

Symposium



R pour ornithologue

Points de vue des utilisateurRs et programmeuRs

R for Ornithologists

Perspectives from useRs to programmeRs (to birdeRs)

10:30-14:15 Salle des Plaines II

Stefanie E. LaZerte



R for Ornithologists

How R can benefit the study of Ornithology



 @steffilazerte

 steffilazerte

 steffilazerte.ca

 sel@steffilazerte.ca

What is R?

@steffilazerte

What is R?

@steffilazerte

A statistical programming language and environment

(free and open source!)

What is R?

@steffilazerte

A statistical programming language and environment

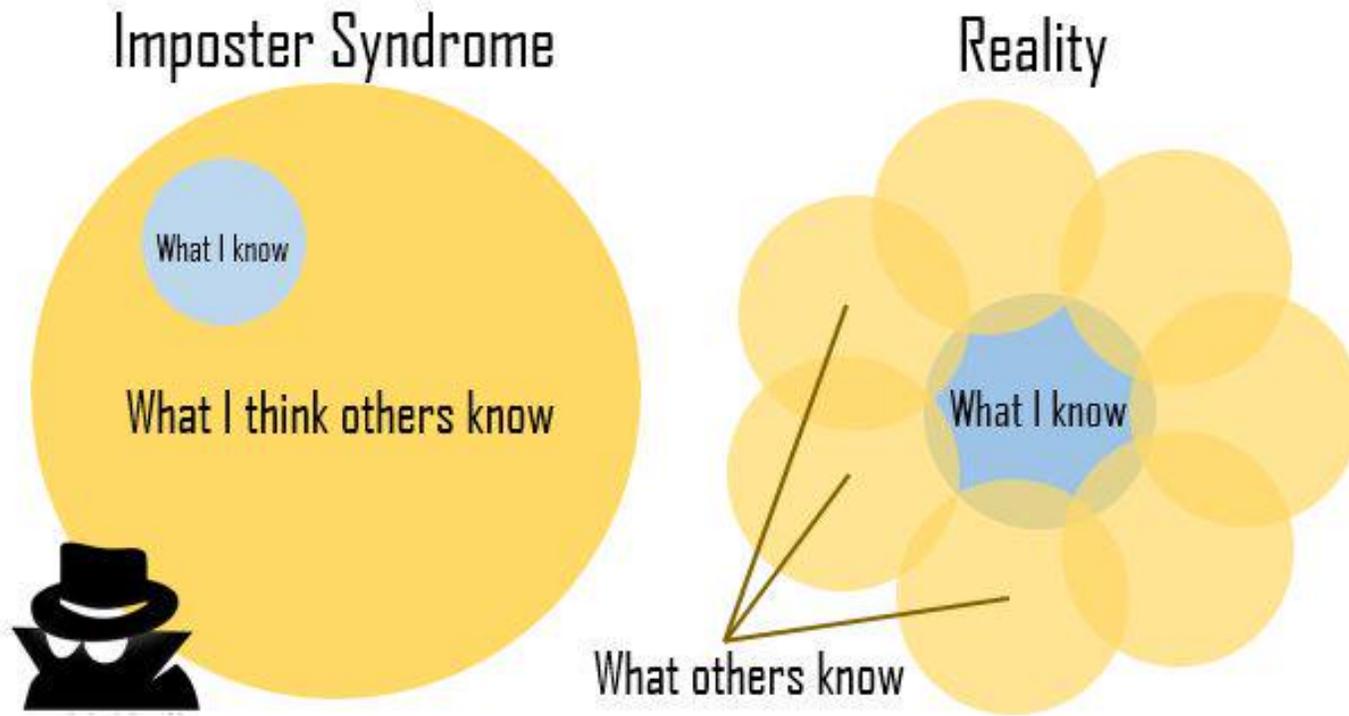
(free and open source!)

R uses packages

- Packages extend R (i.e. `nlme` and `lme4` add mixed modeling)
- Packages can be written by anyone
- Some are ok, some are good, some are AMAZING
- **Base R** is R without any extra packages (also good)

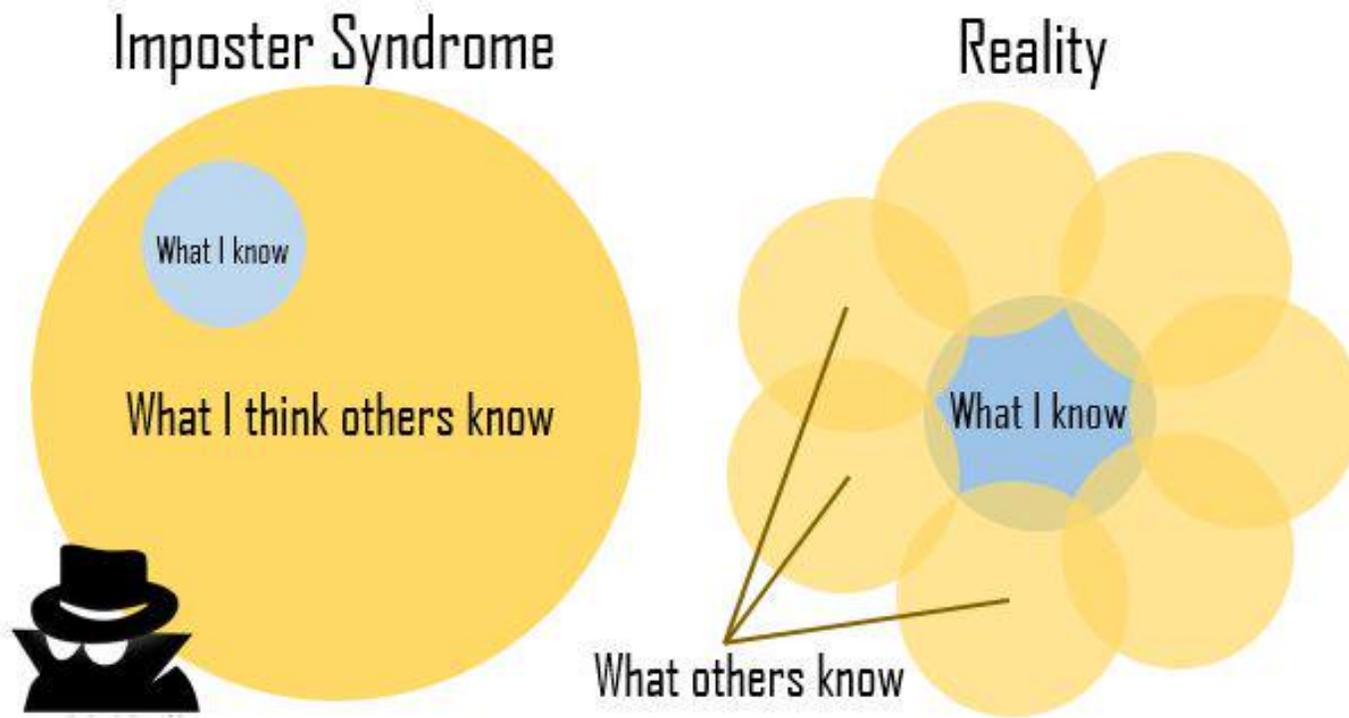
There are 1000's of packages!

