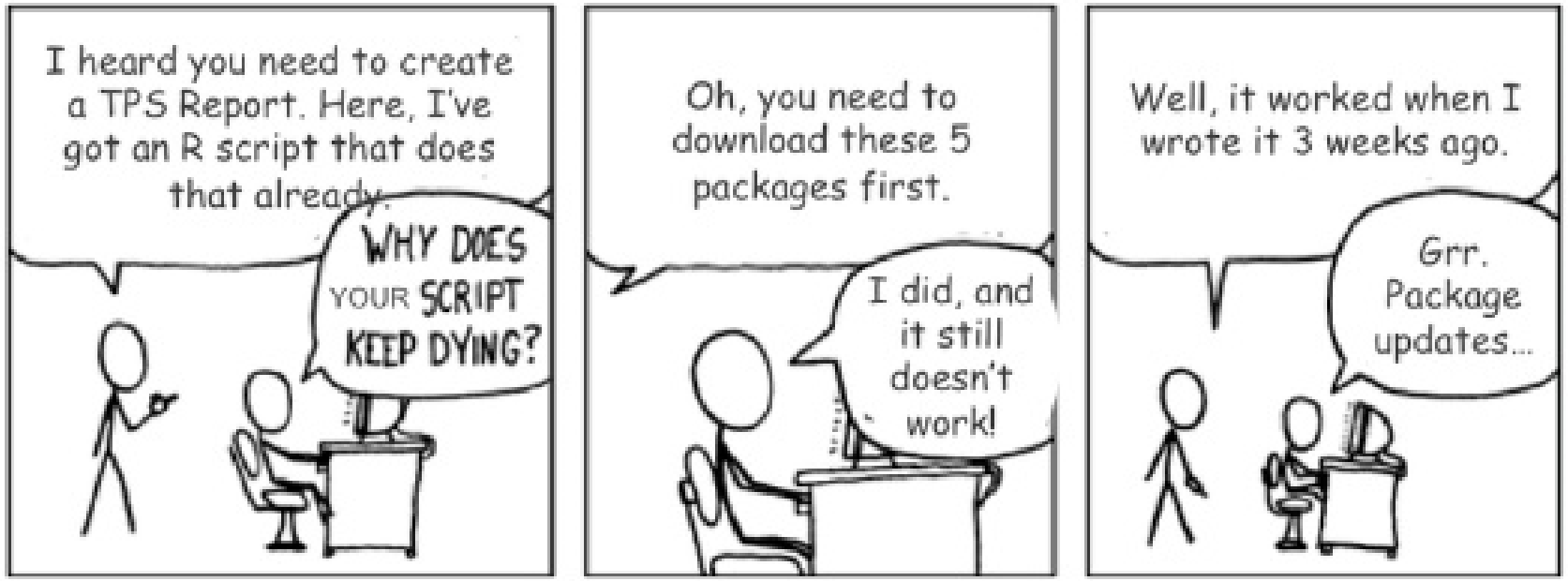
Session 3

Lunch R workshop 2018.03.20

Max Lindmark & Philip Jacobson

lunchR



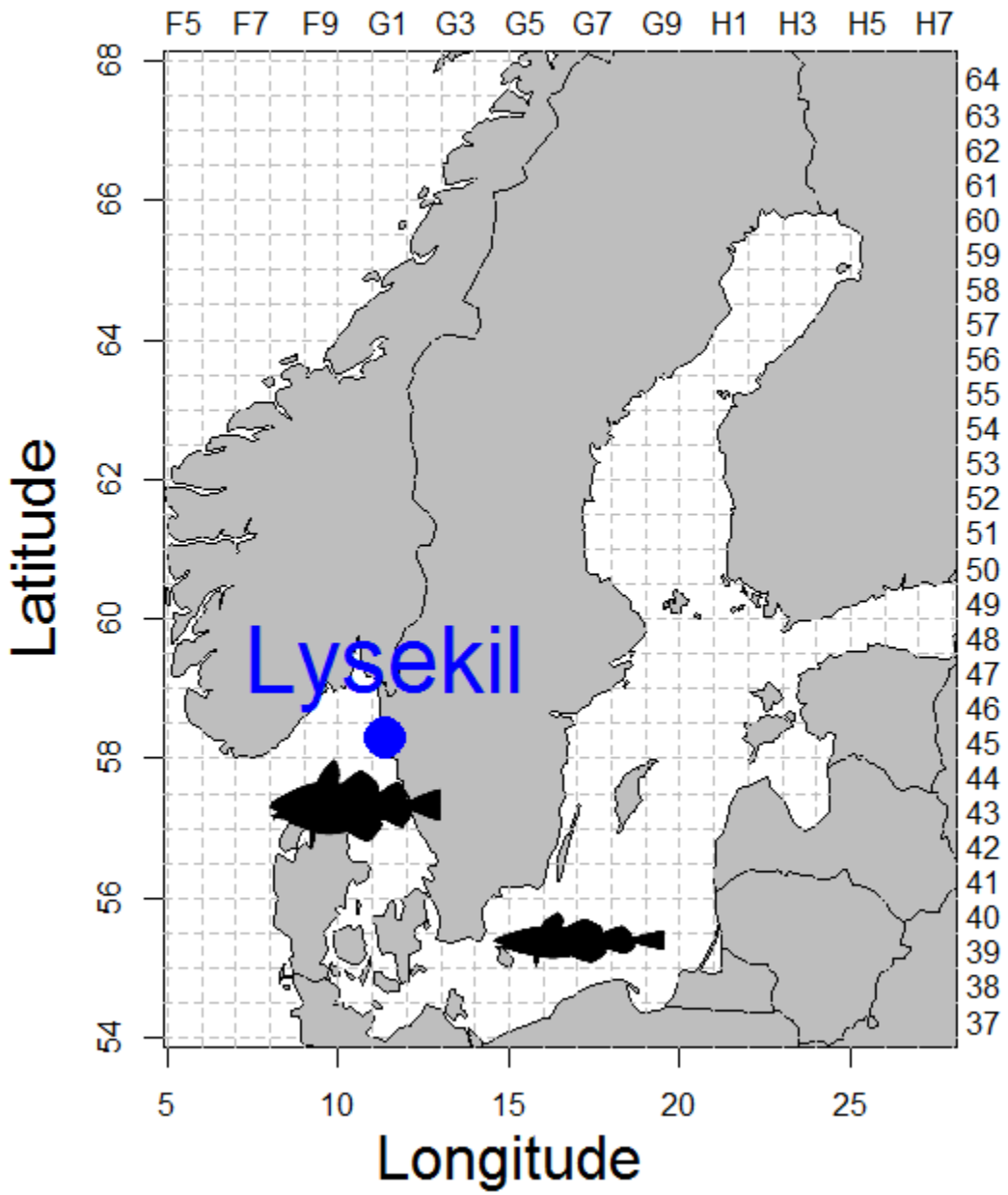
https://www.slideshare.net/continuumio/open-data-science-with-r-and-anaconda

lunchR

Sessions

1. Intro to why R, data formats and plotting styles in R
2. Digging deeper into data exploration in R using base R and ggplot; focus on different graph styles
3. **Tips and tricks in R – format your data for easy plotting**
4. Creating publish-quality figures

