

METHODS 22222 WOooooo

# CLASS 1:

## REGRESSION AND THE GLM: OVERVIEW, DATA AND MEASUREMENT



# AGENDA FOR TODAY

Warming up to R after break + solving GitHub issues + doing exercises

## Getting started

- Hi
- Lecture review

## Important business

- Git setup run-through
- Some meta talk on exercises and workflow

## Exercises!!

- New Spotify playlist



Taking a picture of  
your code



# MENTI TIME: CLASS 1

—



4217  
4745



# MENTI TIME: CLASS 2

---



3300  
2124



# LECTURE REVIEW

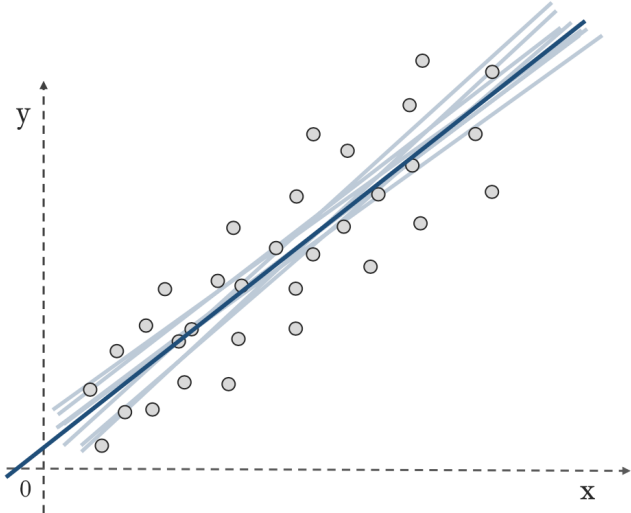
How class 4 (this class) connects to lecture 4



# LECTURE REVIEW

## SLIDES FROM THE LECTURE

### Statistical inference



### Linear Algebra

$$y_i = \alpha + \beta x_i + \epsilon_i$$

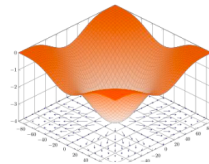
### Probability Theory

#### Bayesian inference

$$P(\alpha|y) = \frac{P(y|\alpha)P(\alpha)}{P(y)}$$



### Calculus



Linear regression:  $y_i = \alpha + \beta x_i + \epsilon_i$

Regression model parameters = coefficients that rep. the relationship between ind. vars (predictors) and dep. var (outcome)

### The generalized linear model (GLM)

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{E}$$

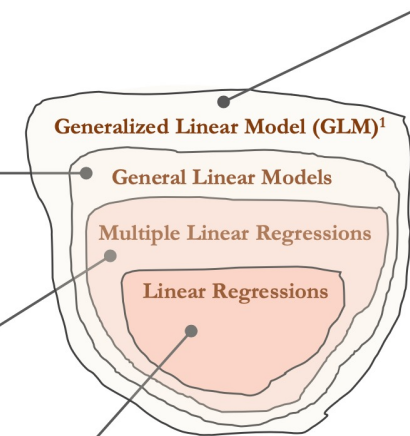
Multiple predictors and outcome variables. Include ANOVA, t-test, ANCOVA...

$$y = \mathbf{X}\boldsymbol{\beta} + \epsilon$$

One outcome, multiple predictors variables

$$y_i = \alpha + \beta x_i + \epsilon_i$$

One predictor, one outcome variable



The residuals can come from any distribution from the exponential family

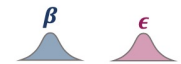
The link between the predictor variables and the outcomes can be any function

$$\mu_i = \mathbf{X}_i^T \cdot \boldsymbol{\beta} \rightarrow g(\mu_i) = \mathbf{X}_i^T \boldsymbol{\beta}$$

### Statistical inference

#### Three challenges of statistics:

1. Generalizing from sample to population
2. Generalizing from treatment to control group
3. Generalizing from observed measurements to underlying constructs of interest