Impost R Syndrome



David Whittaker

ImpostR Syndrome



David Whittaker

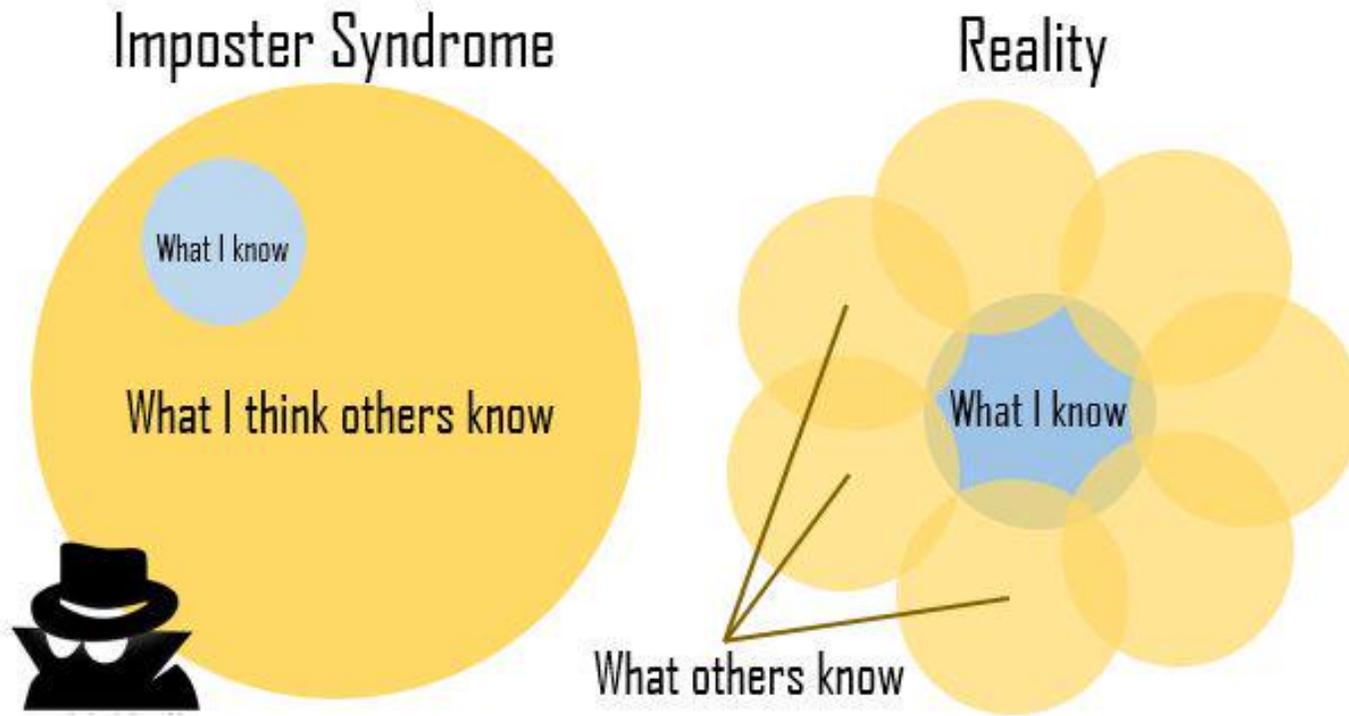
ImpostR Syndrome

Moral of the story?

Make friends, code in groups, learn together and don't beat yourself up

ImpostR Syndrome

@steffilazerte



David Whittaker

Using R in the undergraduate biology classroom: Hurdles, hints, and aha moments
([Here @ 1:45pm](#))

What I **am not** going to do

- Teach you R
- Talk about statistics



What I **am not** going to do

- Teach you R
- Talk about statistics

What I **am** going to do

- Explain how R can benefit ornithologists
- Showcase useful packages
- Give you resources to get started



What I **am not** going to do

- Teach you R
- Talk about statistics

What I **am** going to do

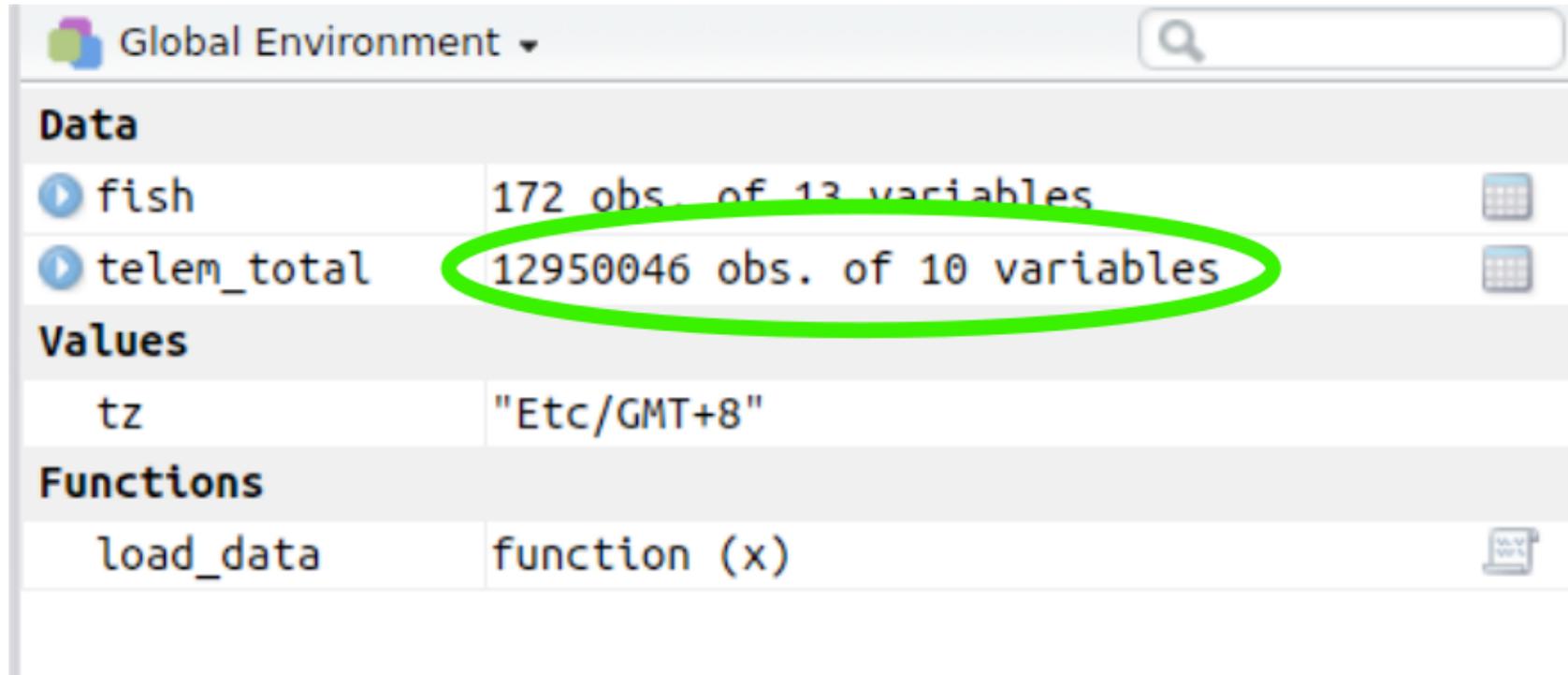
- Explain how R can benefit ornithologists
- Showcase useful packages
- Give you resources to get started
- Inspire you to take your **R** to the next level!



