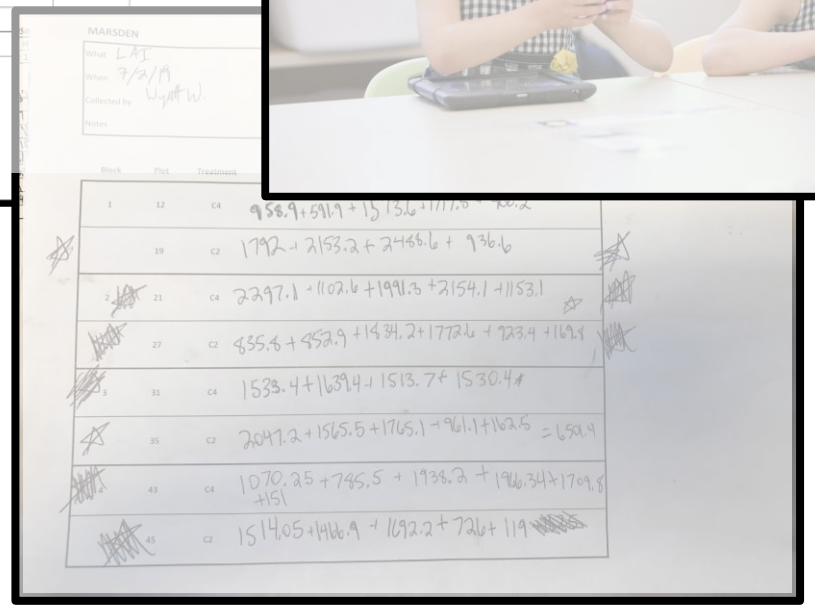


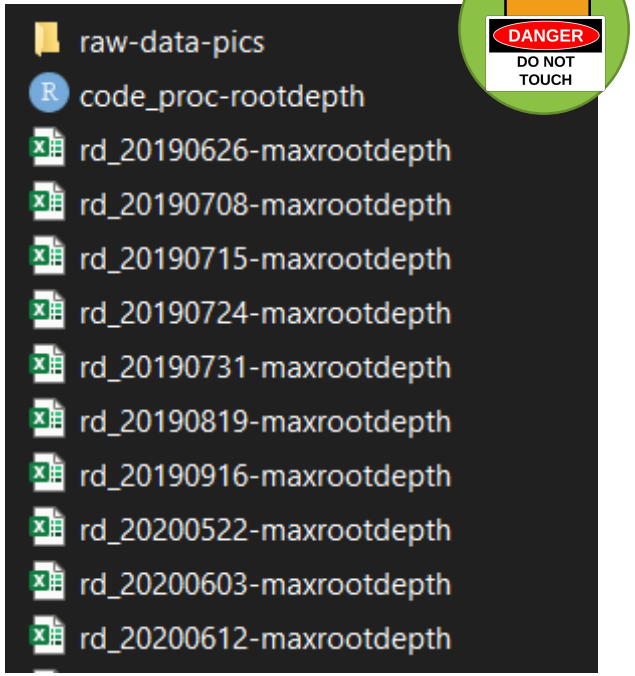
collect data

enter data

wrangle data





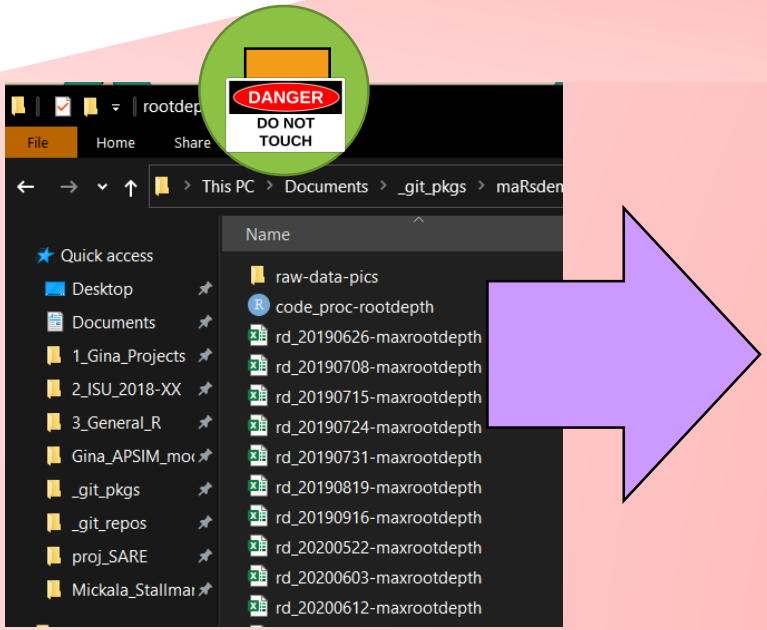


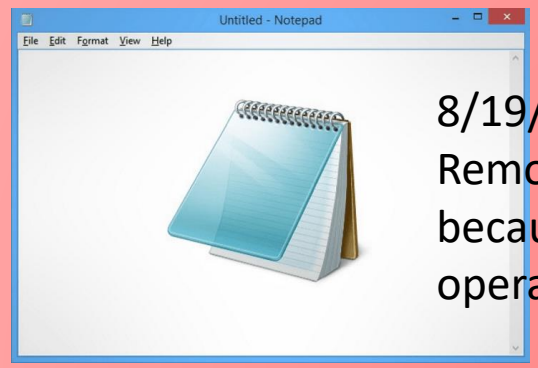
collect data

enter data

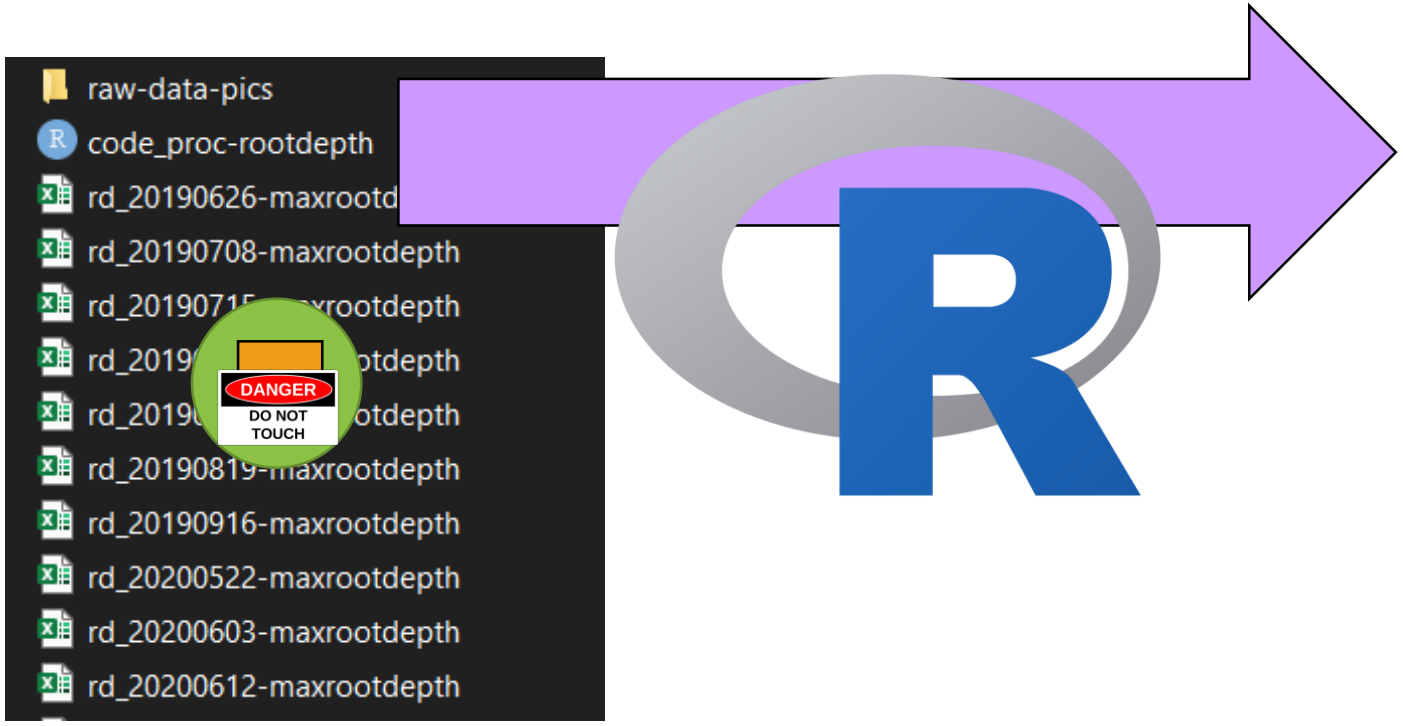
DANGER DO NOT TOUCH

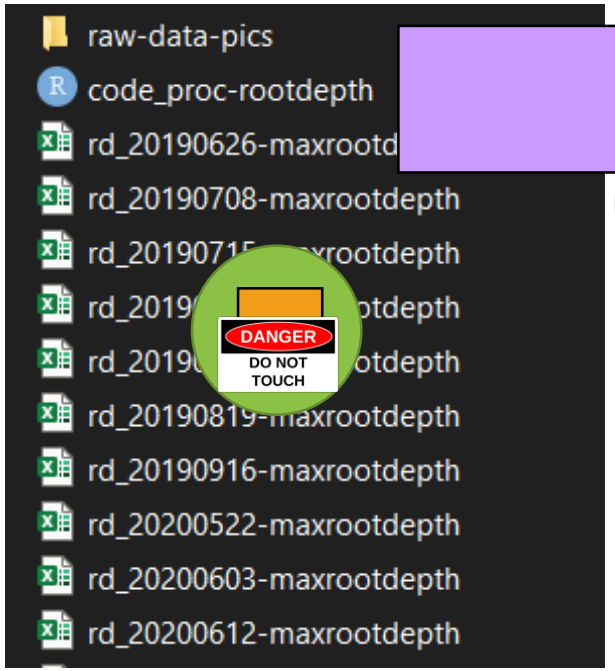
wrangle data