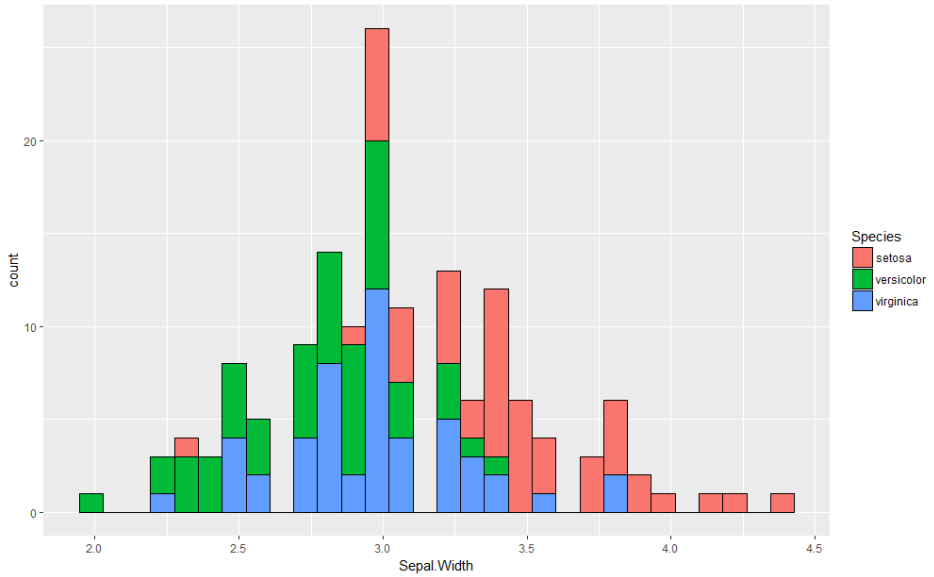
lunchR



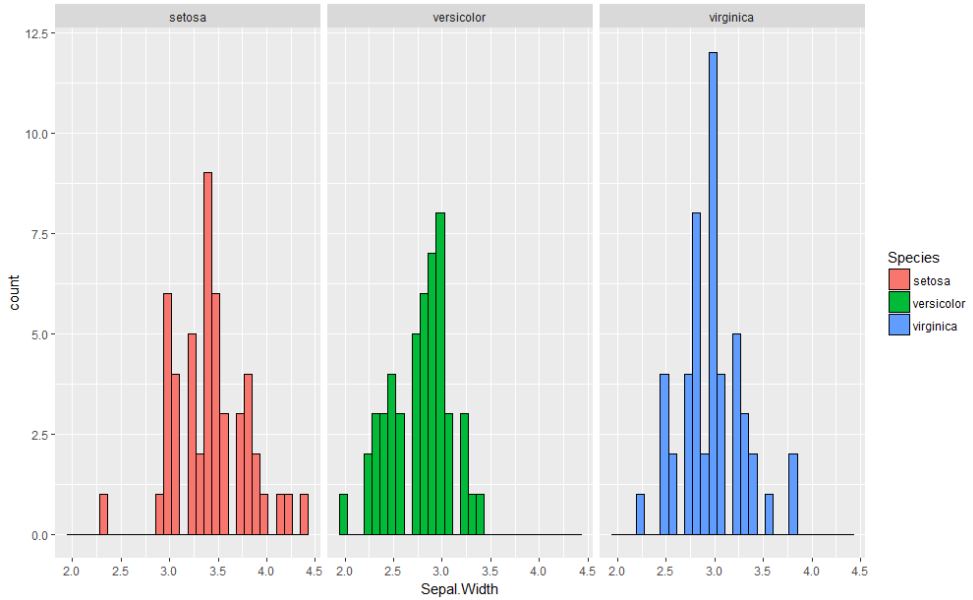
Session 2: Some of you had problems creating the final map

- Windows 10 users.
- Apparently some package in the tidyverse “masked” functions in the map and/or mapplots packages.
- Restart R (or the computer) and then only use library(map) and library(mapplots) in the beginning and see if the code works.

We produced this plot



Same thing but in panels

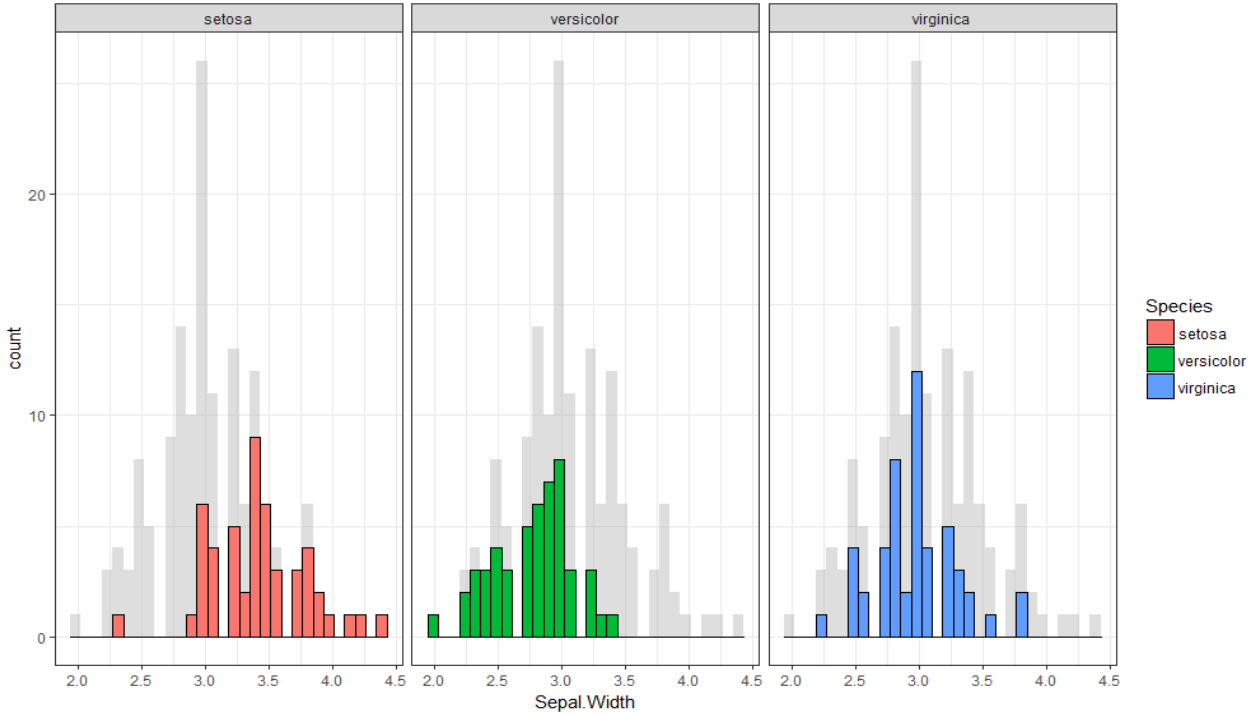


lunchR

Session 2: One more thing..

https://drsimonj.svbtle.com/plotting-background-data-for-groups-with-ggplot2

Same thing in panels with the background data in grey (i.e. the first plot)



lunchR

Sessions

1. Intro to why R, data formats and plotting styles in R
2. Digging deeper into data exploration in R using base R and ggplot; focus on different graph styles
3. **Tips and tricks in R – format your data for easy plotting**
4. Creating publish-quality figures

lunchR

Tips and tricks in R – format your data for easy plotting

- Get your data into R and share it with others
  - Email, Server or URL?
- Handle, re-organize and create data in R
  - Pipes and ggplot = goosebumps
- Cheatsheets

lunchR

How and why do we share data and scripts?