Why ornithologists should use R

@steffilazerte

R is powerful!



The screenshot shows the R Studio Global Environment window. It is titled 'Global Environment' and has a search bar. The window is divided into three main sections: Data, Values, and Functions. In the Data section, two objects are listed: 'fish' with 172 observations and 13 variables, and 'telem_total' with 12950046 observations and 10 variables. The 'telem_total' entry is circled in green. In the Values section, the 'tz' variable is shown with the value 'Etc/GMT+8'. In the Functions section, the 'load_data' function is shown as a function of 'x'.

Global Environment	
Data	
fish	172 obs. of 13 variables
telem_total	12950046 obs. of 10 variables
Values	
tz	"Etc/GMT+8"
Functions	
load_data	function (x)

Why ornithologists should use R

@steffilazerte

R is powerful!

The blessing and curse of automated data collection:

R and dealing with big data in a modern age

[\(Here @ 10:45am\)](#)

Super-computing with R:

Harnessing the power of the cloud to analyze big-bird-data, or just run your simulations, models, and cross-validations faster

[\(Here @ 11:15am\)](#)

Why ornithologists should use R

@steffilazerte

Reproducible Science

- Scripts are records of your work

```
m <- lm(mpg ~ cyl, data = mtcars)
summary(m)
```

Why ornithologists should use R

@steffilazerte

Reproducible Science

- Scripts are records of your work

```
m <- lm(mpg ~ cyl, data = mtcars)
summary(m)
```

- Scripts can be compiled into pdf/html reports with [rmarkdown](#) and [knitr](#)
(In RStudio: File > Compile Report)

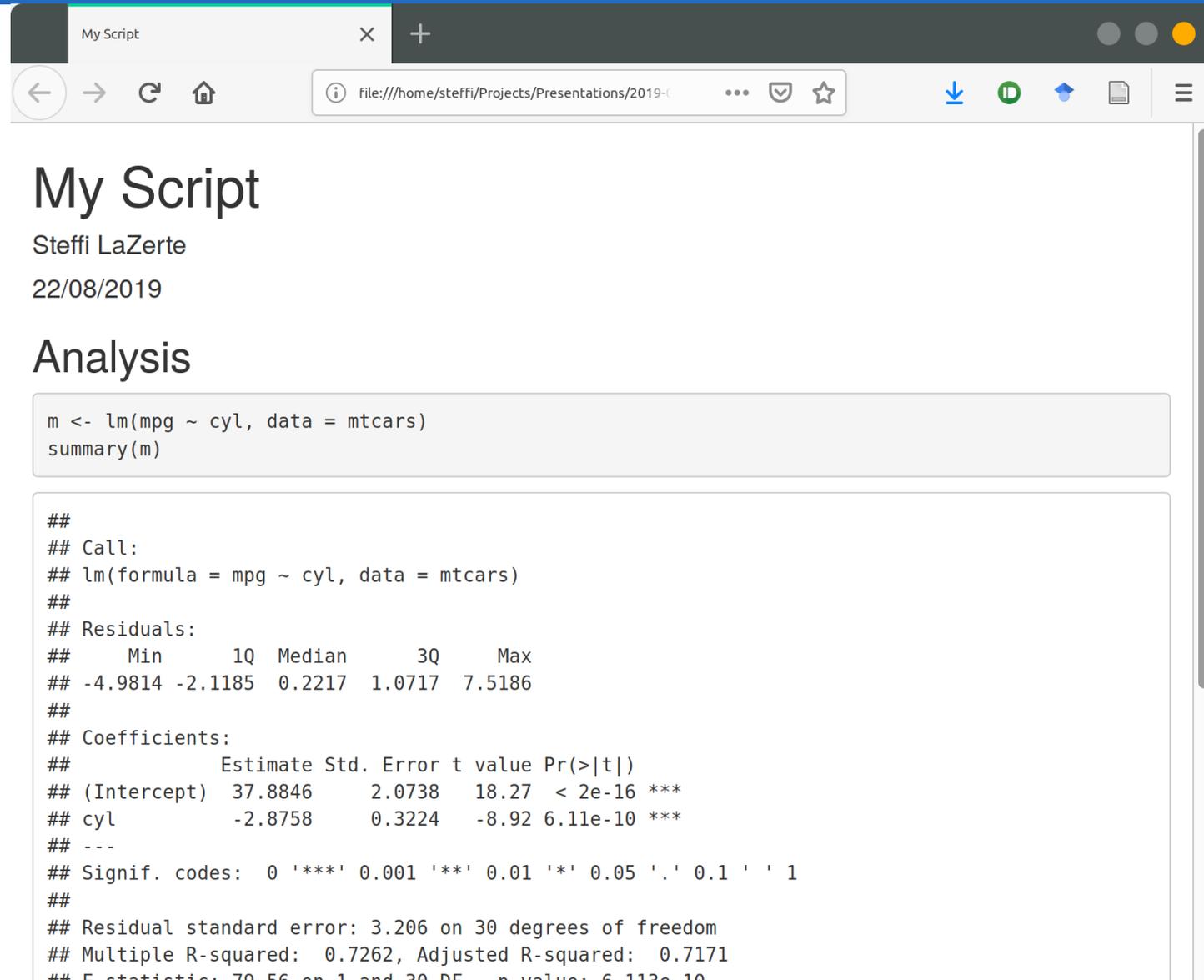


Why ornithologists should use R

@steffilazerte

Reproducible Science

- Keep track of code AND output



The screenshot shows a web browser window with a single tab titled "My Script". The address bar shows the file path: file:///home/steffi/Projects/Presentations/2019-08-22/My Script. The page content is as follows:

```
My Script
Steffi LaZerte
22/08/2019

Analysis

m <- lm(mpg ~ cyl, data = mtcars)
summary(m)

##
## Call:
## lm(formula = mpg ~ cyl, data = mtcars)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.9814 -2.1185  0.2217  1.0717  7.5186
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  37.8846     2.0738   18.27 < 2e-16 ***
## cyl         -2.8758     0.3224   -8.92 6.11e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.206 on 30 degrees of freedom
## Multiple R-squared:  0.7262, Adjusted R-squared:  0.7171
## F-statistic: 79.56 on 1 and 30 DF, p-value: 6.113e-10
```

Why ornithologists should use R

@steffilazerte

Reproducible Science

- Keep track of code AND output
- Keep track of data

The screenshot shows an RStudio window titled "My Script". The browser address bar indicates the file path: `file:///home/steffi/Projects/Presentations/2019-0...`. The main content area is divided into two sections: "Reproducibility" and "Data".