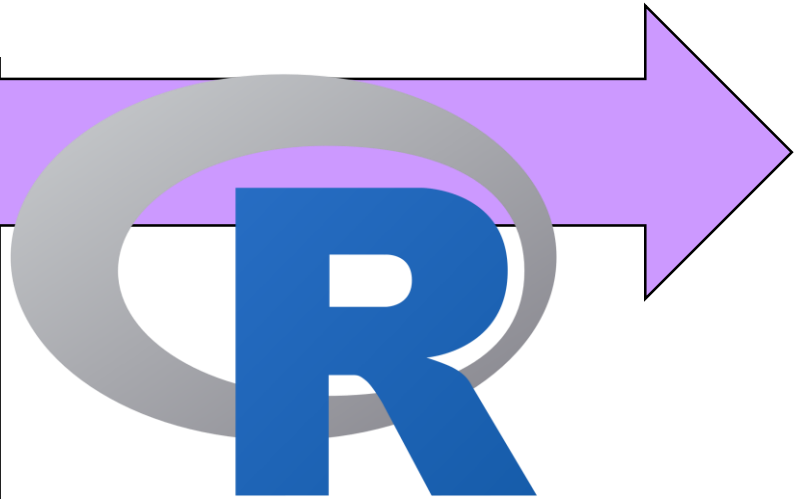




8/19/1945
Removed obs 59 because machine operator sneezed


raw-data-pics
code_proc-rootdepth
rd_20190626-maxrootd
rd_20190708-maxrootdepth
rd_20190715-maxrootdepth
rd_20190819-maxrootdepth
rd_20190916-maxrootdepth
rd_20200522-maxrootdepth
rd_20200603-maxrootdepth
rd_20200612-maxrootdepth
DANGER
DO NOT TOUCH
R





