

# GG606

## Workflows

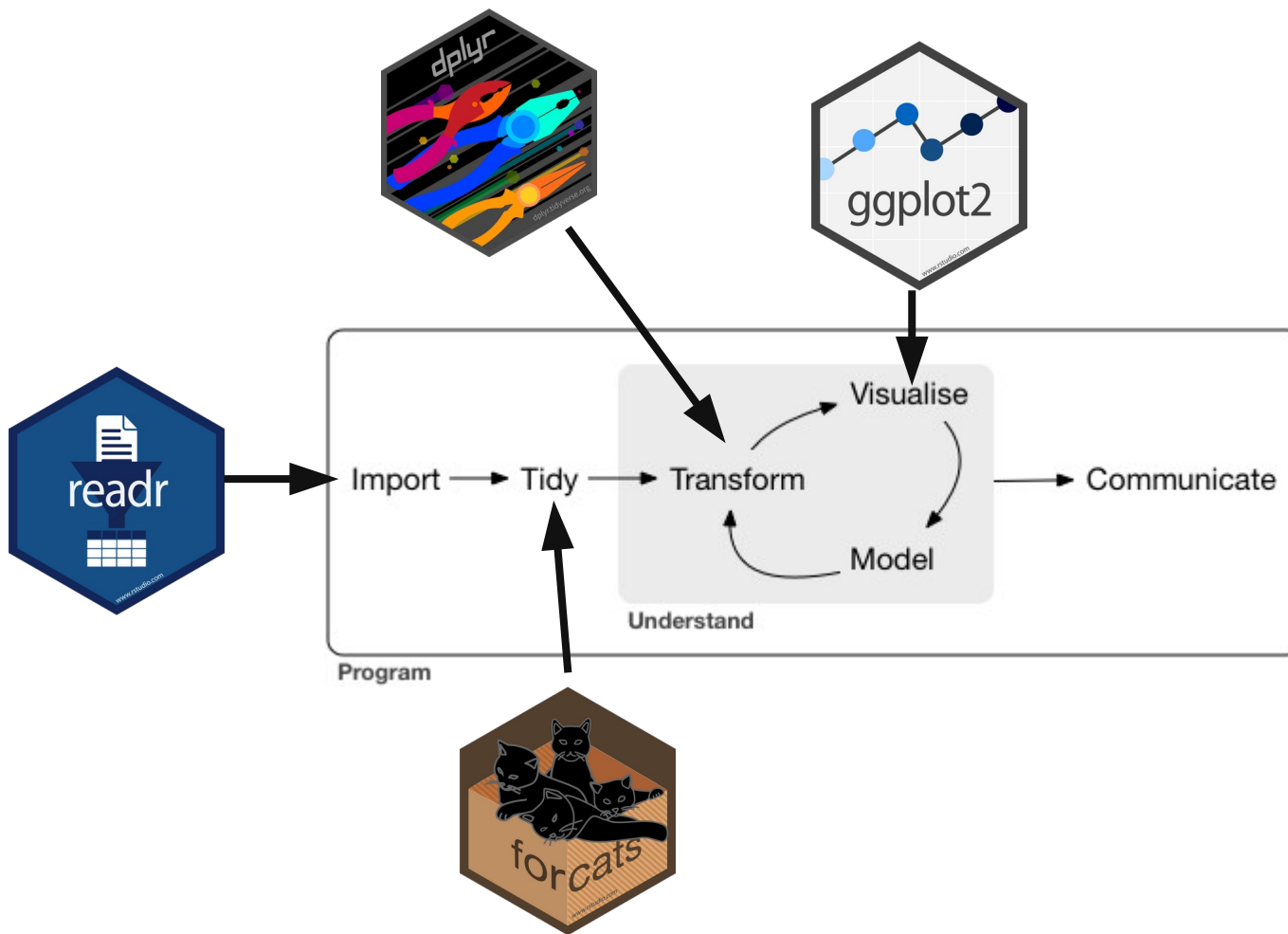
# Homework

- `install.packages("palmerpenguins")`
- Make some interesting plots with new `geom_` and try out some `theme_`
- [ggplot2.tidyverse.org](https://ggplot2.tidyverse.org)
- Try out `ggsave`
- [pangaea.de](https://pangaea.de) & [www.frdr-dfdr.ca](https://www.frdr-dfdr.ca)