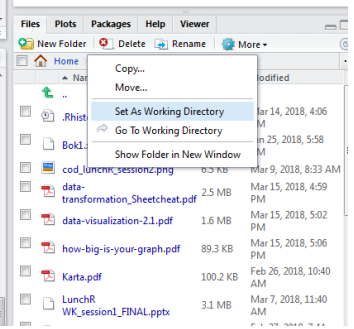
- Co-authors
- Back up
- Reviewers
- Students: teaching, R-labs, R-courses (e.g. LunchR)

lunchR

How do we share data and scripts? – Session 2

```
399  setwd("\\\\storage-og.slu.se/home$/phja0001/My Documents/Philip/R
     /LunchR")
400  getwd()
401
402  # Now R know where to look for the picture we downloaded so we can load
     it into R:
403
```

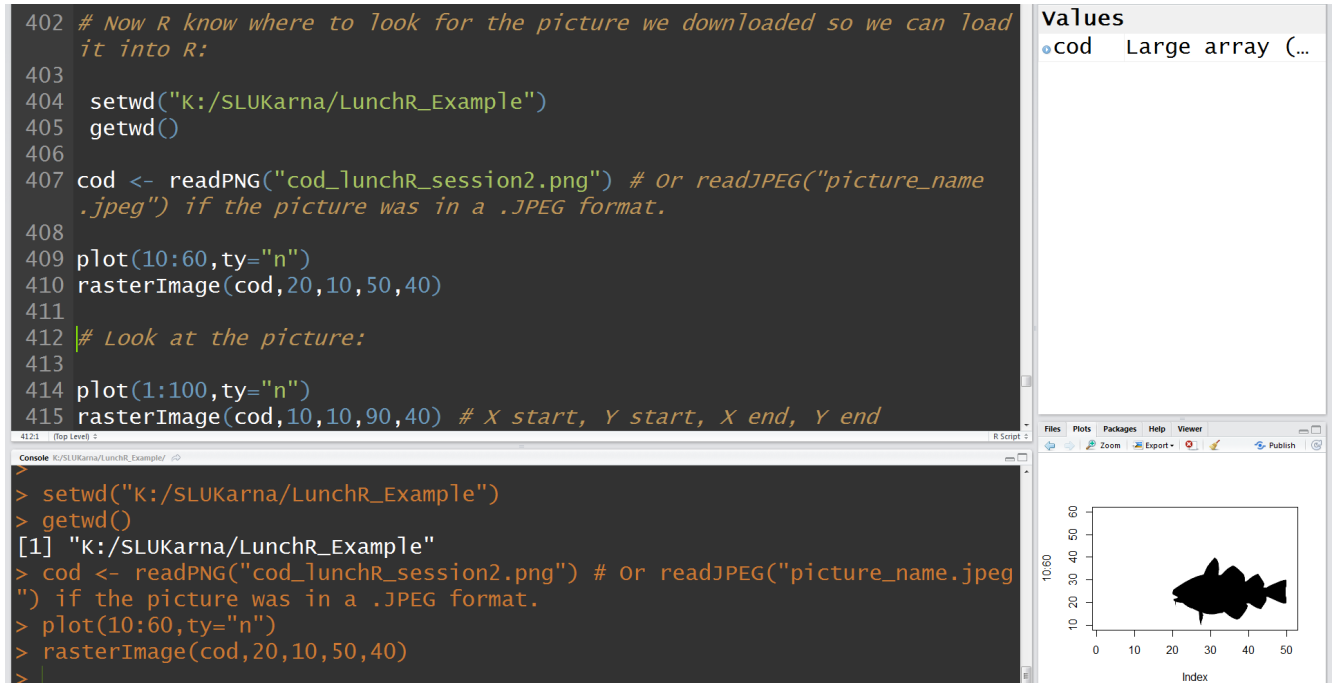
```
> setwd("\\\\storage-og.slu.se/home$/phja0001/My Documents/Philip/R/LunchR")
> getwd()
[1] "\\\\storage-og.slu.se/home$/phja0001/My Documents/Philip/R/LunchR"
```



This is not bad but it can lead to problems. You need to email scripts+data and the user need to manually change the code.

For anyone else to use this script, the person needs to specify where they have stored the object (data or picture, Session 2) we need to read in to R.

How do we share data and scripts?