AutoSave Off

File Home Insert Page Layout Formulas Data Review

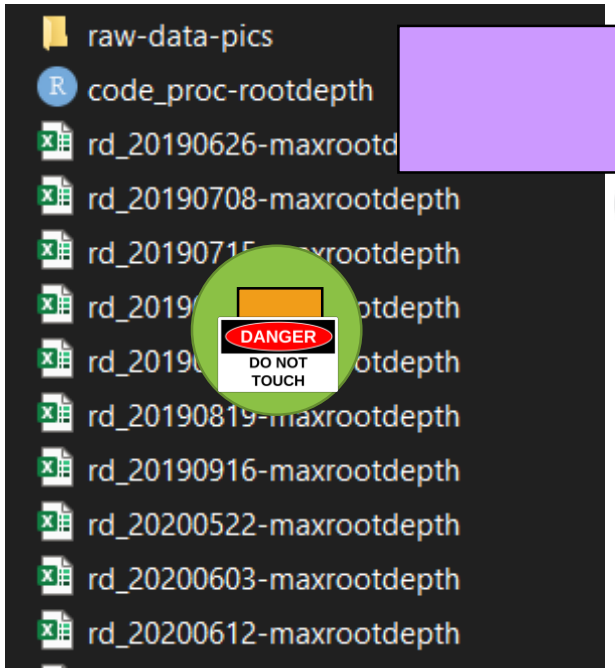
Paste Cut Copy Format Painter

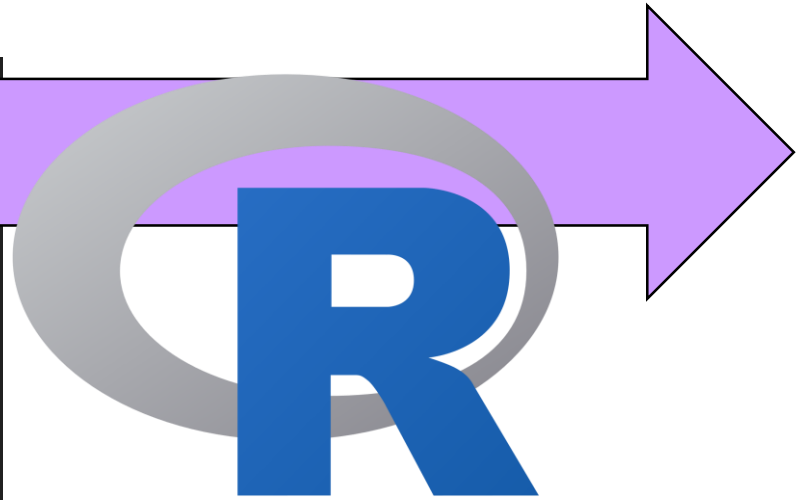
Clipboard Font

A1 year

	A	B	C	D	E	F	G
1	year	date	doy	plot_id	lai_cm2pl		
2	2018	6/1/2018	152	2018_13	382.2788		
3	2018	6/1/2018	152	2018_18	491.8917		
4	2018	6/1/2018	152	2018_24	648.4343		
5	2018	6/1/2018	152	2018_28	462.7012		
6	2018	6/1/2018	152	2018_34	291.1287		
7	2018	6/1/2018	152	2018_37	515.1943		
8	2018	6/1/2018	152	2018_44	487.1675		
9	2018	6/1/2018	152	2018_46	513.0775		
10	2018	6/18/2018	169	2018_13	3256.685		
11	2018	6/18/2018	169	2018_18	3542.381		
12	2018	6/18/2018	169	2018_24	3623.366		
13	2018	6/18/2018	169	2018_28	3533.6		
