Symposium

R pour ornithologue

Points de vue des utilisateuRs aux les programmeuRs

R for Ornithologists

Perspectives from useRs to programmeRs (to birdeRs)

10:30-14:15 Salle des Plaines II

SCO-SOC 2019

Stefanie E. LaZerte

BRANDON UNIVERSITY

R for Ornithologists

How R can benefit the study of Ornithology

@steffilazerte steffilazerte steffilazerte.ca sel@steffilazerte.ca

What is R?

What is R?

A statistical programming language and environment

(free and open source!)

What is R?

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!

R is hard

@steffilazerte

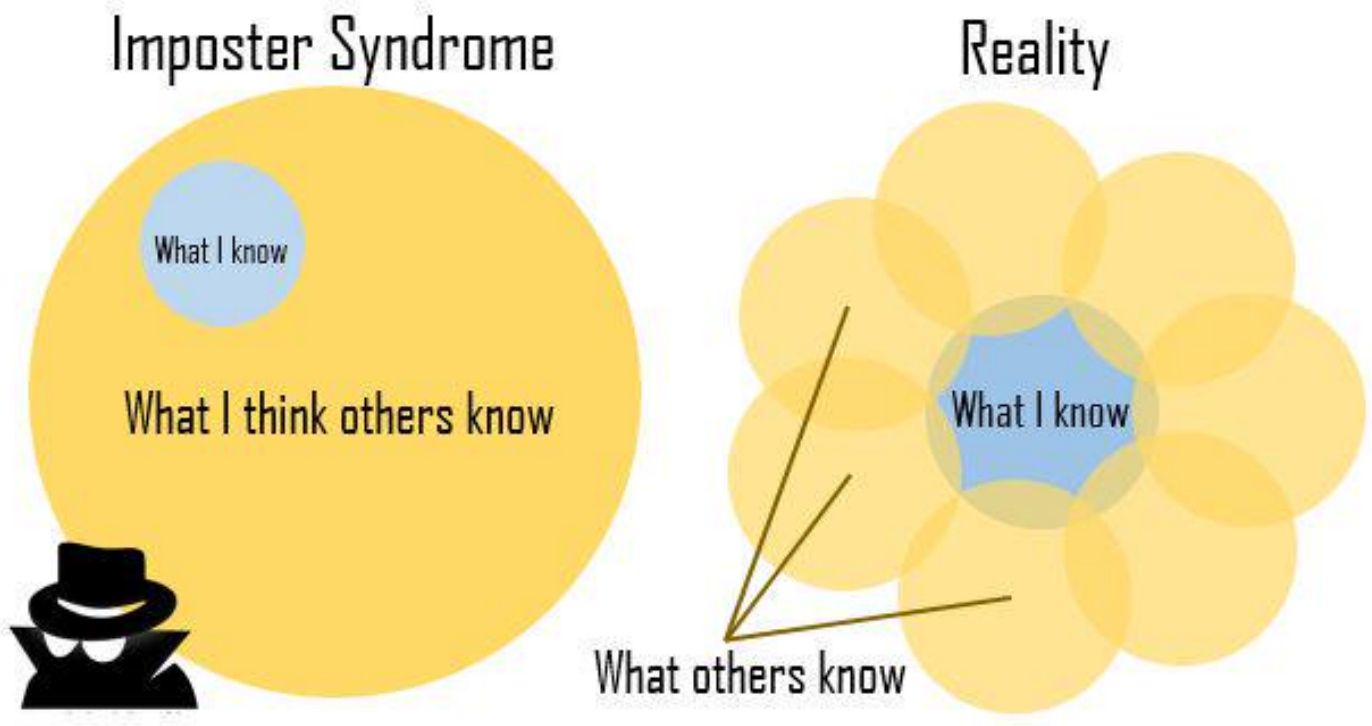
```
# Get in circle around city
circle <- data.frame()
cutoff <- 10
for(i in unique(gps$region)) {
  n <- nrow(gps[gps$region == i,])  ##number of IDs
  if(i == "wil") tmp <- geocode("Williams Lake, Canada")
  if(i == "kam") tmp <- geocode("Kamloops, Canada")
  if(i == "kel") tmp <- geocode("Kelowna, Canada")
  temp <- data.frame()
  for(a in 1:n){
    if(a <= cutoff) temp <- rbind(temp, gcDestination(lon = tmp$lon,
                                                      lat = tmp$lat,
                                                      bearing = (a*(360/(cutoff))-360/(cutoff)),
                                                      dist = 20,
                                                      dist.units = "km",
                                                      model = "WGS84"))
    if(a > cutoff) temp <- rbind(temp, gcDestination(lon = tmp$lon,
                                                     lat = tmp$lat,
                                                     bearing = ((a-cutoff)*(360/(max(table(gps$region
))-10))-360/(max(table(gps$region))-cutoff)),
                                                     dist = 35,
                                                     dist.units = "km",
                                                     model = "WGS84"))
  }
  circle <- rbind(circle, cbind(temp,
                                region = i,
                                hab = gps$hab[gps$region == i],
                                spl = gps$spl.orig[gps$region == i],
```