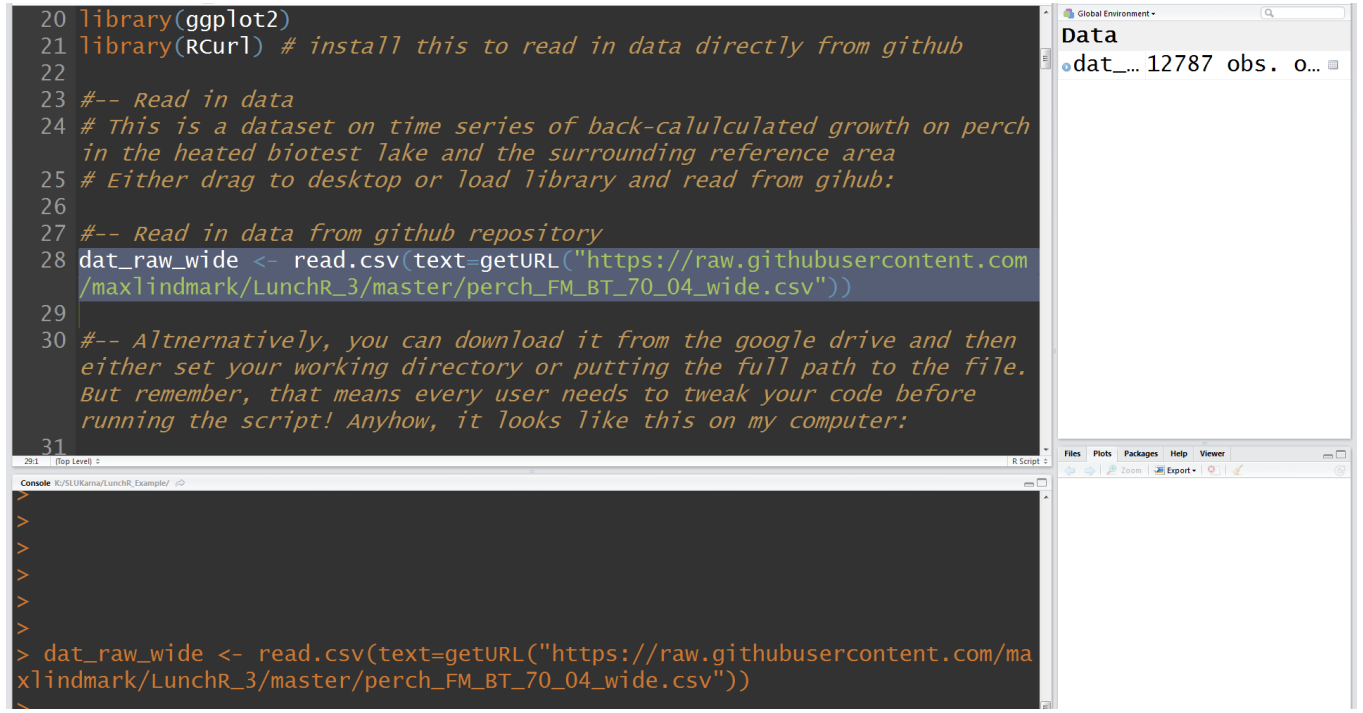
For those that do not have access to the server?

– To bad..

We can use a server. Now everyone that have access to the Öregrund common server K: can produce this plot by using the script.

lunchR

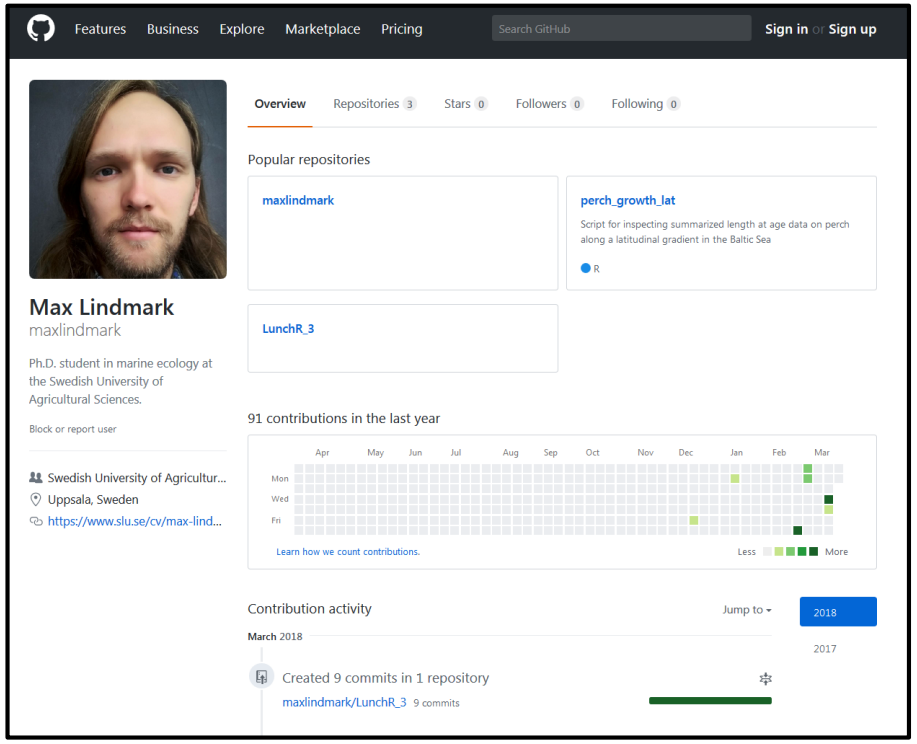
How do we share data and scripts?