14	2018	6/18/2018	169	2018_34	2861.985		
15	2018	6/18/2018	169	2018_37	3814.048		
16	2018	6/18/2018	169	2018_44	2992.357		
17	2018	6/18/2018	169	2018_46	3467.939		
18	2018	7/3/2018	184	2018_13	NA		
19	2018	7/3/2018	184	2018_18	NA		
20	2018	7/3/2018	184	2018_24	NA		
21	2018	7/3/2018	184	2018_28	NA		
22	2018	7/3/2018	184	2018_34	NA		
23	2018	7/3/2018	184	2018_37	NA		
24	2018	7/3/2018	184	2018_44	NA		
25	2018	7/3/2018	184	2018_46	NA		
26	2018	7/17/2018	198	2018_13	4922.826		
27	2018	7/17/2018	198	2018_18	5395.206		
28	2018	7/17/2018	198	2018_24	6000.032		
29	2018	7/17/2018	198	2018_28	5214.014		

cornlai





AutoSave Off

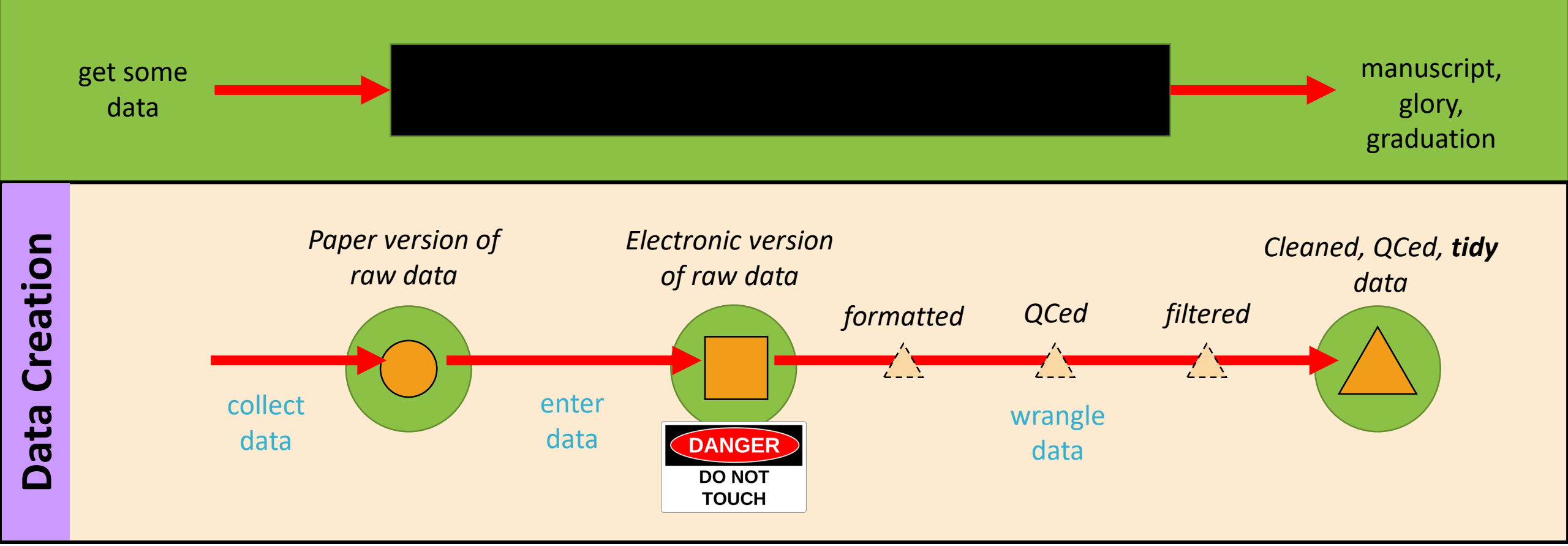
File Home Insert Page Layout Formulas Data Review