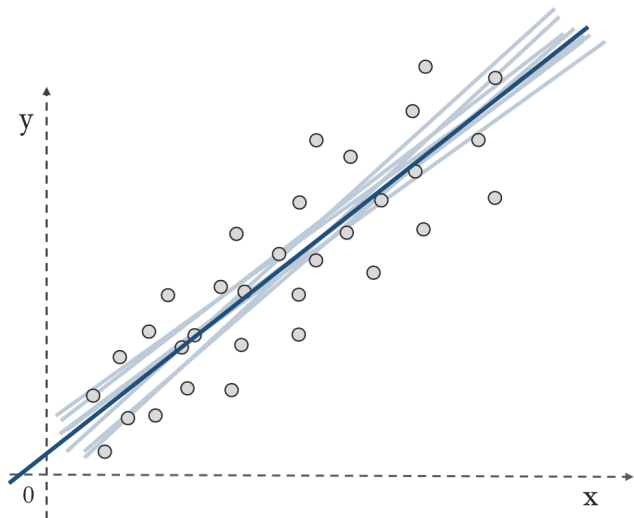


1 - Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized Linear Models. In Journal of the Royal Statistical Society: Series A (General) (Vol. 135, Issue 3, p. 370). IOSTOR. <https://doi.org/10.2307/2344614>

# LECTURE REVIEW

## SLIDES FROM THE LECTURE

### Statistical inference



Linear regression:  $y_i = \alpha + \beta x_i + \epsilon_i$

Regression model parameters = coefficients that rep. the relationship between ind. vars (predictors) and dep. var (outcome)

### Linear Algebra

$$y_i = \alpha + \beta x_i + \epsilon_i$$

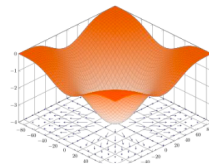
### Probability Theory

#### Bayesian inference

$$P(\alpha|y) = \frac{P(y|\alpha)P(\alpha)}{P(y)}$$



### Calculus



### The generalized linear model (GLM)

$$Y = XB + E$$

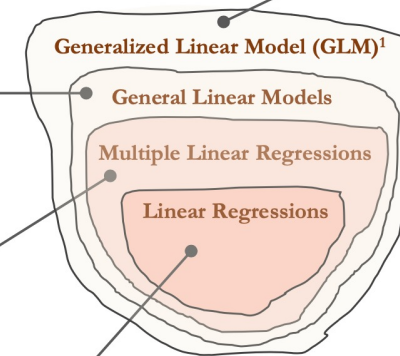
Multiple predictors and outcome variables. Include ANOVA, t-test, ANCOVA...

$$y = X\beta + \epsilon$$

One outcome, multiple predictors variables

$$y_i = \alpha + \beta x_i + \epsilon_i$$

One predictor, one outcome variable



The residuals can come from any distribution from the exponential family

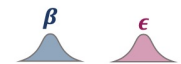
The link between the predictor variables and the outcomes can be any function

$$\mu_i = X_i^T \cdot \beta \rightarrow g(\mu_i) = X_i^T \beta$$

### Statistical inference

#### Three challenges of statistics:

1. Generalizing from sample to population
2. Generalizing from treatment to control group
3. Generalizing from observed measurements to underlying constructs of interest



1 - Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized Linear Models. In Journal of the Royal Statistical Society: Series A (General) (Vol. 135, Issue 3, p. 370). IOSTOR. <https://doi.org/10.2307/2344614>

My take on what this course is:

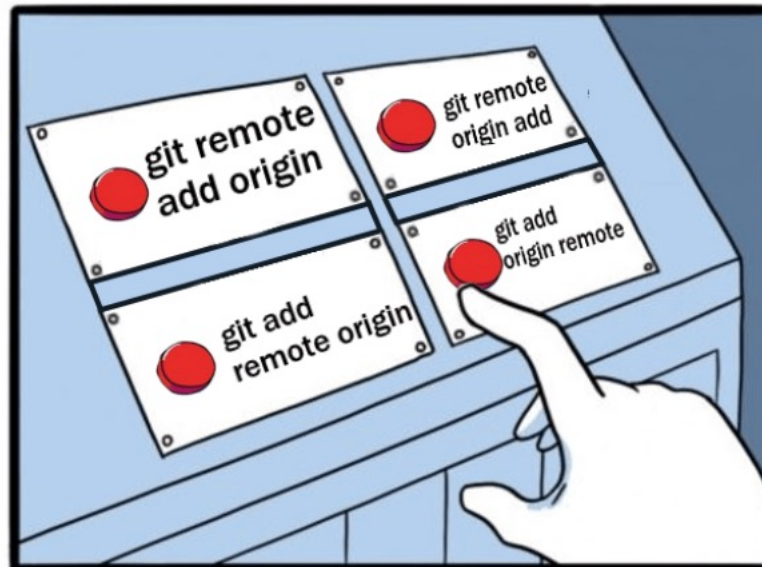




# GITHUB



## GET READY TO GIT



JAKE-CLARK.TUMBLR



SCHOOL OF CULTURE AND SOCIETY

AARHUS UNIVERSITY

METHODS 2 S24  
12 FEBRUARY 2024

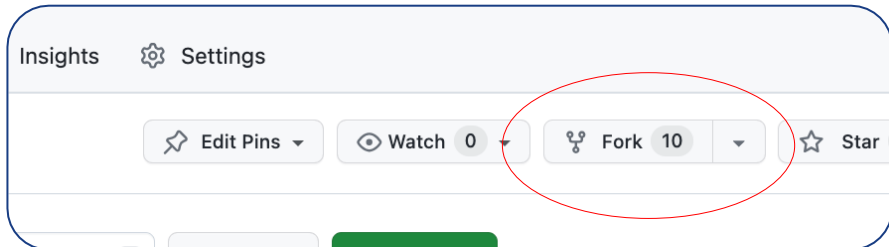
PERNILLE HØJLUND BRAMS  
TEACHING ASSISTANT



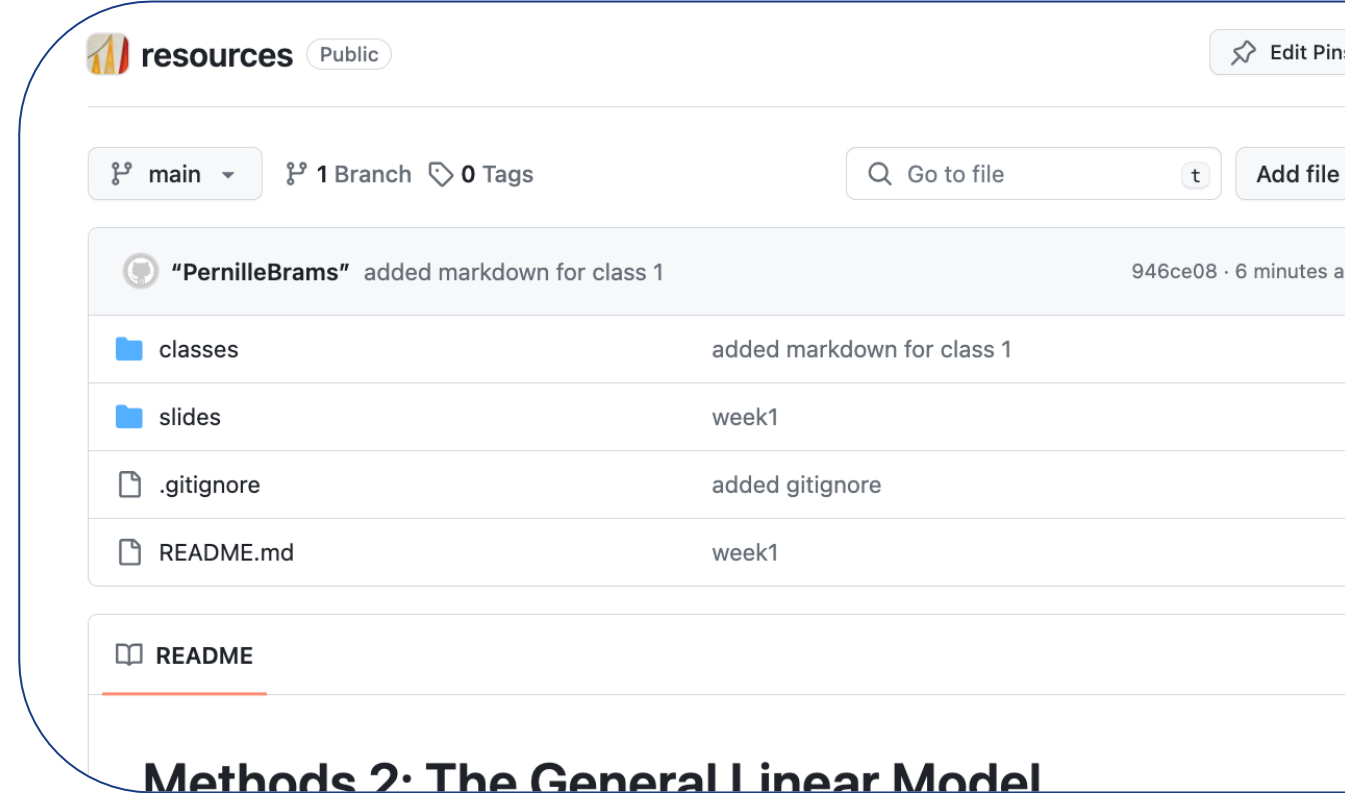
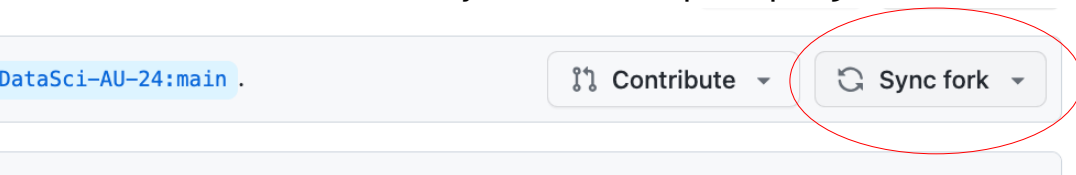