Impost**R** Syndrome

@steffilazerte

ImpostR Syndrome

@steffilazerte

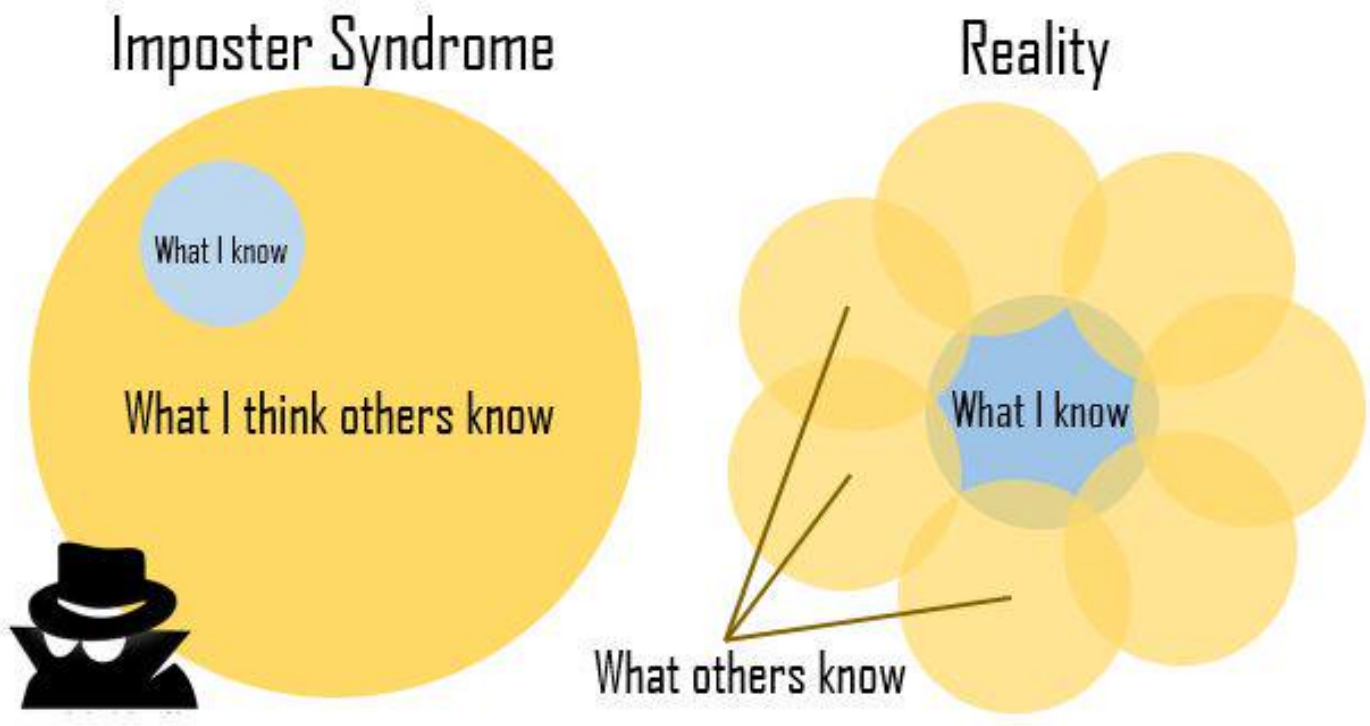


David Whittaker

Impost R Syndrome

ImpostR Syndrome

@steffilazerte



David Whittaker

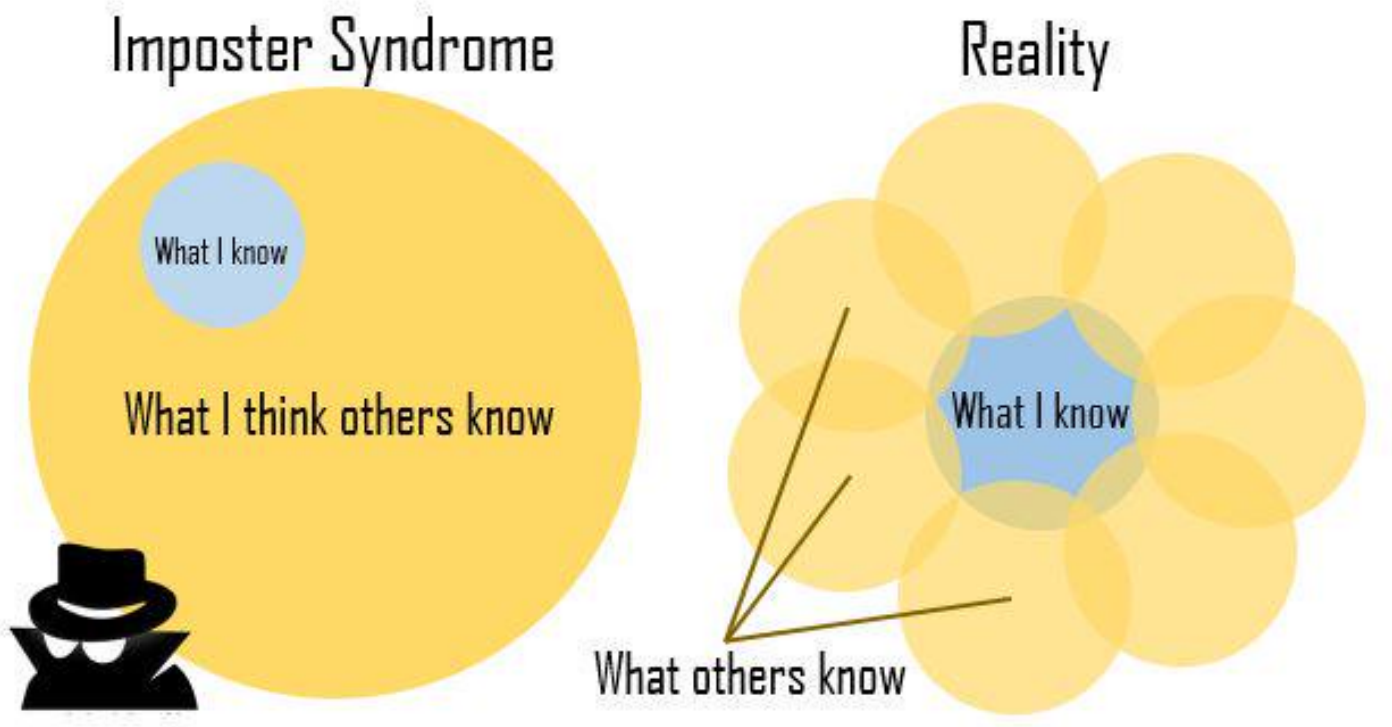
Impost R Syndrome

Moral of the story?