Cut
Copy
Format Painter
Paste
Clipboard
Font

A1 year

	A	B	C	D	E	F	G
1	year	date	doy	plot_id	lai_cm2pl		
2	2018	6/1/2018	152	2018_13	382.2788		
3	2018	6/1/2018	152	2018_18	491.8917		
4	2018	6/1/2018	152	2018_24	648.4343		
5	2018	6/1/2018	152	2018_28	462.7012		
6	2018	6/1/2018	152	2018_34	291.1287		
7	2018	6/1/2018	152	2018_37	515.1943		
8	2018	6/1/2018	152	2018_44	487.1675		
9	2018	6/1/2018	152	2018_46	513.0775		
10	2018	6/18/2018	169	2018_13	3256.685		
11	2018	6/18/2018	169	2018_18	3542.381		
12	2018	6/18/2018	169	2018_24	3623.366		
13	2018	6/18/2018	169	2018_28	3533.6		
14	2018	6/18/2018	169	2018_34	2861.985		
15		6/18/2018	169	2018_37			
	2018	6/18/2018	169	2018_44	2992.357		
17	2018	6/18/2018	169	2018_46	3467.939		
18	2018	7/3/2018	184	2018_13	NA		
19	2018	7/3/2018	184	2018_18	NA		
20	2018	7/3/2018	184	2018_24	NA		
21	2018	7/3/2018	184	2018_28	NA		
22	2018	7/3/2018	184	2018_34	NA		
23	2018	7/3/2018	184	2018_37	NA		
24	2018	7/3/2018	184	2018_44	NA		
25	2018	7/3/2018	184	2018_46	NA		
	2018	7/17/2018	198	2018_13	4922.826		
27		7/17/2018	198	2018_18			
28	2018	7/17/2018	198	2018_24	6000.032		
29	2018	7/17/2018	198	2018_28	5214.014		

cornlai


get some data
manuscript, glory, graduation
Data Creation
Paper version of raw data
Electronic version of raw data
Cleaned, QCed, tidy data
formatted
QCed
filtered
collect data
enter data
wrangle data
DANGER
DO NOT TOUCH

get some data

manuscript, glory, graduation

Data Creation

Paper version of raw data

Electronic version of raw data

formatted

QCed

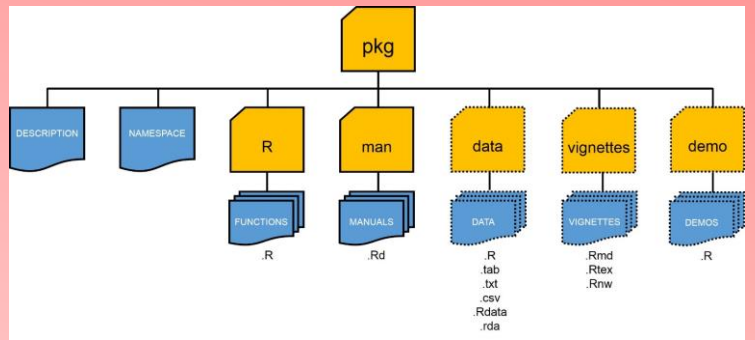
filtered

*Cleaned, QCed, **tidy** data*

collect data

enter data

wrangle data



get some data

[redacted]

manuscript, glory, graduation

Data Creation

Paper version of raw data

collect data

Electronic version of raw data

enter data

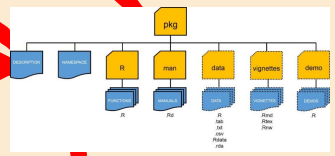
formatted

QCed

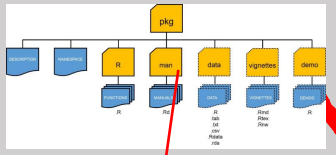
filtered

wrangle data

Cleaned, QCed, tidy format data



Data Analysis



Give to poor grad student taking up your project

Analyze data

Summarized data (ex. stats, figures)

Contextualize data

Some form of communication about your data (ex. manuscript)

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function | Addins | Project: (None)

Untitled1*

Source on Save | Run | Source

```
1 library(maRsden)
2 mrs
```

mrs_cornbio
mrs_cornbio_wo
mrs_cornlai
mrs_earrows
mrs_elevation
mrs_penetrom
mrs_phen
mrs_plotkey
mrs_residue