# GITHUB STUFF

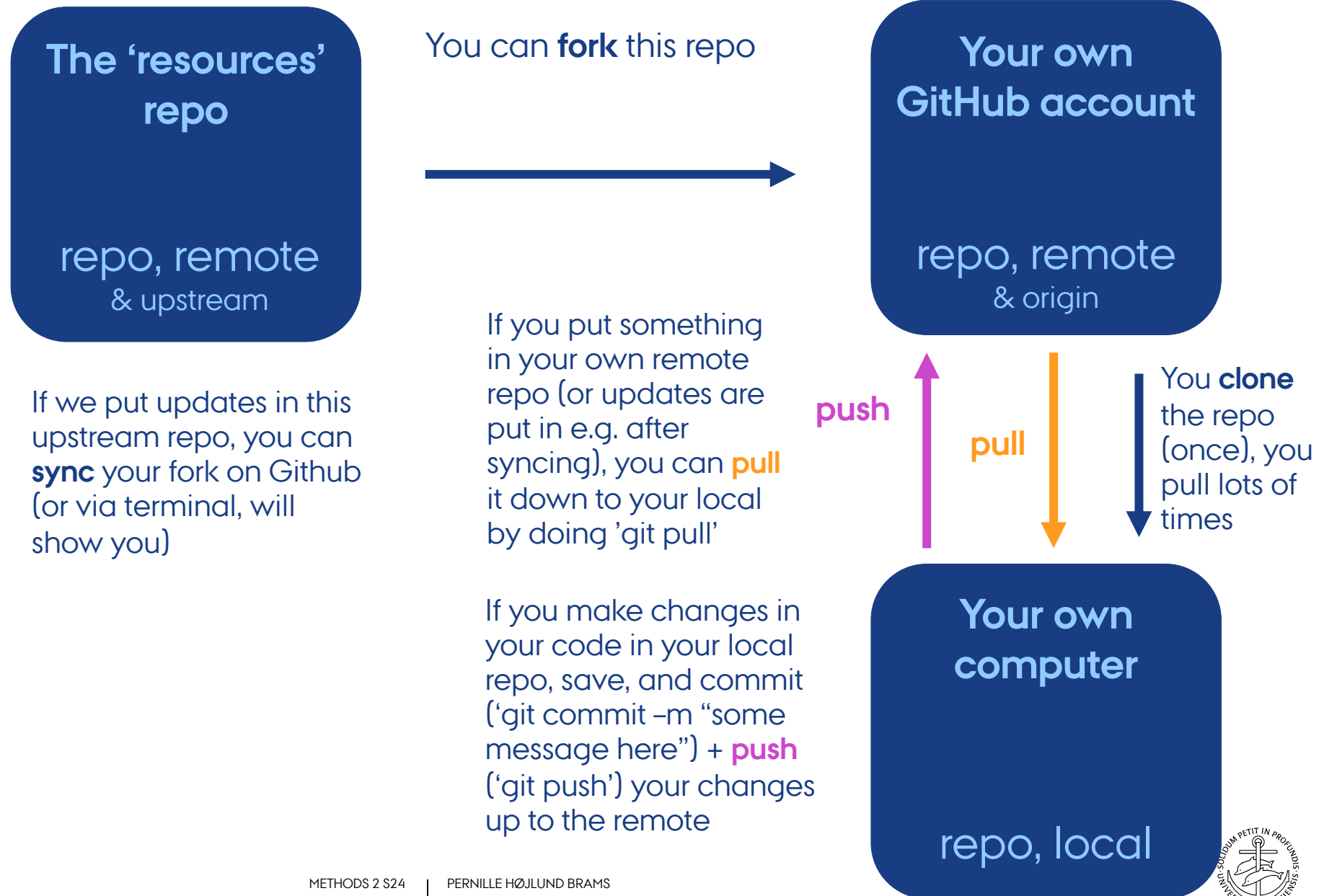
1. If you haven't already, go to the [course git-page](#) and find the 'resources' repo (scroll down) →
2. Fork it (make a copy) to your own account, and name it something