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

ImpostR Syndrome

@steffilazerte



David Whittaker

Impost R Syndrome

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

Ornithologists and R

@steffilazerte

What I **am not** going to do

- Teach you R
- Talk about statistics

Ornithologists and R

@steffilazerte

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

Ornithologists and R

@steffilazerte

What I **am not** going to do

- Teach you R
- Talk about statistics

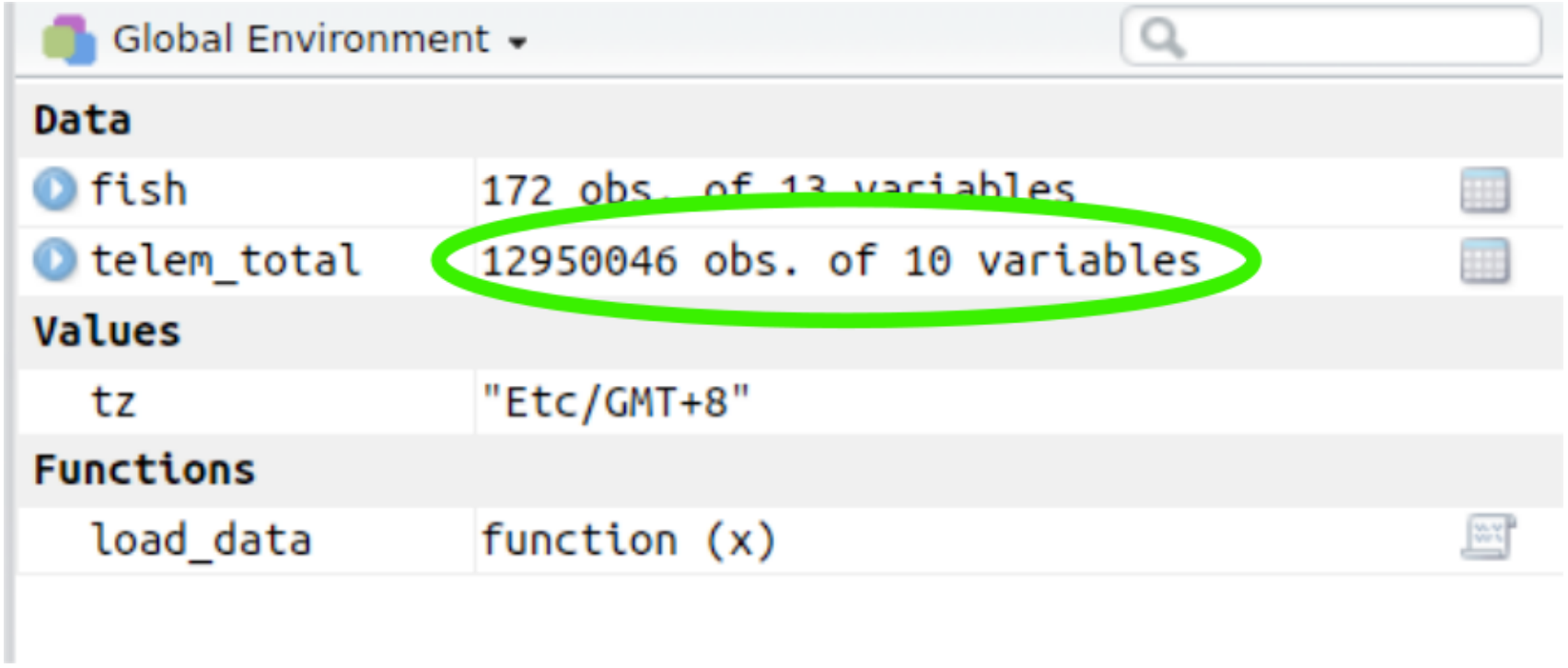
What I **am** going to do

- 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!