2:4 (Top Level) R Script

Console Terminal Jobs

~/

```
> 
> 
> 
> library(maRsden)
> 
```

Environment History Connections

Import Dataset | List

Global Environment

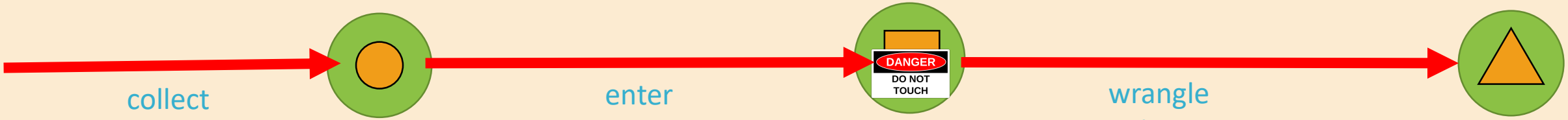
Environment is empty

Files Plots Packages Help Viewer

Zoom | Export

Type here to search

10:41 AM
9/6/2020

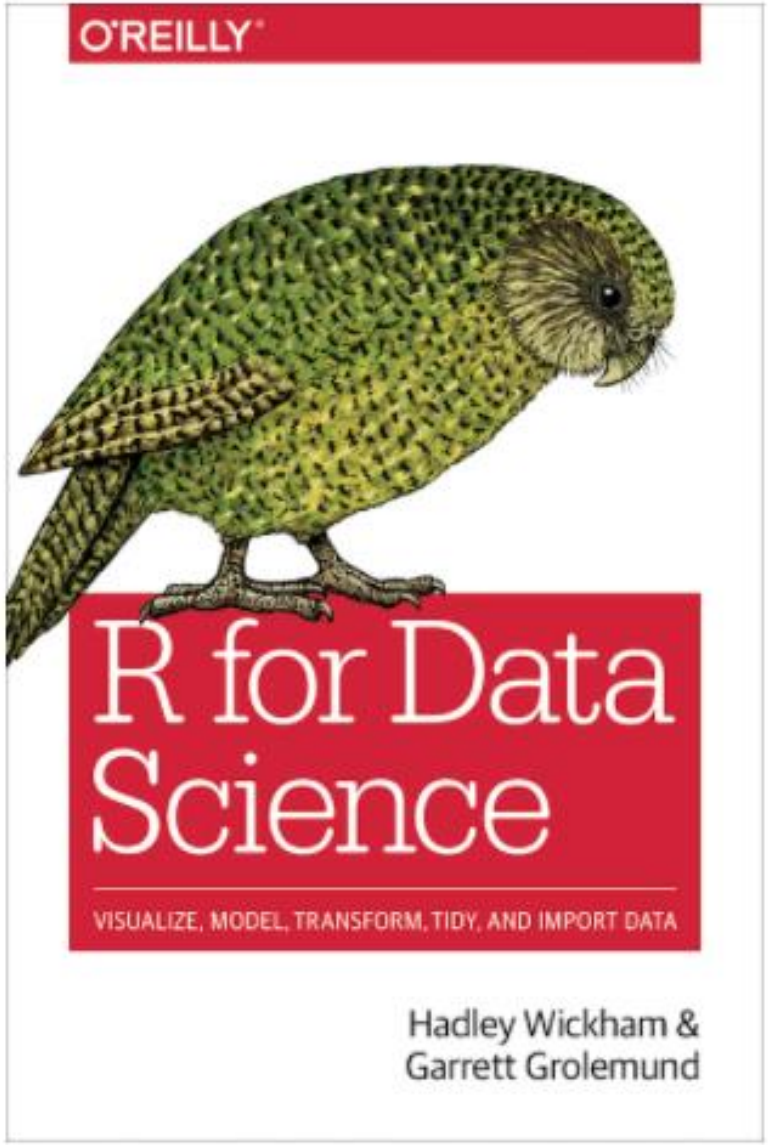