Reproducibility

```
DT::datatable(mtcars, options = list(pageLength = 5))
```

Below the code, there is a "Show 5 entries" control and a "Search:" input field.

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21	6	160	110	3.9	2.62	16.46	0	1	4	4
Mazda RX4 Wag	21	6	160	110	3.9	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108	93	3.85	2.32	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360	175	3.15	3.44	17.02	0	0	3	2

Showing 1 to 5 of 32 entries

Navigation: Previous 1 2 3 4 5 6 7 Next

Software

```
devtools::session_info()
```

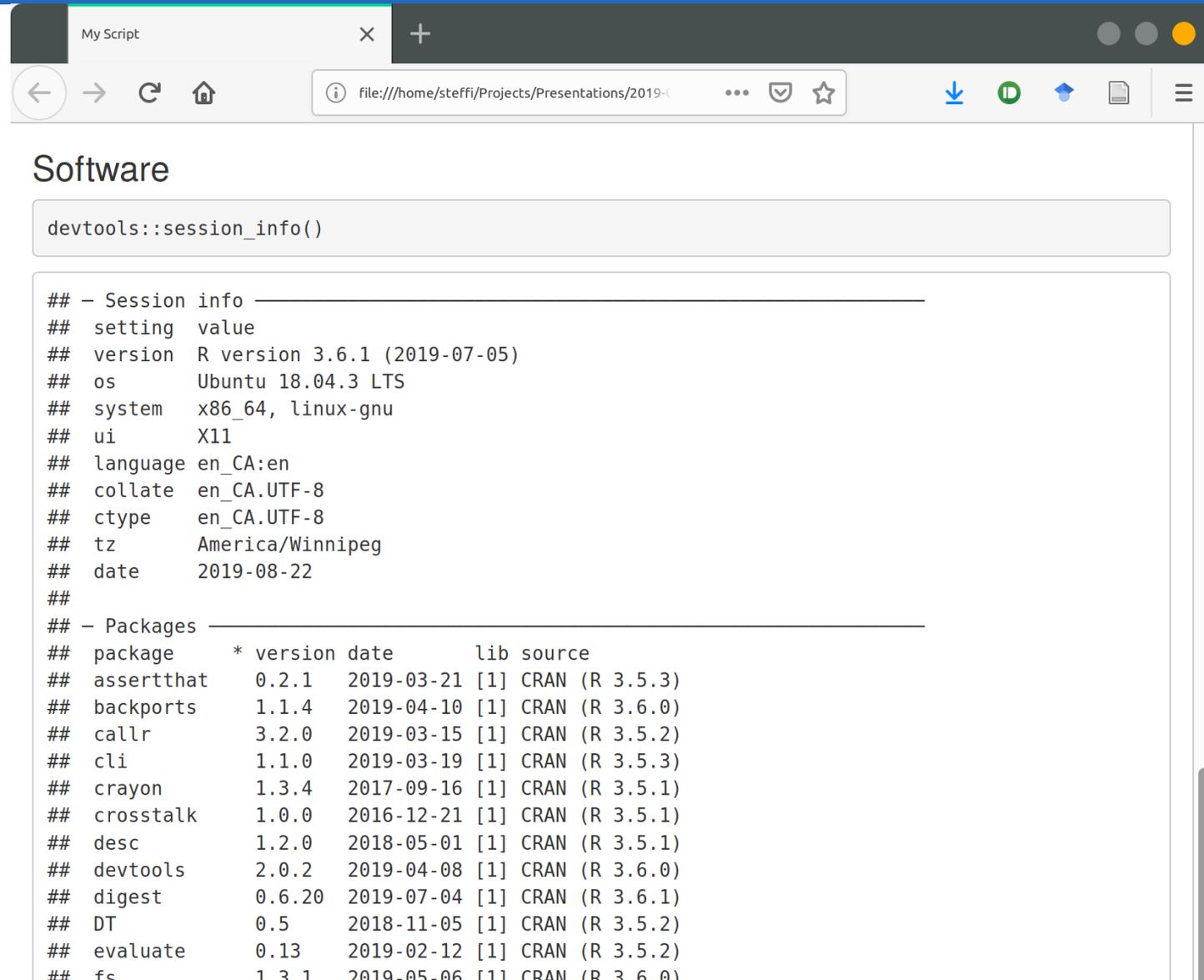
```
## - Session info -  
## setting value  
## version R version 3.6.1 (2019-07-05)  
## os      Ubuntu 18.04.3 LTS
```

Why ornithologists should use R

@steffilazerte

Reproducible Science

- Keep track of code AND output
- Keep track of data
- Keep track of software



```
devtools::session_info()

## - Session info -----
## setting value
## version R version 3.6.1 (2019-07-05)
## os      Ubuntu 18.04.3 LTS
## system  x86_64, linux-gnu
## ui      X11
## language en_CA:en
## collate en_CA.UTF-8
## ctype   en_CA.UTF-8
## tz      America/Winnipeg
## date    2019-08-22
##
## - Packages -----
## package      * version date       lib source
## assertthat   0.2.1   2019-03-21 [1] CRAN (R 3.5.3)
## backports    1.1.4   2019-04-10 [1] CRAN (R 3.6.0)
## callr        3.2.0   2019-03-15 [1] CRAN (R 3.5.2)
## cli          1.1.0   2019-03-19 [1] CRAN (R 3.5.3)
## crayon       1.3.4   2017-09-16 [1] CRAN (R 3.5.1)
## crosstalk    1.0.0   2016-12-21 [1] CRAN (R 3.5.1)
## desc         1.2.0   2018-05-01 [1] CRAN (R 3.5.1)
## devtools     2.0.2   2019-04-08 [1] CRAN (R 3.6.0)
## digest       0.6.20  2019-07-04 [1] CRAN (R 3.6.1)
## DT           0.5     2018-11-05 [1] CRAN (R 3.5.2)
## evaluate     0.13    2019-02-12 [1] CRAN (R 3.5.2)
## fs           1.3.1   2019-05-06 [1] CRAN (R 3.6.0)
```

Why ornithologists should use R

@steffilazerte

Find Data!