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)
```

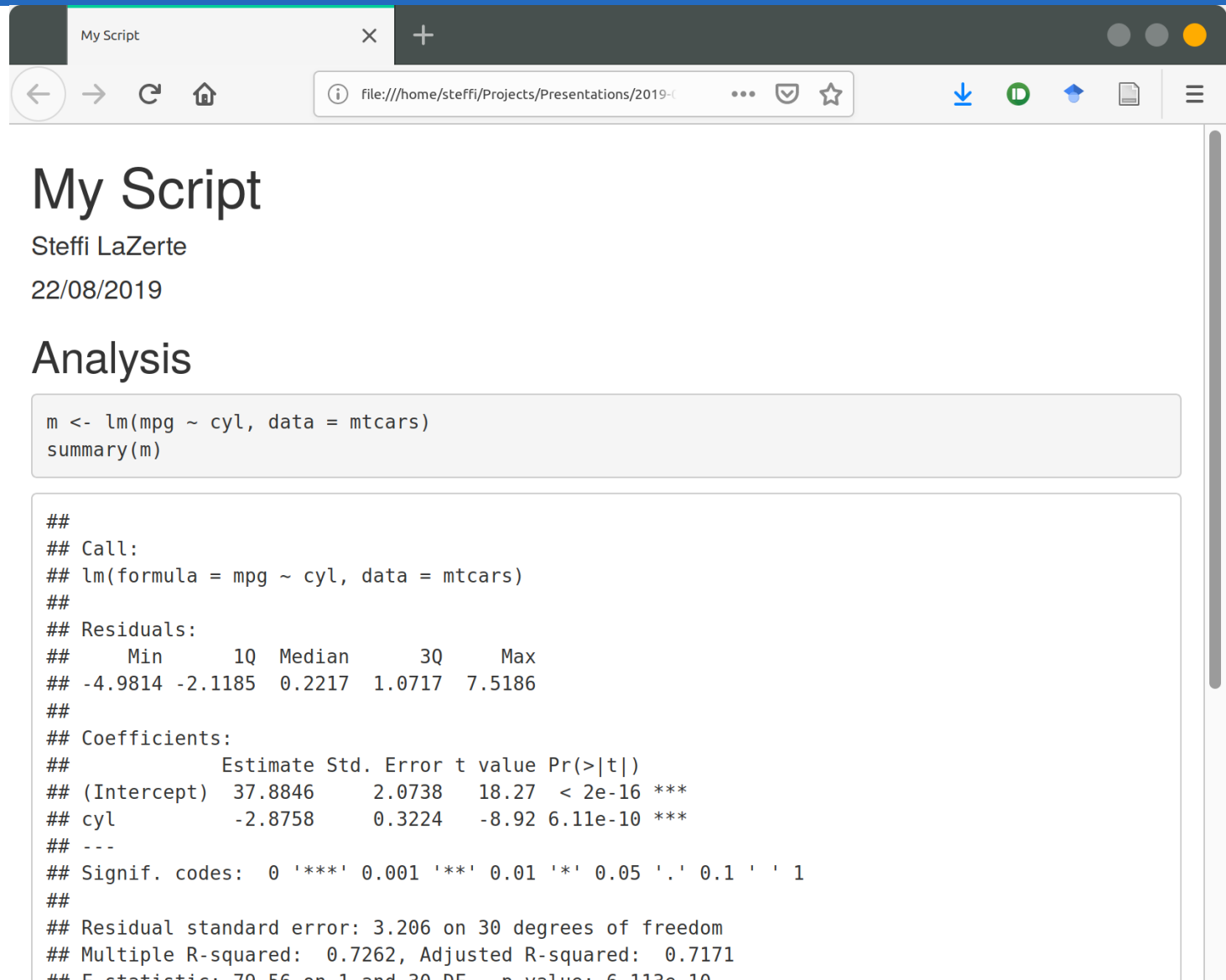
- Scripts can be compiled into pdf/html reports with **`rmarkdown`** and **`knitr`**
  (In RStudio: `File > Compile Report`)

Why ornithologists should use R

@steffilazerte

Reproducible Science

- Keep track of code AND output

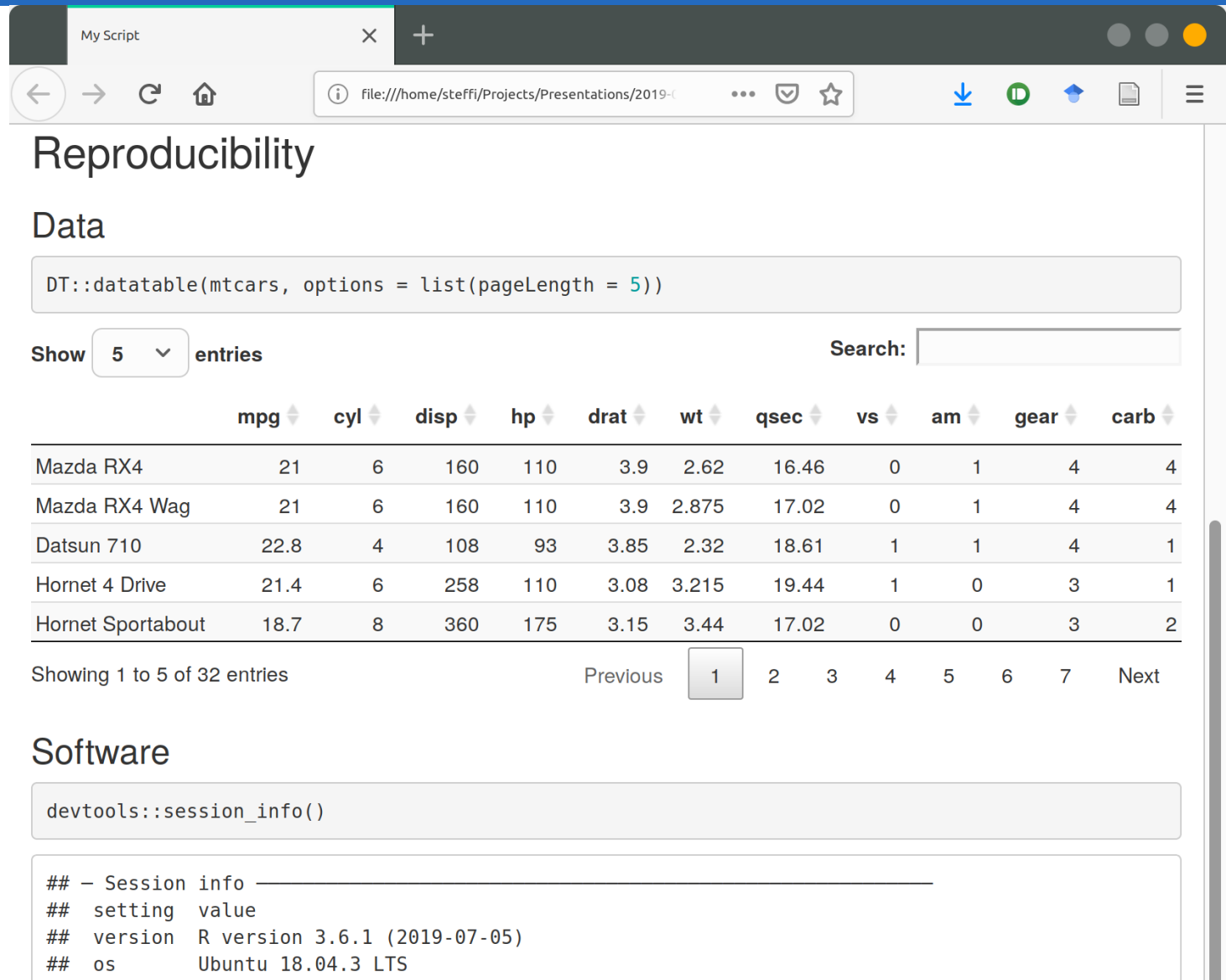


Why ornithologists should use R

@steffilazerte

Reproducible Science

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

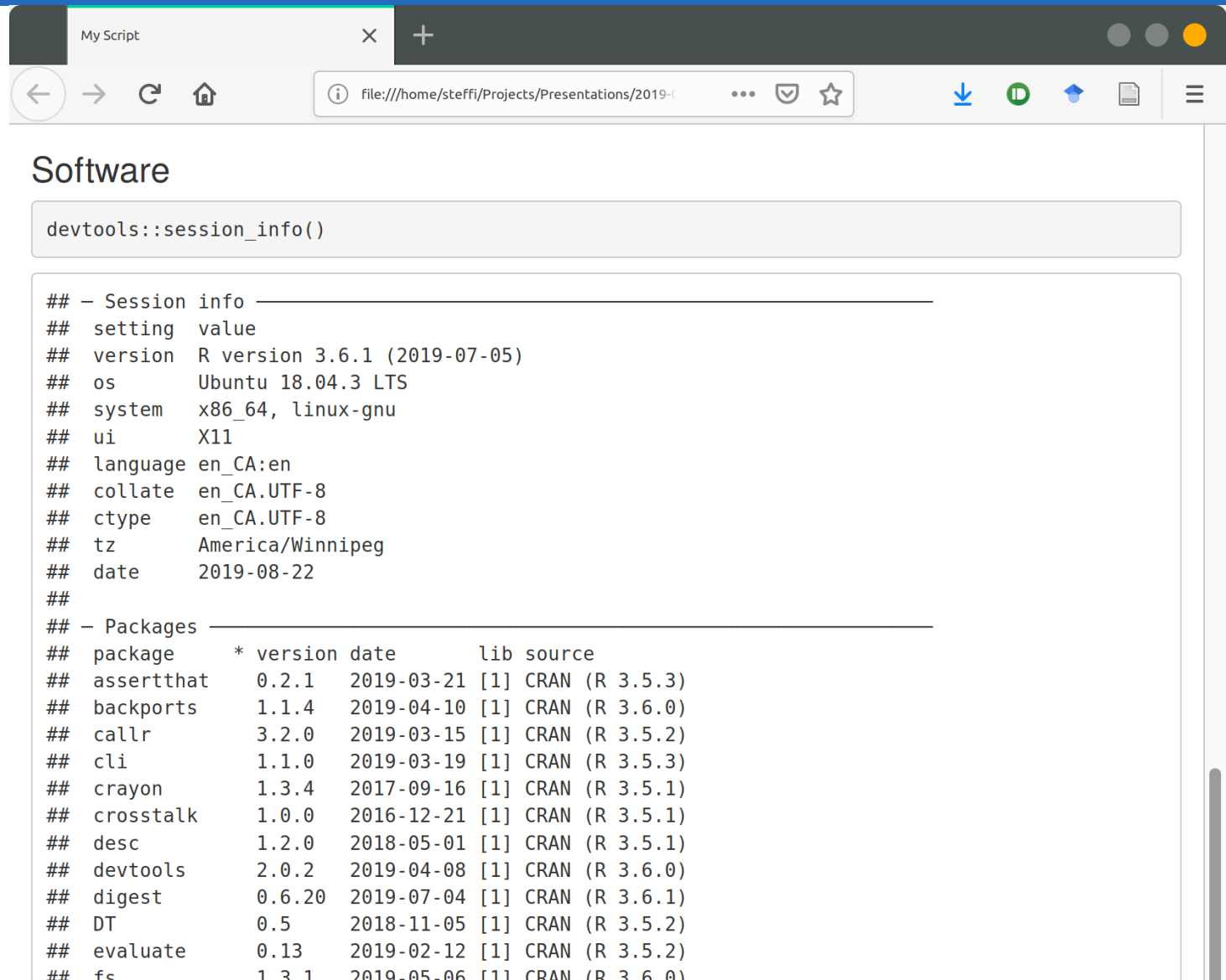


Why ornithologists should use R

@steffilazerte

Reproducible Science

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



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!

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!

warbleR