3. Go to your forked repo. When we update something in 'resources' (like putting an exercise there) you can 'sync' your new forked repo to get newest changes on this button at your own repo's page:



# GITHUB CLONING, PUSHING, & PULLING



# MORE GIT HELP FOR WHEN YOU NEED IT

## Cloning / Forking

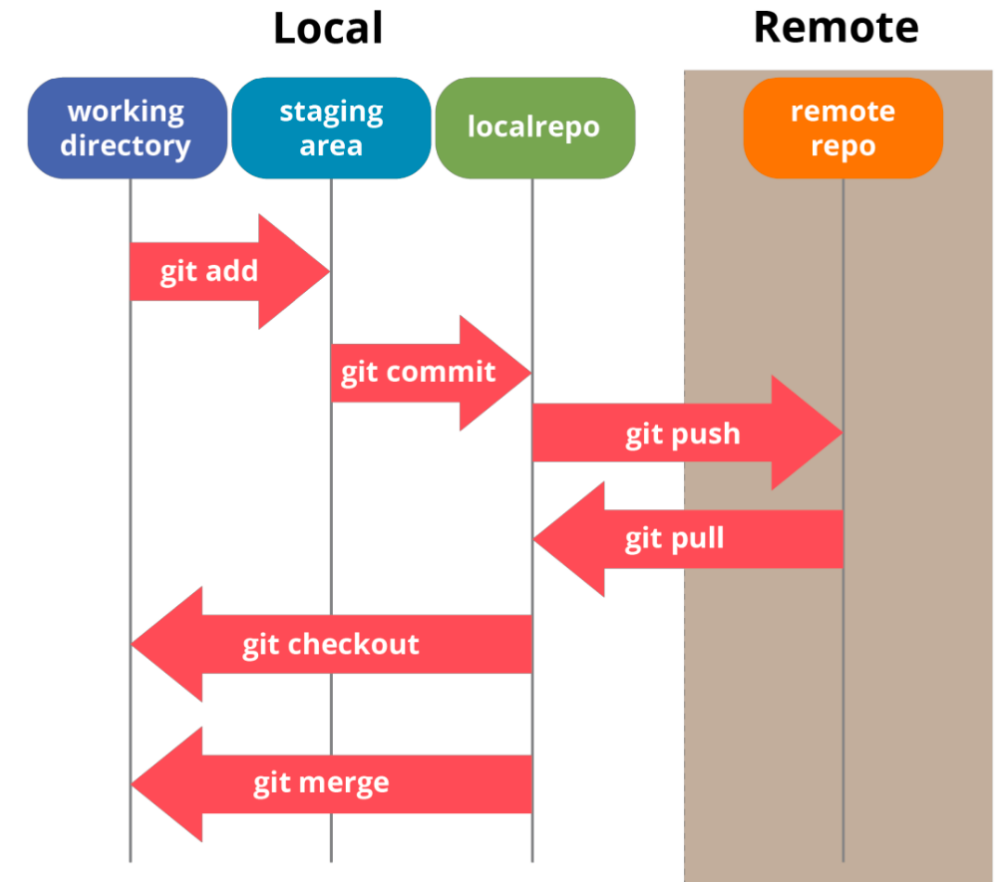
```
## go to the path where you want to clone YOUR fork  
git clone <url_to_your_repository>
```