- Many online data sources are accessible through R
- Reproducible science includes data sources!

Why ornithologists should use R

@steffilazerte

Find Data!

- Many online data sources are accessible through R
- Reproducible science includes data sources!

Packages

- Observations from [ebird](#) with [auk](#)



Why ornithologists should use R

@steffilazerte

Find Data!

- Many online data sources are accessible through R
- Reproducible science includes data sources!



Packages

- Observations from [ebird](#) with [auk](#)
- Observations from [NatureCounts](#) with [naturecounts](#) (**Here @ 11:00am**)

Why ornithologists should use R

@steffilazerte

Find Data!

- Many online data sources are accessible through R
- Reproducible science includes data sources!



warbleR

Packages

- Observations from [ebird](#) with [auk](#)
- Observations from [NatureCounts](#) with [naturecounts](#) ([Here @ 11:00am](#))
- Recordings of vocalizations from [xeno-canto](#) with [warbleR](#)

Why ornithologists should use R

@steffilazerte

Find Data!

- Many online data sources are accessible through R
- Reproducible science includes data sources!



warbleR

Packages

- Observations from [ebird](#) with [auk](#)
- Observations from [NatureCounts](#) with [naturecounts](#) ([Here @ 11:00am](#))
- Recordings of vocalizations from [xeno-canto](#) with [warbleR](#)
- Weather data from [Environment and Climate Change Canada](#) with [weathercan](#)



Why ornithologists should use R

@steffilazerte

Find Data!

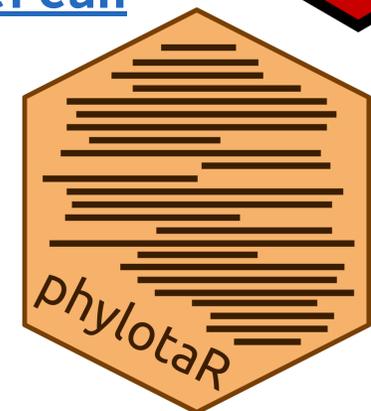
- Many online data sources are accessible through R
- Reproducible science includes data sources!

Packages

- Observations from [ebird](#) with [auk](#)
- Observations from [NatureCounts](#) with [naturecounts](#) ([Here @ 11:00am](#))
- Recordings of vocalizations from [xeno-canto](#) with [warbleR](#)
- Weather data from [Environment and Climate Change Canada](#) with [weathercan](#)
- DNA Sequences from [GenBank](#) with [phylotaR](#)



warbleR



Why ornithologists should use R

@stefilazerte

Find Data!

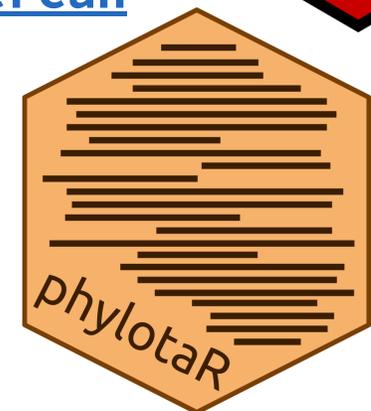
- Many online data sources are accessible through R
- Reproducible science includes data sources!

Packages

- Observations from [ebird](#) with [auk](#)
- Observations from [NatureCounts](#) with [naturecounts](#) ([Here @ 11:00am](#))
- Recordings of vocalizations from [xeno-canto](#) with [warbleR](#)
- Weather data from [Environment and Climate Change Canada](#) with [weathercan](#)
- DNA Sequences from [GenBank](#) with [phylotaR](#)
- Taxonomic data with [taxize](#)



warbleR



Why ornithologists should use R

@steffilazerte

Find Data!

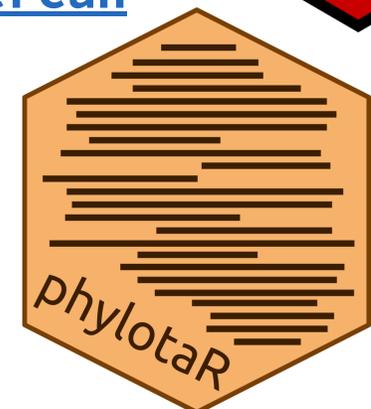
- Many online data sources are accessible through R
- Reproducible science includes data sources!

Packages

- Observations from [ebird](#) with [auk](#)
- Observations from [NatureCounts](#) with [naturecounts](#) ([Here @ 11:00am](#))
- Recordings of vocalizations from [xeno-canto](#) with [warbleR](#)
- Weather data from [Environment and Climate Change Canada](#) with [weathercan](#)
- DNA Sequences from [GenBank](#) with [phylotaR](#)
- Taxonomic data with [taxize](#)
- IUCN Red Lists with [rredlist](#)