- 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**

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**
- DNA Sequences from GenBank with **phylotaR**

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`**

Why ornithologists should use R

@steffilazerte

Find Data!

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

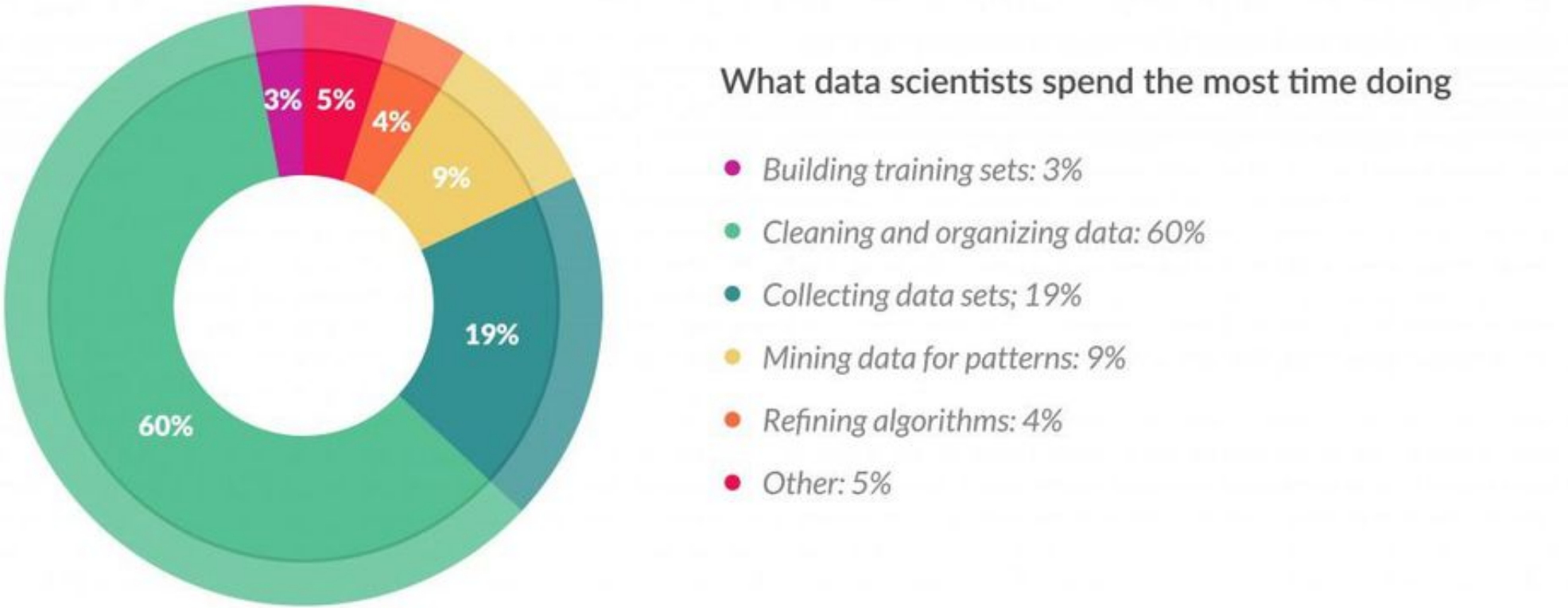
Packages

- 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**

Why ornithologists should use R

@steffilazerte

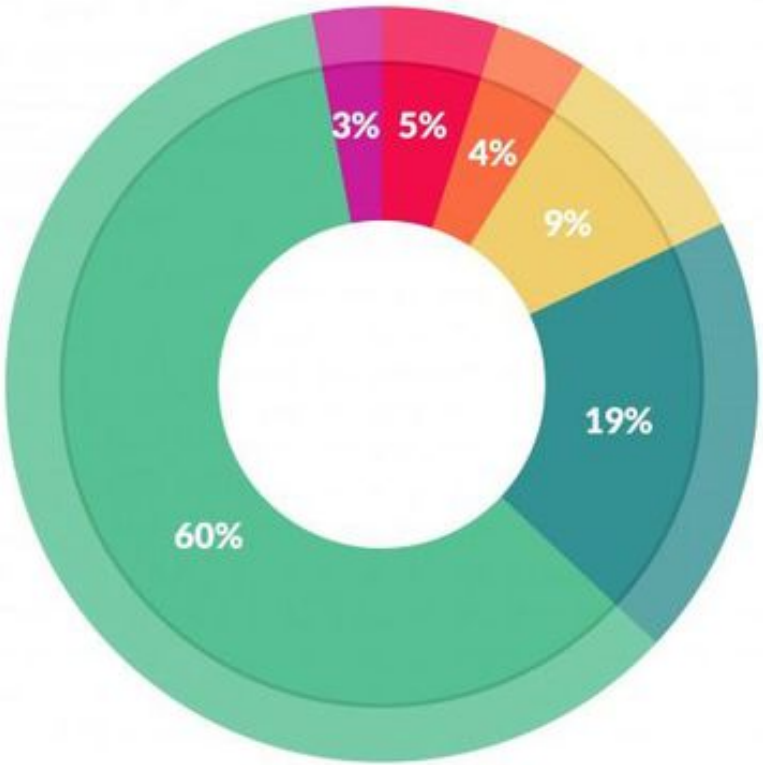
Dealing with Data



Forbes.com

Why ornithologists should use R

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%*

Forbes.com

Why ornithologists should use R

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)
- `tidyverse` (http://tidyverse.org)