## Pushing local changes to remote

```
cd <your_local_directory> # the folder you cloned your repository into  
git status ## check which files differ  
git add <name_of_your_solution_file.Rmd> # add files to what you will later commit  
git commit -m "<your message>" # commit all the files you have added above  
git push origin main ## push the committed changes to your origin (your remote location)
```

## Pull or sync fork + pull

```
git pull upstream main ## call this from your local copy to pull my latest changes ...
```



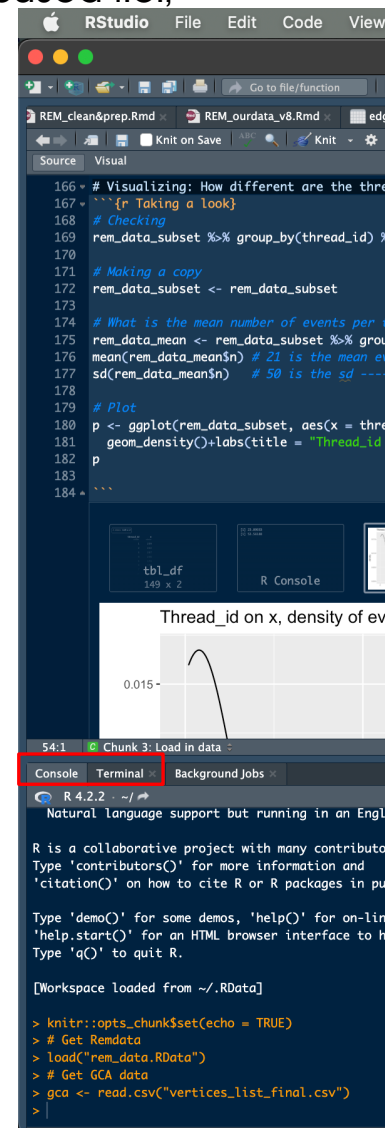
# OVERVIEW: BASIC COMMANDS FOR USING THE TERMINAL

**First of all:** In the simplest terms: the working directory is the folder on your computer where Rstudio is currently focused i.e., the specific folder you have opened and are currently in.

Terminal command	Use	Example
cd	Access directory (folder)	To access folder 'Project': cd Documents/Project
cd ..	Go up one directory	If you're in work/renv_cogsci_e23, and you want to look in work/, run cd .. in terminal
cd /.	Go to the root directory	cd /.
tab	Autocomplete a file or folder name	Type 'Do' + 'tab' to complete as 'Documents' (if you have that folder)
ls	List contents of the current directory	ls
pwd	Print the current working directory	pwd
dir	Same as ls, but mostly used in Windows	dir
mkdir	Create a new directory	mkdir MyNewFolder

In a chunk in RStudio: run `getwd()` (= "get working directory") to check where Rstudio says you are 😊 or you can run `pwd`, see table  
←

Terminal can be viewed down to the left, but in this picture, the console view is shown – you need to press the 'Terminal' tab right next to it to access the terminal 😊