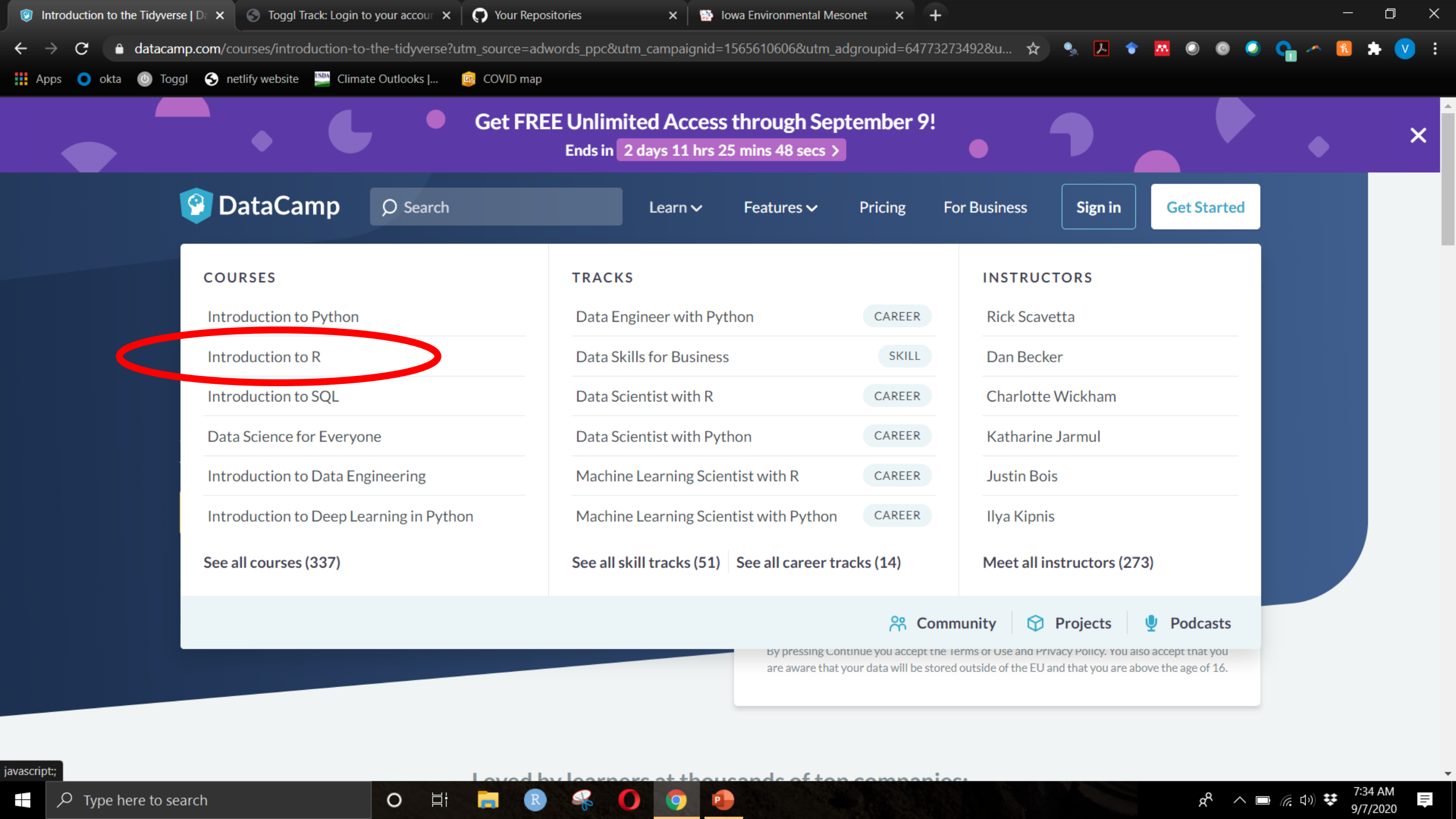
- Separate data collection and data analysis
- Consider publishing your data!
- Take pictures of paper copy of data
- Never touch the raw data
- Never touch the raw data
- If using Excel, consider learning R, but in the meantime keep a log of differences between raw data and tidy data