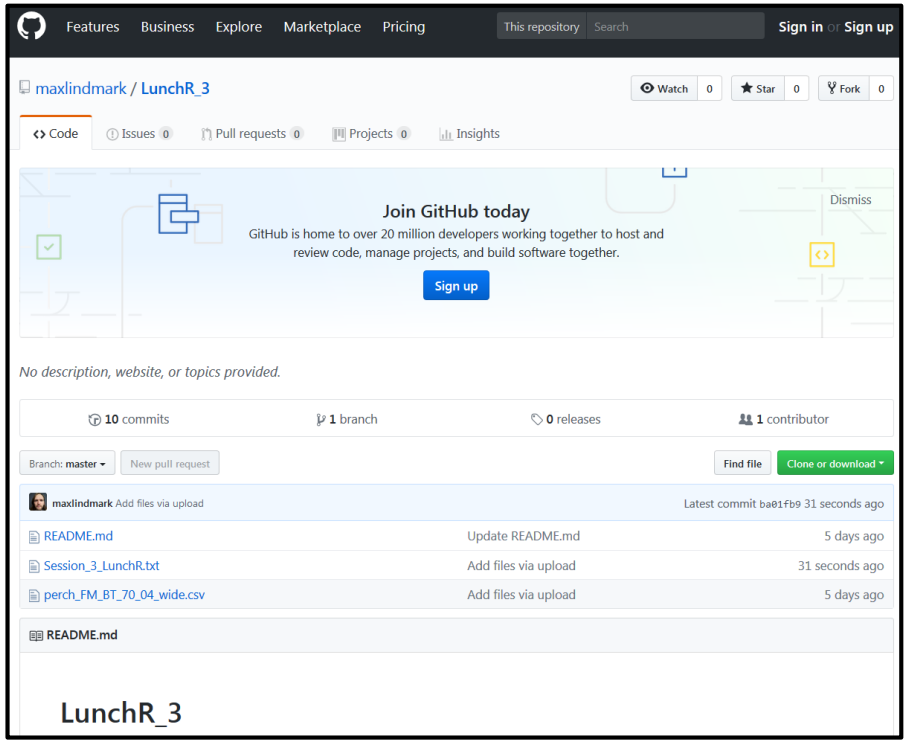


Available to everyone with an internet connection

URL! We can upload our script and data to an URL like github.org and load the data directly from there (we will do this later today).

How do we share data and scripts?





https://github.com/maxlindmark/

lunchR

How do we share data and scripts?

Email: Send script+data.
The user need to manually change the working directory = can be problematic. Automatic back-up.

Server: As long as everyone have access to it, go ahead! Automatic back-up. No need for manual editing.

URL: Everyone can access script and data. Automatic back-up. No need for manual editing.

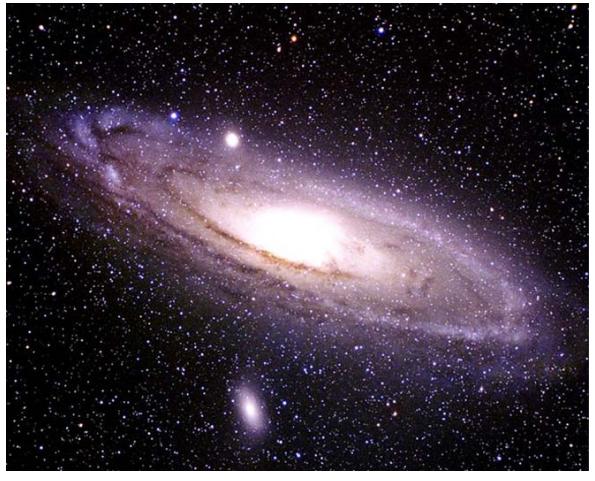
We will use URL today

lunchR

Handle, re-organize and create data in R

- Used to be tricky and time consuming
- You had to learn a lot of functions and code
- Very different from other programs (e.g. Excel)

This is not the case anymore thanks to:



%<%

Hadley Wickham

From Session 1

Tidy data with one line of code!

“tidy” data

```
> l
    x           y group
1   1 0.497868052     a
2   2 0.691927672     a
3   3 0.760313282     a
4   4 0.155401223     a
5   5 0.849457093     a
6   6 0.946817819     a
7   7 0.588419190     a
8   8 0.502250815     a
9   9 0.189779918     a
10 10 0.001836858     a
11  1 1.077578062     b
12  2 0.334111338     b
13  3 0.222741224     b
14  4 1.139136706     b
15  5 0.492948723     b
16  6 0.364326574     b
17  7 0.599102556     b
18  8 0.659575412     b
19  9 0.634030849     b
20 10 0.717009826     b
```

“wide data”

```
> w
    x           a         b
1   1 0.497868052 1.0775781
2   2 0.691927672 0.3341113
3   3 0.760313282 0.2227412
4   4 0.155401223 1.1391367
5   5 0.849457093 0.4929487
6   6 0.946817819 0.3643266
7   7 0.588419190 0.5991026
8   8 0.502250815 0.6595754
9   9 0.189779918 0.6340308
10 10 0.001836858 0.7170098
```

Several ways to switch between tidy and wide data!