# EXERCISES AND META-TALK

1. Statisticians love giving different names to things, and these books are no exception:

- “Sketch” = simulate = make fake data

2. The exercises in the course vary a bit in difficulty in general

- Some are R-based, some (esp. later on) will be more pen-paper-mathy

3. Always remember:



The solution to a difficult exercise is to break it up into smaller parts

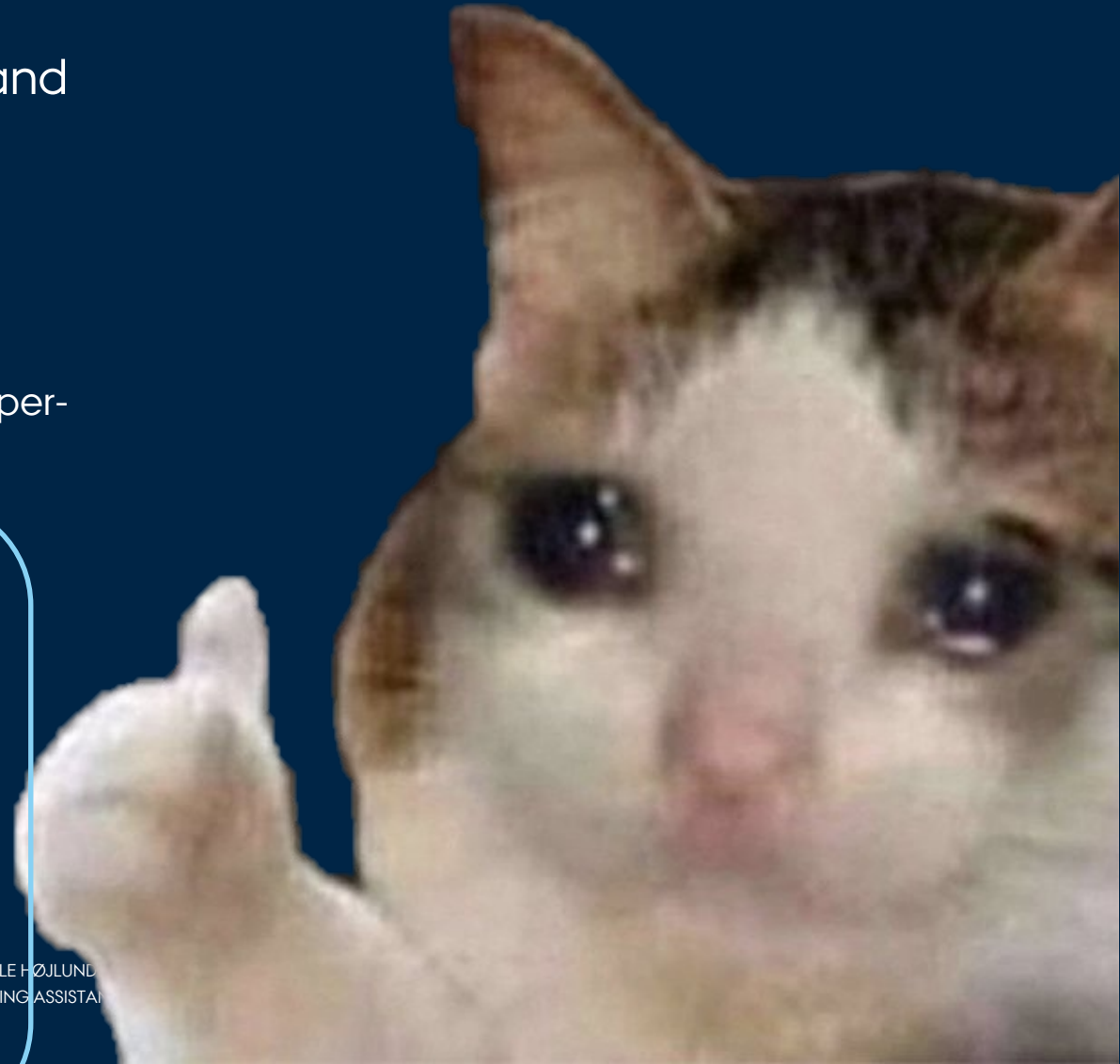


SCHOOL OF CULTURE AND SOCIETY

AARHUS UNIVERSITY

METHODS 2 S24  
12 FEBRUARY 2024

PERNILLE HØJLUND  
TEACHING ASSISTANT



# EXERCISES FOR TODAY

## IN GROUPS OR BY YOURSELF

3. ✨ Always remember:

The solution to a difficult exercise is to break it up into smaller parts

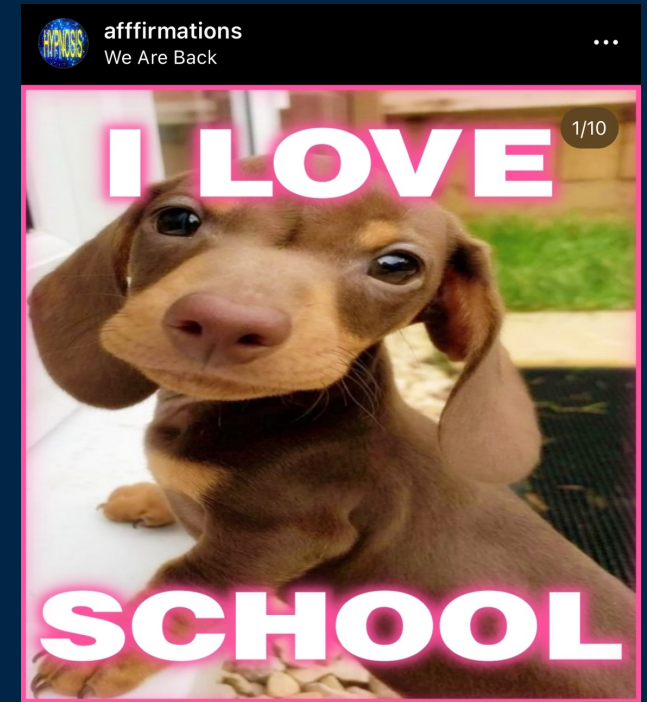
- Do some of the 0-exercises explained in the markdown, and then:
- **From Regression and Other Stories (ROS):**
  - 1.2\*
  - 1.5
  - 2.3
  - 2.7 b)

If you finish early:

- Try to optimise 1.2 e.g. by making a function to sample data and generate plots
- 2.7 a) from ROS
- all the ones from the first chapter in ROS

I will share solutions by tomorrow 🖋️

\*(if it is a bit hard at first, try and write down what you know from the start / from the question. Write down exactly what they ask you to do in this question – and think about what parts a regression needs, see lecture review slide)



Ask me or Nicolas for help if you're stuck - or go on to another exercise ☺️ *"If you are confused it is only because you are trying to understand"* – Riccardo Fusaroli, another CogSci prof

# BREAK



SCHOOL OF CULTURE AND SOCIETY

AARHUS UNIVERSITY

METHODS 2 S24  
12 FEBRUARY 2024

PERNILLE HØJLUND BRAMS  
TEACHING ASSISTANT







AARHUS  
UNIVERSITY



**You**

what are some cringe proverbs to say that to break a problem down into smaller ones makes it easier to solve



**ChatGPT**

Creating "cringe" versions of proverbs to emphasize breaking down a problem into smaller, more manageable parts might involve playful exaggeration or mixing metaphors in a way that's humorously awkward. Here are a few examples with a light-hearted, cringe-worthy twist:

1. **"Chop the mountain into pebbles, and it's easier to carry home."** - Overstating the divide-and-conquer strategy with an impractical solution.
2. **"Tame the dragon one scale at a time."** - Mixing fantasy elements into problem-solving advice for added melodrama.
3. **"Bite-sized cookies are easier to digest than the whole jar at once."** - Using a slightly awkward food analogy to describe task management.
4. **"Turn the ocean into droplets, and even a spoon is enough to tackle it."** - An exaggerated way to suggest making a task so small that it becomes trivially easy.
5. **"Unravel the sweater of problems one thread at a time, just watch out for the naked sheep."** - Adding a nonsensical consequence to the idea of deconstructing a problem.
6. **"You can't eat an elephant in one bite, but who's eating elephants anyway?"** - A twist on a classic proverb, introducing an absurd element to highlight the message's strangeness.