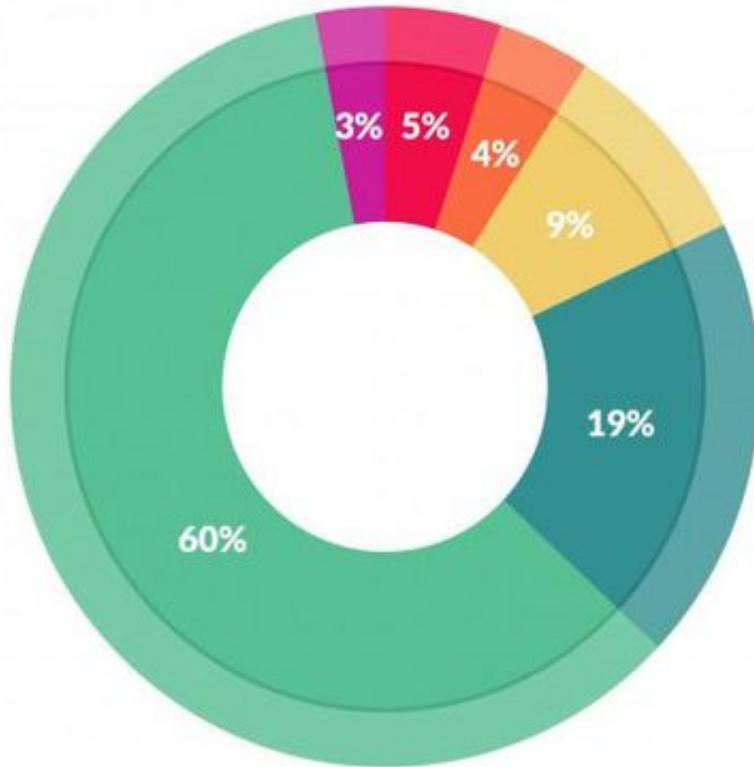
warbleR



Why ornithologists should use R

@steffilazerte

Dealing with Data



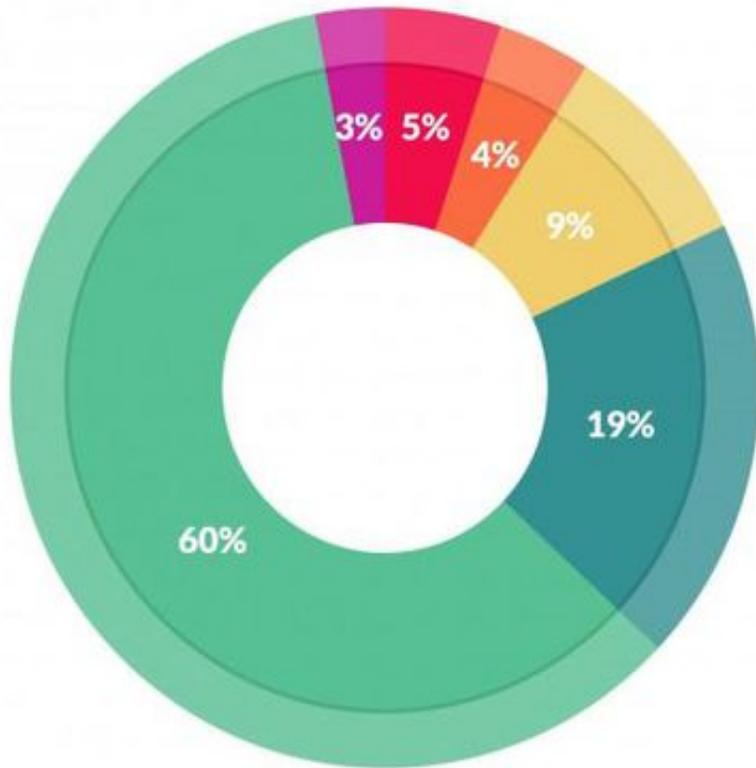
What data scientists spend the most time doing

- Building training sets: 3%
- Cleaning and organizing data: 60%
- Collecting data sets; 19%
- Mining data for patterns: 9%
- Refining algorithms: 4%
- Other: 5%

Why ornithologists should use R

@steffilazerte

Dealing with Data



What data scientists spend the most time doing

● Building training sets: 3%

● Cleaning and organizing data: 60%

● Collecting data sets; 19%

● Mining data for patterns: 9%

● Refining algorithms: 4%

● Other: 5%

Dealing with Data

- Reproducible!
- Cleaning
 - Fix typos
 - Fix/explore odd/missing values
- Filtering
- Summarizing
- Transforming
- Exploring

Why ornithologists should use R

@steffilazerte

Dealing with Data

- Reproducible!
- Cleaning
 - Fix typos
 - Fix/explore odd/missing values
- Filtering
- Summarizing
- Transforming
- Exploring

Packages to Use

- Base R (i.e. no special packages)
- [data.table](http://r-datatable.com) (<http://r-datatable.com>)
- [tidyverse](http://tidyverse.org) (<http://tidyverse.org>)
 - Suite of packages
 - Learn more: [R for Data Science](#)



Why ornithologists should use R

@steffilazerte

100's of Specialized packages

For example...

Why ornithologists should use R

@steffilazerte

100's of Specialized packages

For example...

- Phylogenetic comparative analyses [adephylo](#)

Why ornithologists should use R

@steffilazerte

100's of Specialized packages

For example...

- Phylogenetic comparative analyses [adephylo](#)
- Bioacoustic analyses with [seewave](#)

seewave 

Why ornithologists should use R

@steffilazerte

100's of Specialized packages

For example...

- Phylogenetic comparative analyses [adephylo](#)
- Bioacoustic analyses with [seewave](#)
- Automatic detection of vocalizations with [monitoR](#)

The logo for the R package 'seewave'. The word 'seewave' is written in a bold, lowercase, sans-serif font. The 'e' is notably larger than the other letters. To the right of the text is a green wavy line that resembles a sound wave or a stylized 'e'.

Why ornithologists should use R

@steffilazerte

100's of Specialized packages

For example...

- Phylogenetic comparative analyses [adephylo](#)
- Bioacoustic analyses with [seewave](#)
- Automatic detection of vocalizations with [monitoR](#)
- Animal home ranges with [adehabitatHR](#)

seewave The logo for the 'seewave' R package, featuring the word 'seewave' in a bold, black, sans-serif font, followed by a green wavy line that resembles a sound wave.

Why ornithologists should use R

@steffilazerte

100's of Specialized packages

For example...

- Phylogenetic comparative analyses [adephylo](#)
- Bioacoustic analyses with [seewave](#)
- Automatic detection of vocalizations with [monitoR](#)
- Animal home ranges with [adehabitatHR](#)
- Hierarchical Bayesian modelling of Breeding Bird Survey data with [bbsBayes](#) ([Here @ 11:30am](#))

seewave 



Why ornithologists should use R

@steffilazerte

100's of Specialized packages

For example...

- Phylogenetic comparative analyses [adephylo](#)
- Bioacoustic analyses with [seewave](#)
- Automatic detection of vocalizations with [monitoR](#)
- Animal home ranges with [adehabitathR](#)
- Hierarchical Bayesian modelling of Breeding Bird Survey data with [bbsBayes](#) ([Here @ 11:30am](#))
- Systematic reviews with [litsearchr](#) ([Here @ 1:30pm](#))