  - Suite of packages
  - Learn more: R for Data Science

Why ornithologists should use R

100's of Specialized packages

For example...

Why ornithologists should use R

100's of Specialized packages

For example...

- Phylogenetic comparative analyses **`adephylo`**

Why ornithologists should use R

100's of Specialized packages

For example...

- Phylogenetic comparative analyses **`adephylo`**
- Bioacoustic analyses with **`seewave`**

Why ornithologists should use R

100's of Specialized packages

For example...

- Phylogenetic comparative analyses **`adephylo`**
- Bioacoustic analyses with **`seewave`**
- Automatic detection of vocalizations with **`monitoR`**

Why ornithologists should use R

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`**

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...

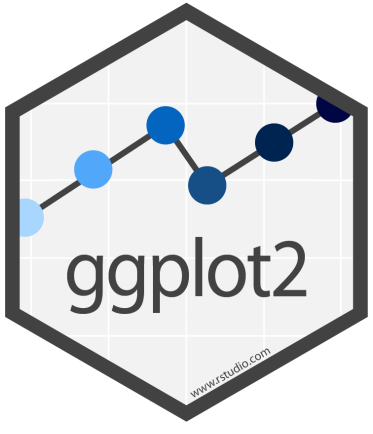
- Phylogenetic comparative analyses **`adephylo`**
- Bioacoustic analyses with **`seewave`**
- Automatic detection of vocalizations with **`monitoR`**
- Animal home ranges with **`adehabitatHR`**
- Hierarchical Bayesian modelling of Breeding Bird Survey data with **`bbsBayes`** (**Here @ 11:30am**)
-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

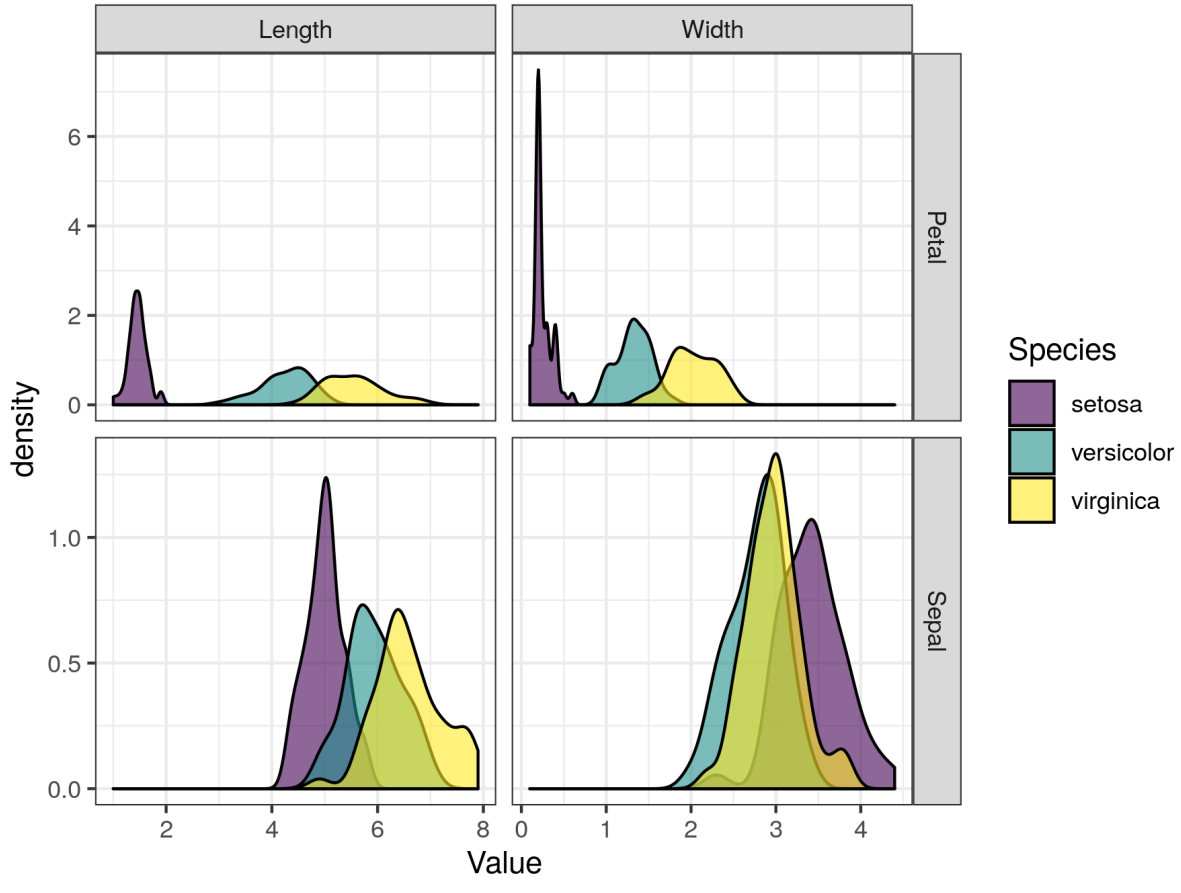
Why ornithologists should use R

@steffilazerte

Dissemination and Visualizations

- Beautiful plots with **ggplot2**



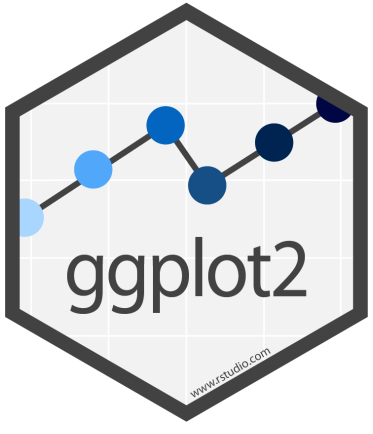


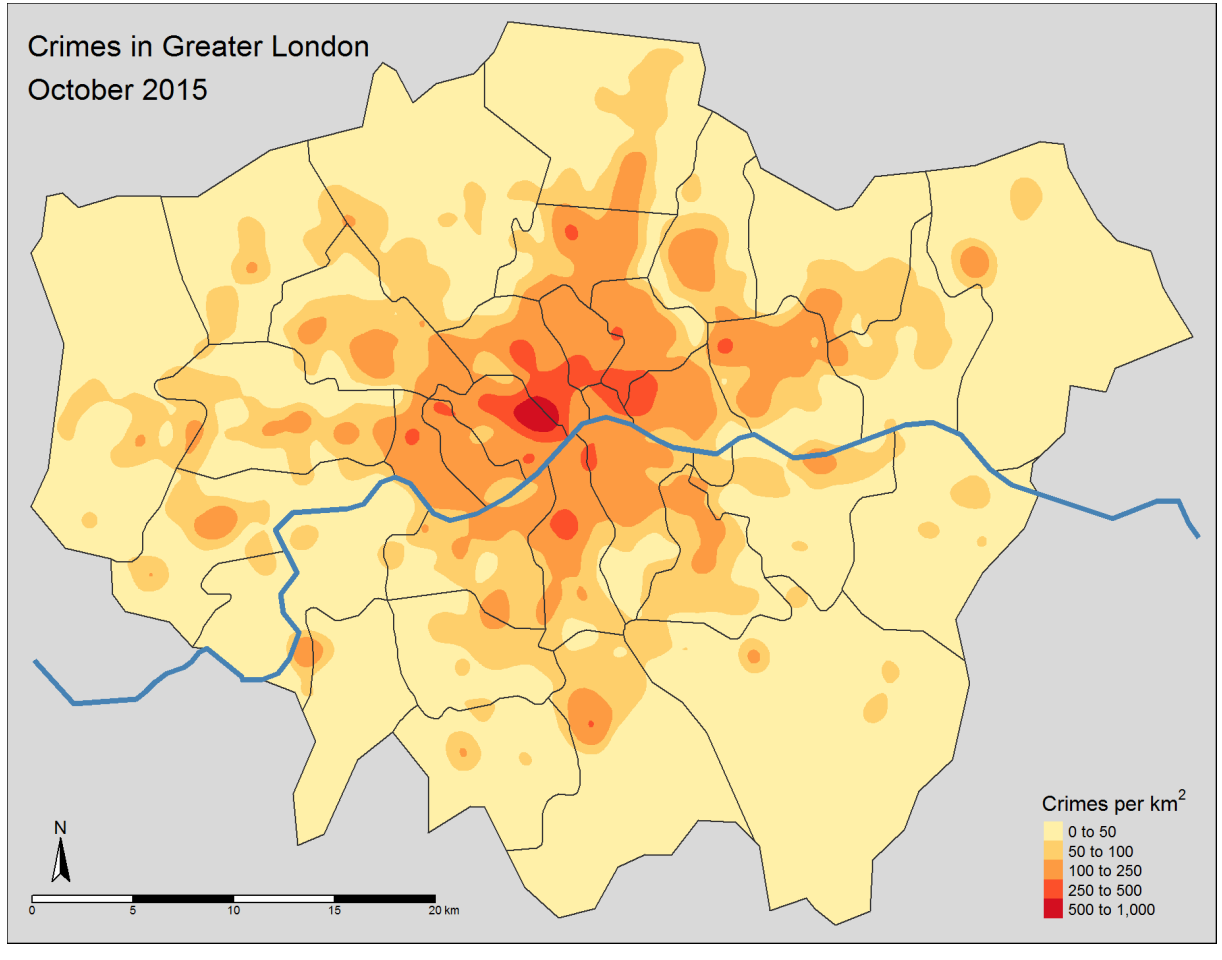
Why ornithologists should use R

@steffilazerte

Dissemination and Visualizations

- Beautiful plots with **`ggplot2`**