# Homework

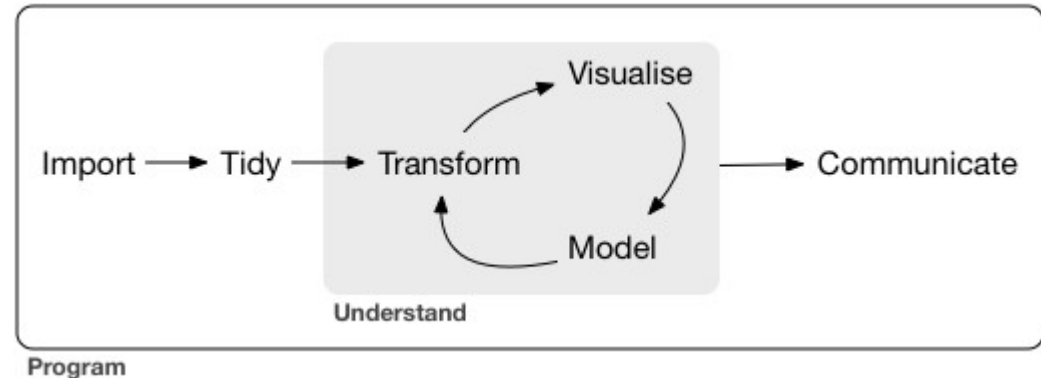
# Workflows

- Organisation
- Reproducible
- Inputs & Outputs
- R-Script vs R-Markdown vs Function



# R-Script

- File > New File > R Script
- Series of command *and comments*
- Sequence & Story



# New Computer/Collaborator

- You are your own (future|past) collaborator
- *New computer?*
- *Sent files to collaborators?*

# "FINAL".doc



FINAL.doc!



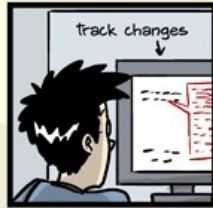
FINAL\_rev.2.doc



FINAL\_rev.6.COMMENTS.doc



FINAL\_rev.8.comments5.  
CORRECTIONS.doc



FINAL\_rev.18.comments7.  
corrections9.MORE.30.doc



FINAL\_rev.22.comments49.  
corrections.10.#@\$%WHYDID  
ICOMETOGRADSCHOOL????.doc

JORGE CHAM © 2012



# A STORY TOLD IN FILE NAMES:

Location: C:\user\research\data

Filename	Date Modified	Size	Type
data_2010.05.28_test.dat	3:37 PM 5/28/2010	420 KB	DAT file
data_2010.05.28_re-test.dat	4:29 PM 5/28/2010	421 KB	DAT file
data_2010.05.28_re-re-test.dat	5:43 PM 5/28/2010	420 KB	DAT file
data_2010.05.28_calibrate.dat	7:17 PM 5/28/2010	1,256 KB	DAT file
data_2010.05.28_huh??.dat	7:20 PM 5/28/2010	30 KB	DAT file
data_2010.05.28_WTF.dat	9:58 PM 5/28/2010	30 KB	DAT file
data_2010.05.29_aaarrgh.dat	12:37 AM 5/29/2010	30 KB	DAT file
data_2010.05.29_#\$@*&!!..dat	2:40 AM 5/29/2010	0 KB	DAT file
data_2010.05.29_crap.dat	3:22 AM 5/29/2010	437 KB	DAT file
data_2010.05.29_notbad.dat	4:16 AM 5/29/2010	670 KB	DAT file
data_2010.05.29_woohoo!!..dat	4:47 AM 5/29/2010	1,349 KB	DAT file
data_2010.05.29_USETHISONE.dat	5:08 AM 5/29/2010	2,894 KB	DAT file
analysis_graphs.xls	7:13 AM 5/29/2010	455 KB	XLS file
ThesisOutline!.doc	7:26 AM 5/29/2010	38 KB	DOC file
Notes_Meeting_with_ProfSmith.txt	11:38 AM 5/29/2010	1,673 KB	TXT file
JUNK...	2:45 PM 5/29/2010		Folder
data_2010.05.30_startingover.dat	8:37 AM 5/30/2010	420 KB	DAT file

Type: Ph.D Thesis Modified: too many times

Copyright: Jorge Cham

www.phdcomics.com

# New Computer/Collaborator

- Will your files always be here?
- Where should your data, code, outputs go?
- File paths can be a nightmare

# New Computer/Collaborator

- Will your files always be here?
- Where should your data, code, outputs go?
- File paths can be a nightmare
- RStudio Project helps

```
> getwd()  
[1] "/home/jason/school/Laurier teaching/GG606 Winter 2023"
```

Microsoft Excel



This workbook contains links to one or more external sources that could be unsafe.

If you trust the links, update them to get the latest data. Otherwise, you can keep working with the data you have.

Upnate

Don't Upnate

Help

Microsoft Excel



We can't update some of the links in your workbook right now.

You can continue without updating their values, or edit the links you think are wrong.

Continue

Edit Links...

# Project-Oriented Workflow

- Commands
- R Scripts (series of commands)
- R Project (series of related things)
- here package  
(project-oriented workflow & portability)

here: find your  
PATH!