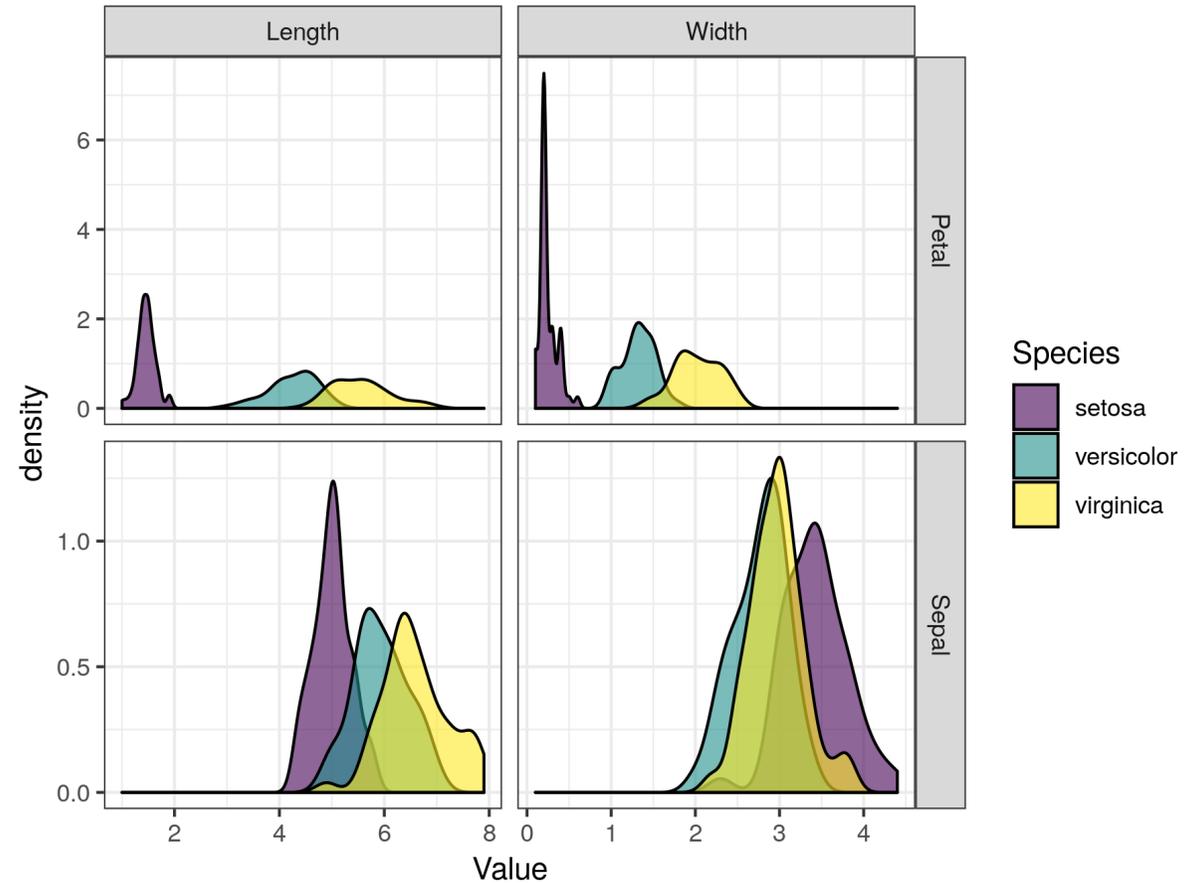
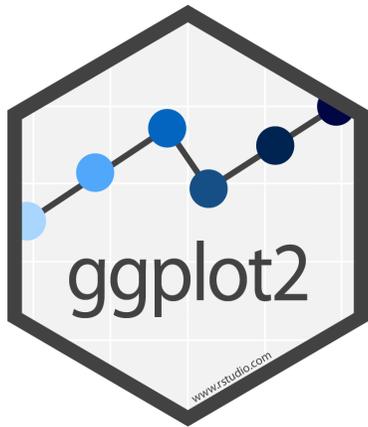
The logo for the 'seewave' R package. It features the word 'seewave' in a bold, black, sans-serif font. A green wavy line is positioned at the end of the word, resembling a sound wave.

Why ornithologists should use R

@steffilazerte

Dissemination and Visualizations

- Beautiful plots with [ggplot2](#)

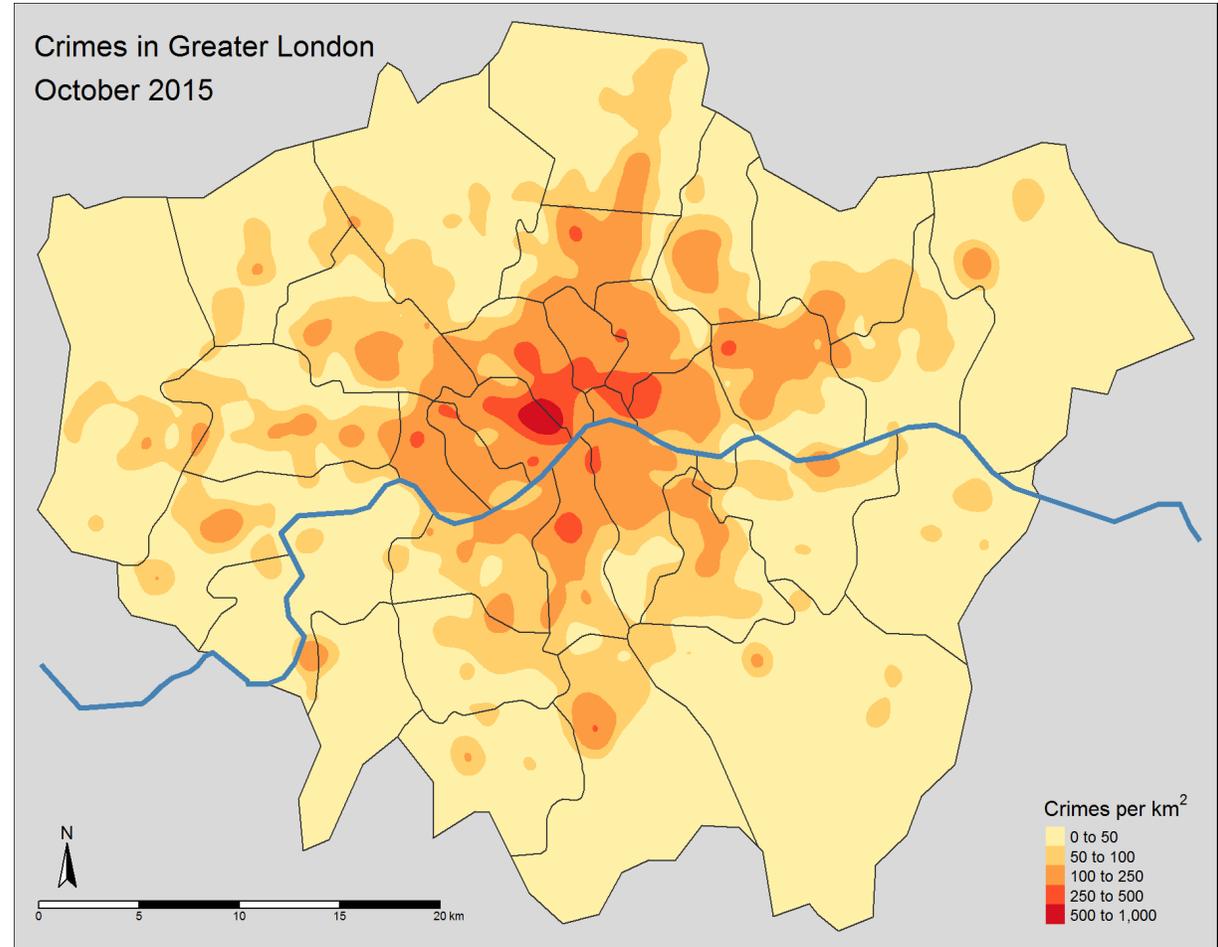
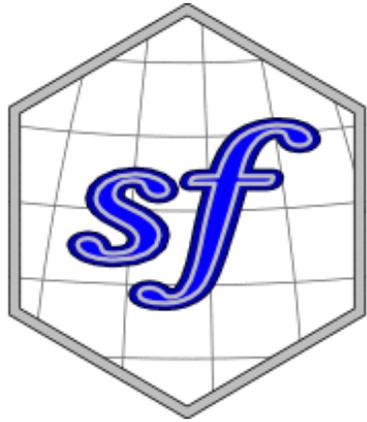
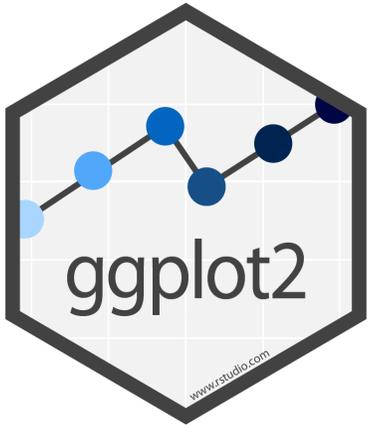