`gather()` `spread()`
(from the “tidyr” package which is included in the tidyverse)

lunchR

Handle, re-organize and create data in R

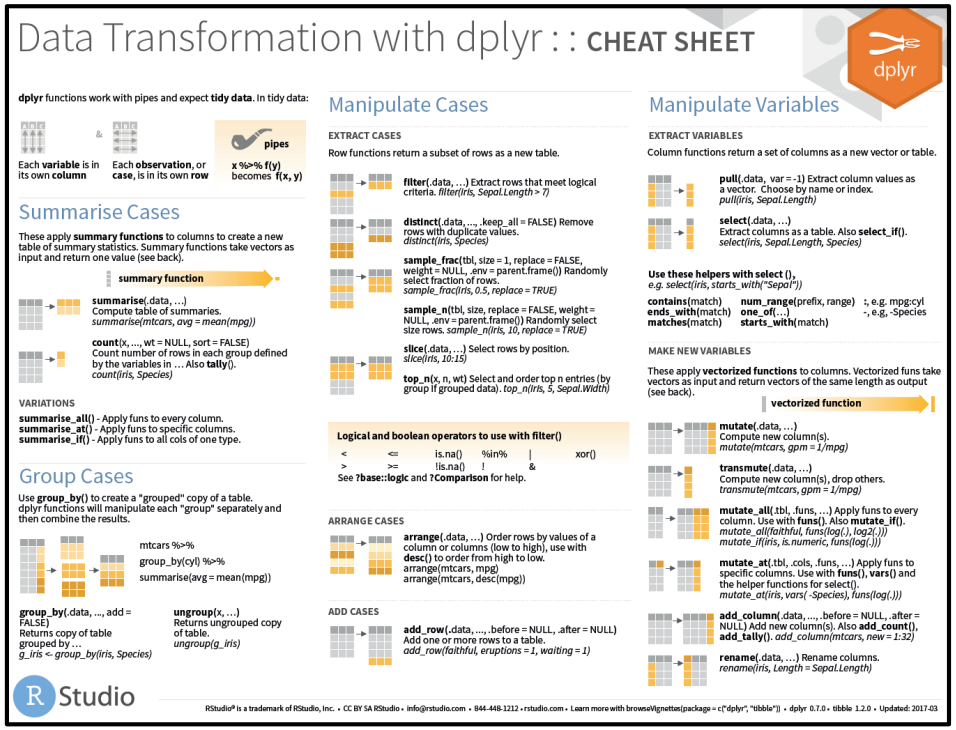
- Use pipes (%>%) to:
  - Convert data from a “wide” format to a “long” format (tidy data)
  - Summarize data (similar to the Pivot function in Excel)
  - Create new data
  - Filter out data
  - Plot and visualize specific parts of your data (both raw and data summaries)
  - Sort your data

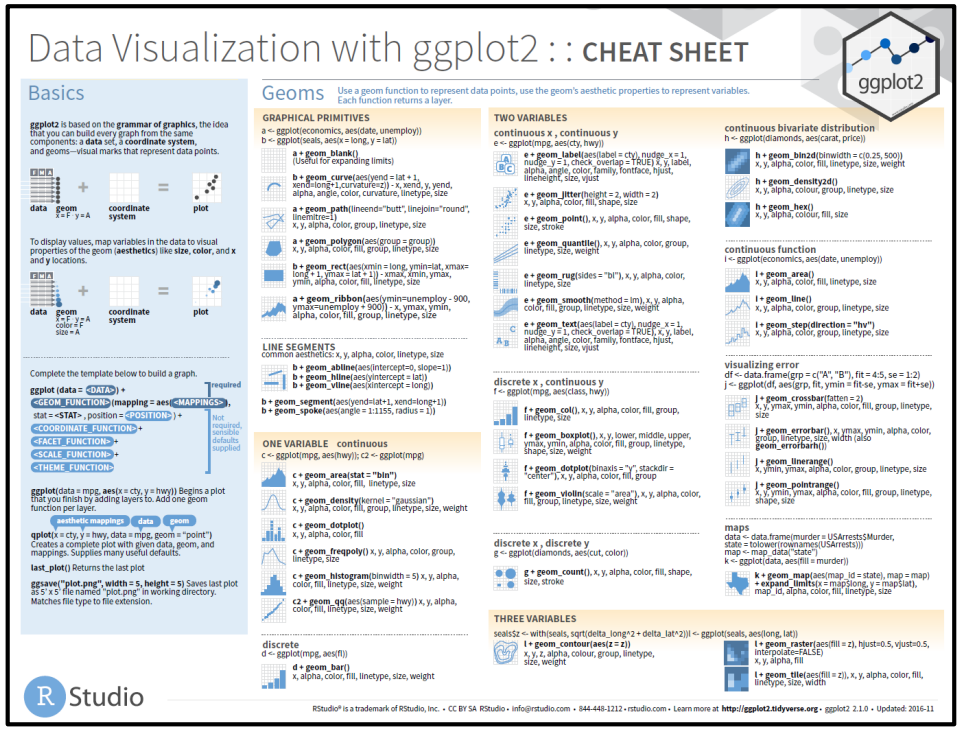
- Show how to organize your scripts in R using clickable headers (RStudio feature)