- Complex maps with **`sf`**, **`tmap`**



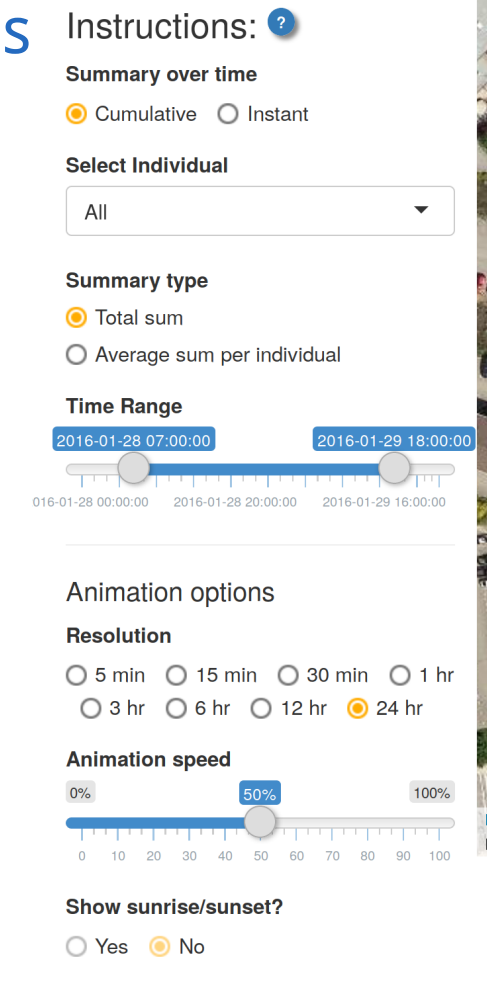


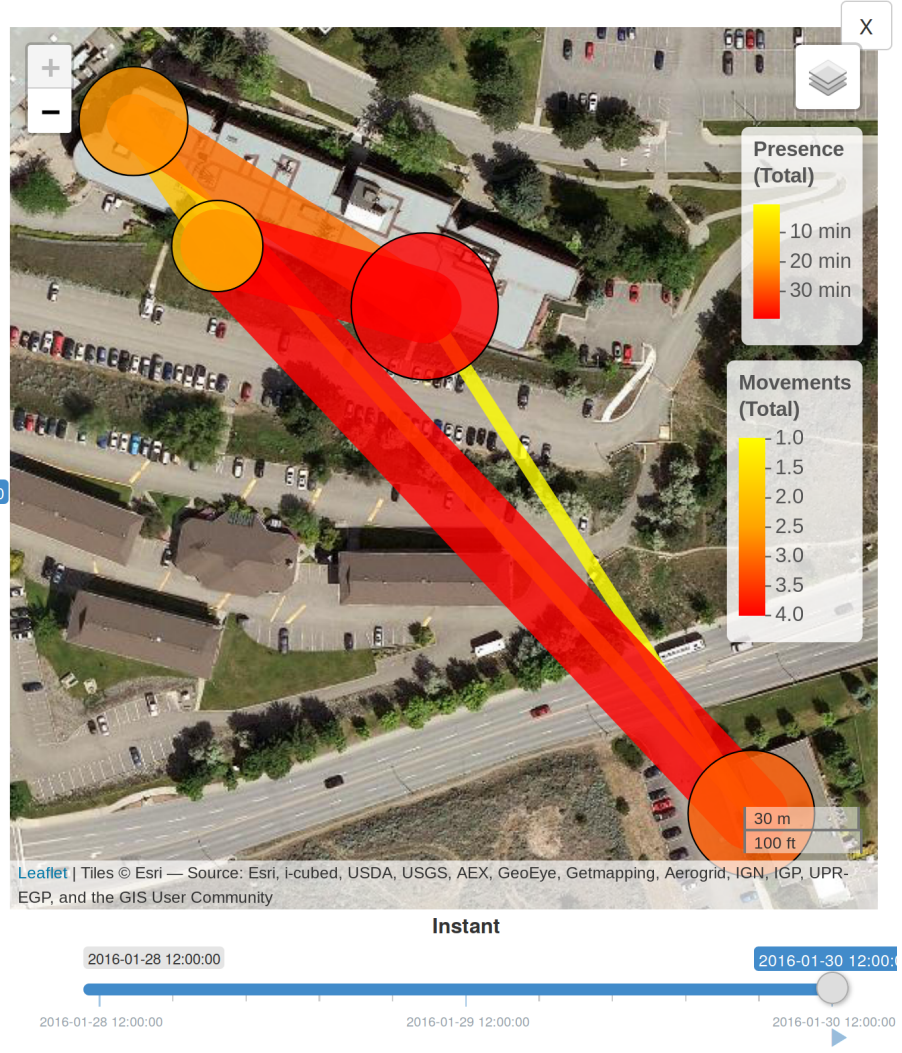
Why ornithologists should use R

@steffilazerte

Dissemination and Visualizations

- Beautiful plots with **ggplot2**
- Complex maps with **sf**, **tmap**
- Interactive visualizations with **shiny**





Finch movements

Why ornithologists should use R

Find More Packages!

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

Why ornithologists should use R

Find More Packages!

- Not an exhaustive list!
- Check out package collections
  - metaverse (https://rmetaverse.github.io/)
  - ropensci (https://ropensci.org/)
  - tidyverse (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

BRANDON UNIVERSITY

Thank you!

@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 2019-08-26