Why ornithologists should use R

@steffilazerte

Dissemination and Visualizations

- Beautiful plots with [ggplot2](#)
- Complex maps with [sf](#), [tmap](#)



Why ornithologists should use R

@steffilazerte

Dissemination and Visualizations

- Beautiful plots with [ggplot2](#)
- Complex maps with [sf](#), [tmap](#)
- Interactive visualizations with [shiny](#)



Instructions: ?

Summary over time
 Cumulative Instant

Select Individual
All

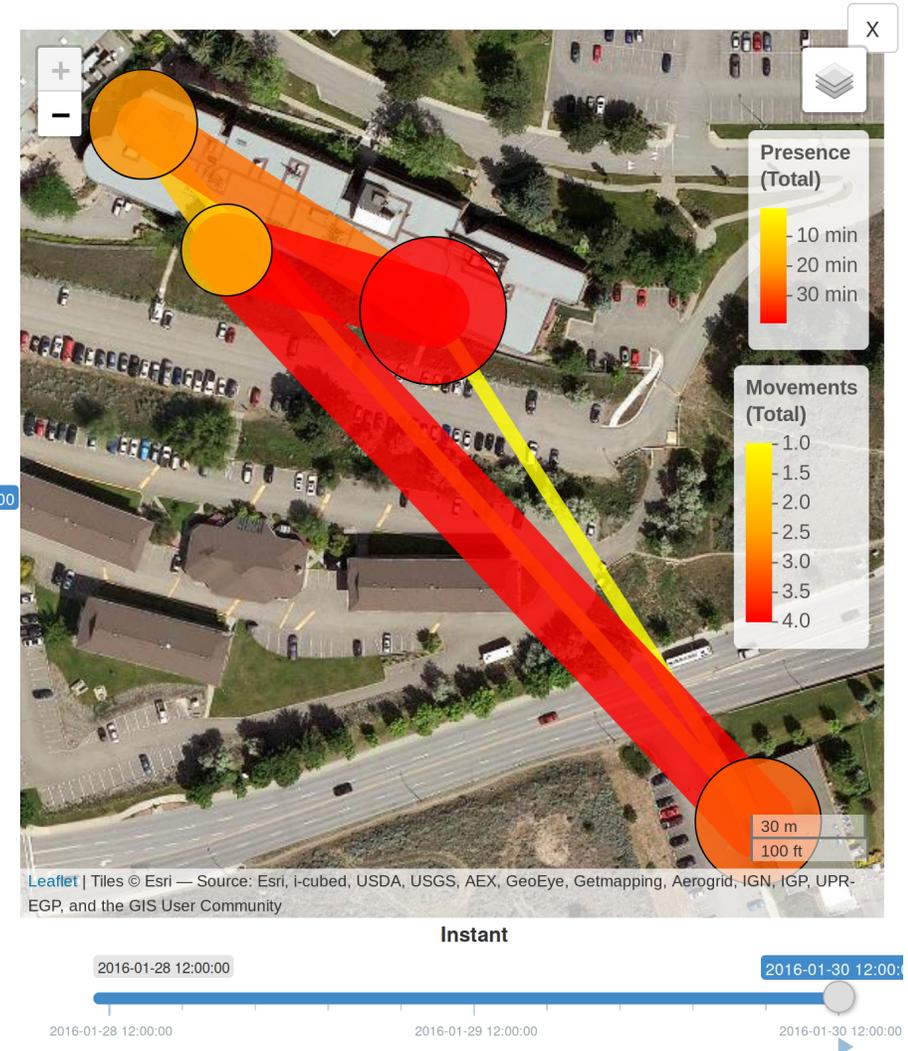
Summary type
 Total sum
 Average sum per individual

Time Range
2016-01-28 07:00:00 2016-01-29 18:00:00

Animation options
Resolution
 5 min 15 min 30 min 1 hr
 3 hr 6 hr 12 hr 24 hr

Animation speed
0% 50% 100%

Show sunrise/sunset?
 Yes No



Finch movements

Find More Packages!

- Not an exhaustive list!
- Check out package collections
 - [metaverse](https://rmetaverse.github.io/) (<https://rmetaverse.github.io/>)
 - [ropensci](https://ropensci.org/) (<https://ropensci.org/>)
 - [tidyverse](https://tidyverse.org/) (<https://tidyverse.org/>)
- Look in journals, i.e. Methods in Ecology and Evolution
- Look in methods sections of articles

Find More Packages!

- Not an exhaustive list!
- Check out package collections
 - [metaverse](https://rmetaverse.github.io/) (<https://rmetaverse.github.io/>)
 - [ropensci](https://ropensci.org/) (<https://ropensci.org/>)
 - [tidyverse](https://tidyverse.org/) (<https://tidyverse.org/>)
- Look in journals, i.e. Methods in Ecology and Evolution
- Look in methods sections of articles

Always cite packages and package versions!

Symposium: R for Ornithologists

@steffilazerte

Stay tuned for 6 more R-related presentations

2:00pm Round-Table Discussion on Ornithological perspectives on R

Symposium: R for Ornithologists

@steffilazerte

Stay tuned for 6 more **R**-related presentations

2:00pm Round-Table Discussion on Ornithological perspectives on R



Thank you!

Thanks to Denis LePage for help on the French version



@steffilazerte



steffilazerte



steffilazerte.ca



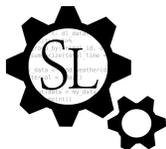
sel@steffilazerte.ca



Presentation Available: <https://steffilazerte.github.io/Presentations/>

Slides created with the R package [xaringan](#), using [remark.js](#), [knitr](#), and [R Markdown](#)

Dr. Steffi LaZerte



Analysis and Data Tools for Science

Compiled on 2019-08-28