lunchR

Handle, re-organize and create data in R

- If you want to learn more (which we hope after this session), have a look at these Cheat Sheets: https://www.rstudio.com/resources/cheatsheets/





- **Data Import Cheat Sheet**
- **Base R Cheat Sheet**
- **Advanced R Cheat Sheet**
- **R Markdown Cheat Sheet**

and many many more..

lunchR

KUL data from the biotest basin loaded from an URL

Database for Coastal Fish - KUL

LAST CHANGED: 22 NOVEMBER 2017

The database KUL, with data from fishing with net and fyke net in coastal Waters, has been in use since year 2006 and provides quality assured catch data of coastal fish. The base also store information such as individual gender, length, weight and age. The database complements the internal database FiRRe, which stores data collected since the 1960s. Continuous work is done to transfer older data to the database KUL. The database is in Swedish.

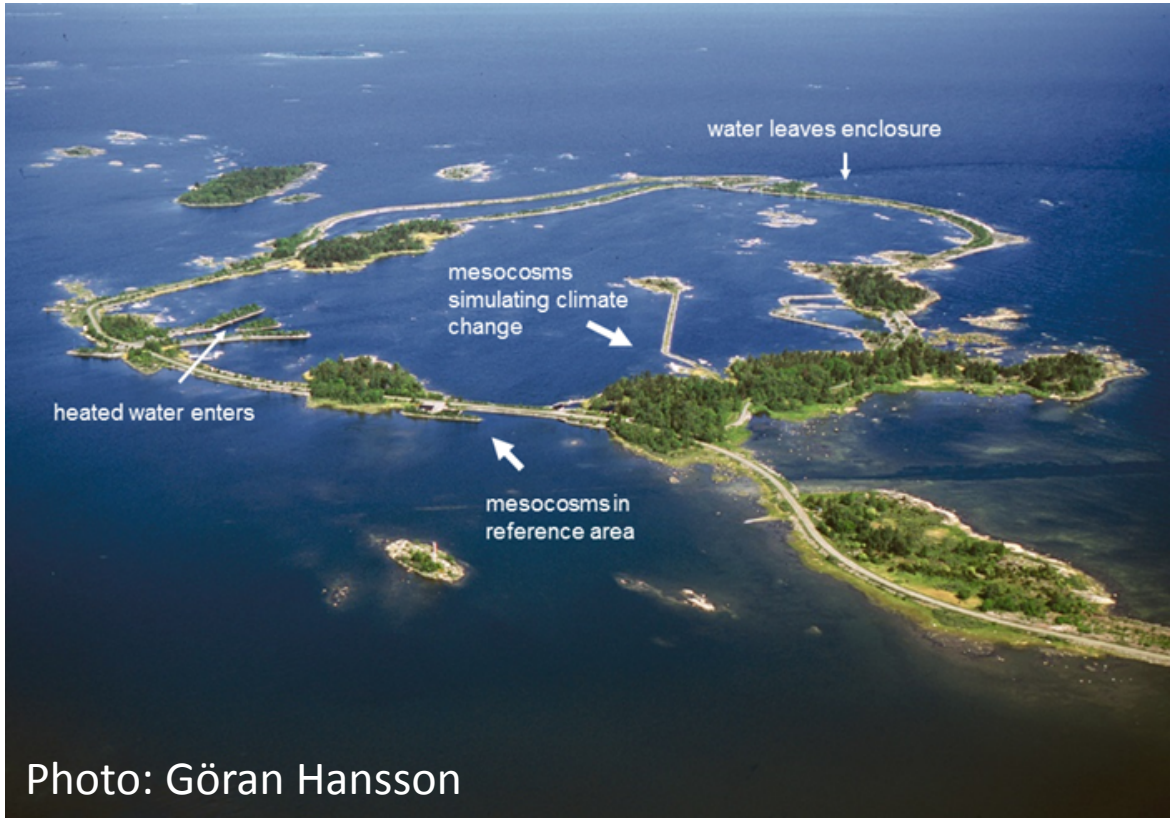


Photo: Göran Hansson

lunchR

Lets work with some KUL fish data!