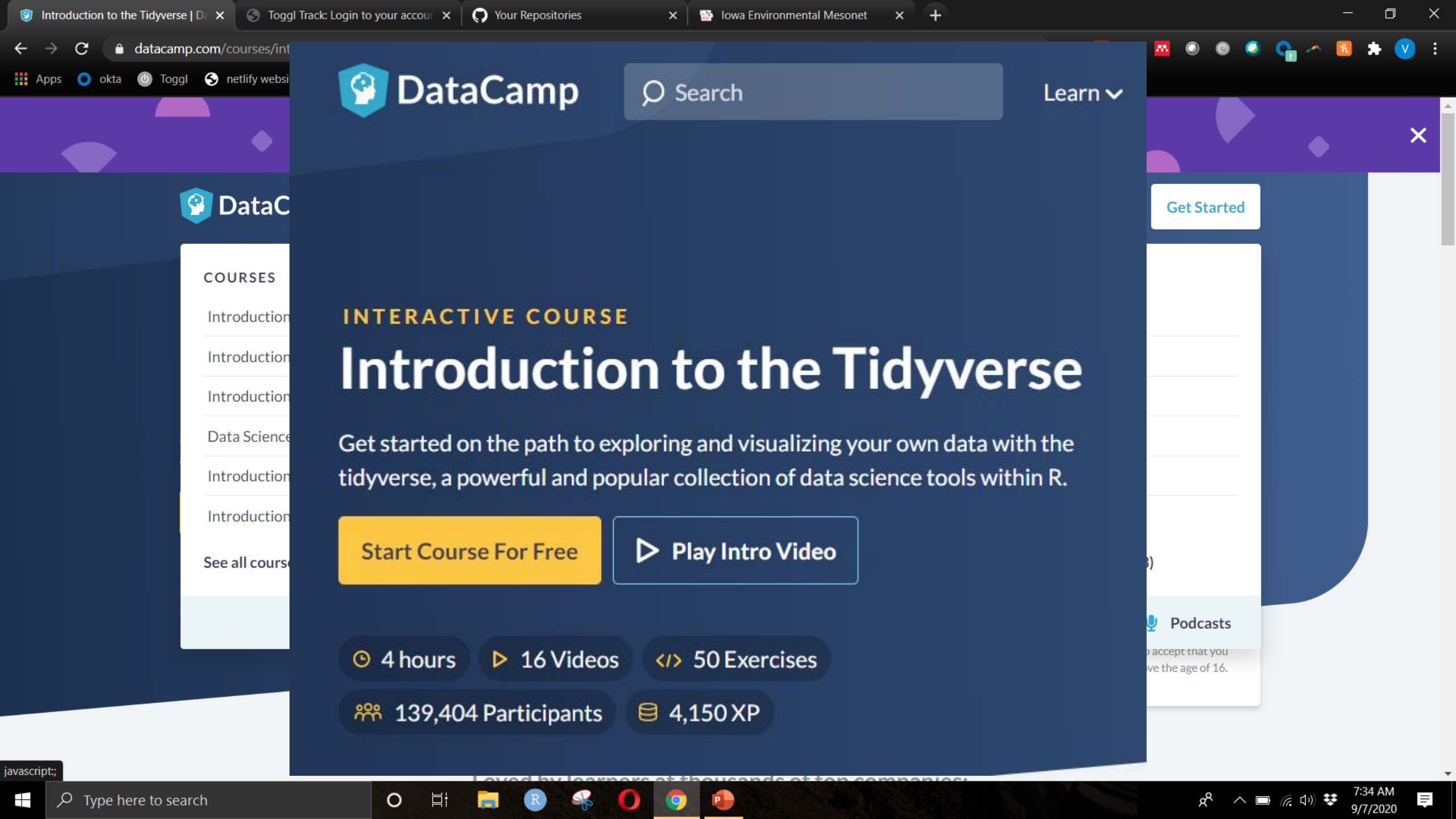
40 hours

- https://r4ds.had.co.nz/
- Stat579 (Intro to R and the tidyverse)
  - Every fall
- Stat585 (More advanced R, R packages)
  - Odd springs (spring 2021!)

- DataFEWSion program…?
- DataCamp




Introduction to the Tidyverse | Da
Toggl Track: Login to your accour
Your Repositories
Iowa Environmental Mesonet
datacamp.com/courses/introduction-to-the-tidyverse?utm_source=adwords_ppc&utm_campaignid=1565610606&utm_adgroupid=64773273492&u...
Apps
okta
Toggl
netlify website
Climate Outlooks |...
COVID map
Get FREE Unlimited Access through September 9!
Ends in 2 days 11 hrs 25 mins 48 secs
DataCamp
Search
Learn
Features
Pricing
For Business
Sign in
Get Started
COURSES
Introduction to Python
Introduction to R
Introduction to SQL
Data Science for Everyone
Introduction to Data Engineering
Introduction to Deep Learning in Python
See all courses (337)
TRACKS
Data Engineer with Python
CAREER
Data Skills for Business
SKILL
Data Scientist with R
CAREER
Data Scientist with Python
CAREER
Machine Learning Scientist with R
CAREER
Machine Learning Scientist with Python
CAREER
See all skill tracks (51)
See all career tracks (14)
INSTRUCTORS
Rick Scavetta
Dan Becker
Charlotte Wickham
Katharine Jarmul
Justin Bois
Ilya Kipnis
Meet all instructors (273)
Community
Projects
Podcasts
By pressing Continue you accept the Terms of Use and Privacy Policy. You also accept that you are aware that your data will be stored outside of the EU and that you are above the age of 16.
javascript:;
Type here to search
7:34 AM
9/7/2020


Introduction to the Tidyverse | Da
Toggl Track: Login to your accour
Your Repositories
Iowa Environmental Mesonet
datacamp.com/courses/int
Apps
okta
Toggl
netlify websi
DataCamp
Search
Learn
Get Started
COURSES
Introduction
Introduction
Introduction
Data Science
Introduction
Introduction
See all cours
INTERACTIVE COURSE
Introduction to the Tidyverse
Get started on the path to exploring and visualizing your own data with the tidyverse, a powerful and popular collection of data science tools within R.
Start Course For Free
Play Intro Video
4 hours
16 Videos
50 Exercises
139,404 Participants
4,150 XP
Podcasts
javascript:;
Type here to search
7:34 AM
9/7/2020