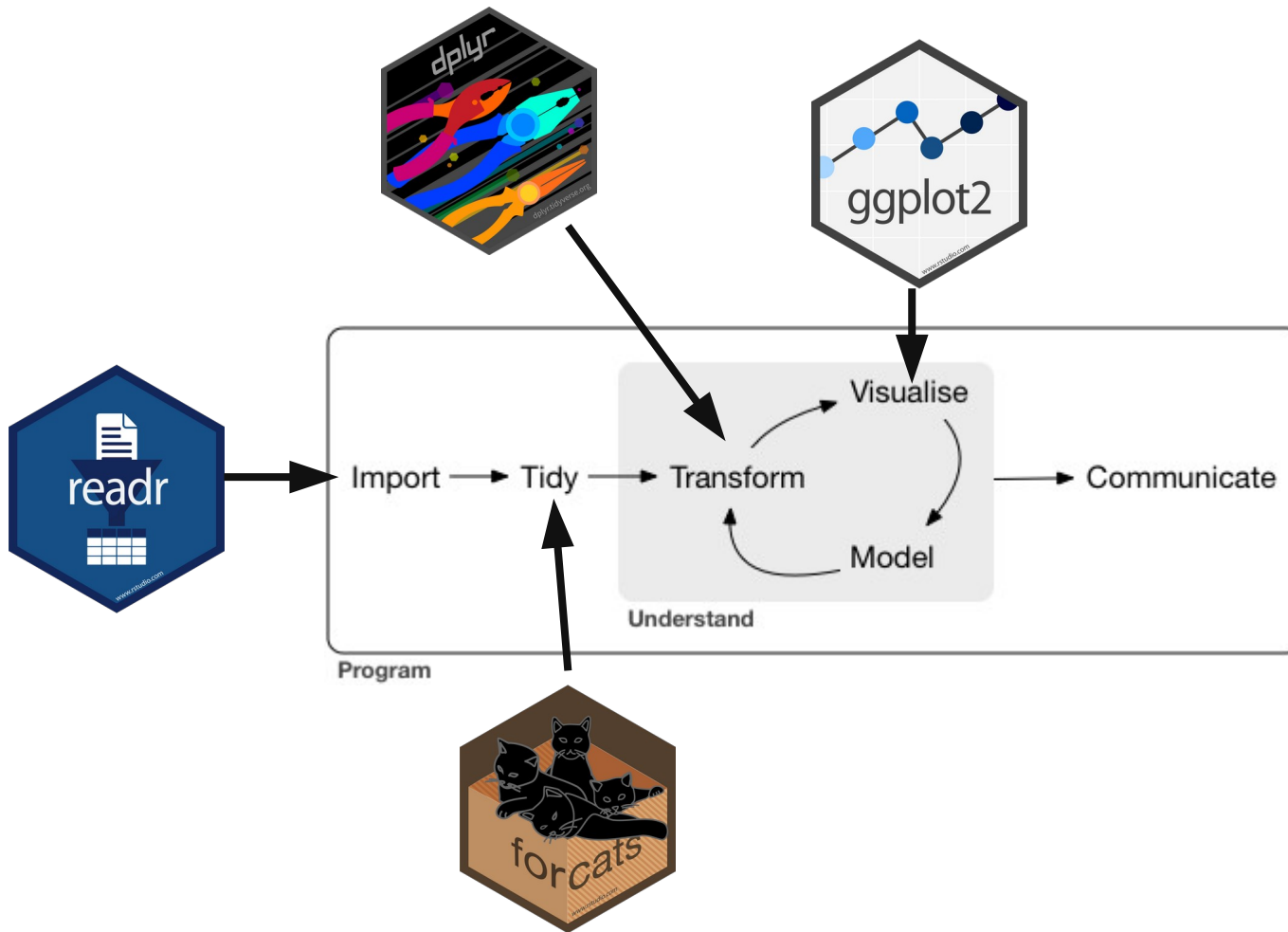
# here

- It figures out *real* path to project
  - "C:\Users\jenny\path\that\only\I\have"
  - "/Users/jenny/cuddly\_broccoli/  
verbose\_funicular/foofy/data"



# here

- It figures out *real* path to project
  - `here("folder", "file")`
  - `ggsave(here("figures",  
"beak_size_by_species.pdf"))`



# Template

- What folders do you want in each project?
- Inputs & Outputs
- Code
- Documents

*Build a basic template*



www.rstudio.com

# Wrap Up

- Make folders that make sense
- Put a data file in there
- `library(readr)`
- `read_csv()` for your found data and `penguins_raw.csv`
- Update your code to use here, load data, save figures
- Think about metadata

Why are Jenny Bryan & Timothée Poisot  
So Worked Up?

If the first line of your R script is

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

I\* will come into your office and  
SET YOUR COMPUTER ON FIRE 🔥.

\* or maybe Timothée Poisot will

If the first line of your R script is

```
rm(list = ls())
```

I will come into your office and  
SET YOUR COMPUTER ON FIRE 🔥.

# Homework

- folder structure for the workflow
  - we spoke about keeping raw data separate from processed data and keeping figures and/or tables together)
- use the here package and function
- R script to load data from pangaea.de or [www.frdr-dfdr.ca](http://www.frdr-dfdr.ca)
  - example, `read_csv(here("folder", "file"))`
- create and save a figure to an appropriate folder
  - hint, use the `ggsave` and `here` functions
- look at the metadata fields & make list of metadata required to deposit data



