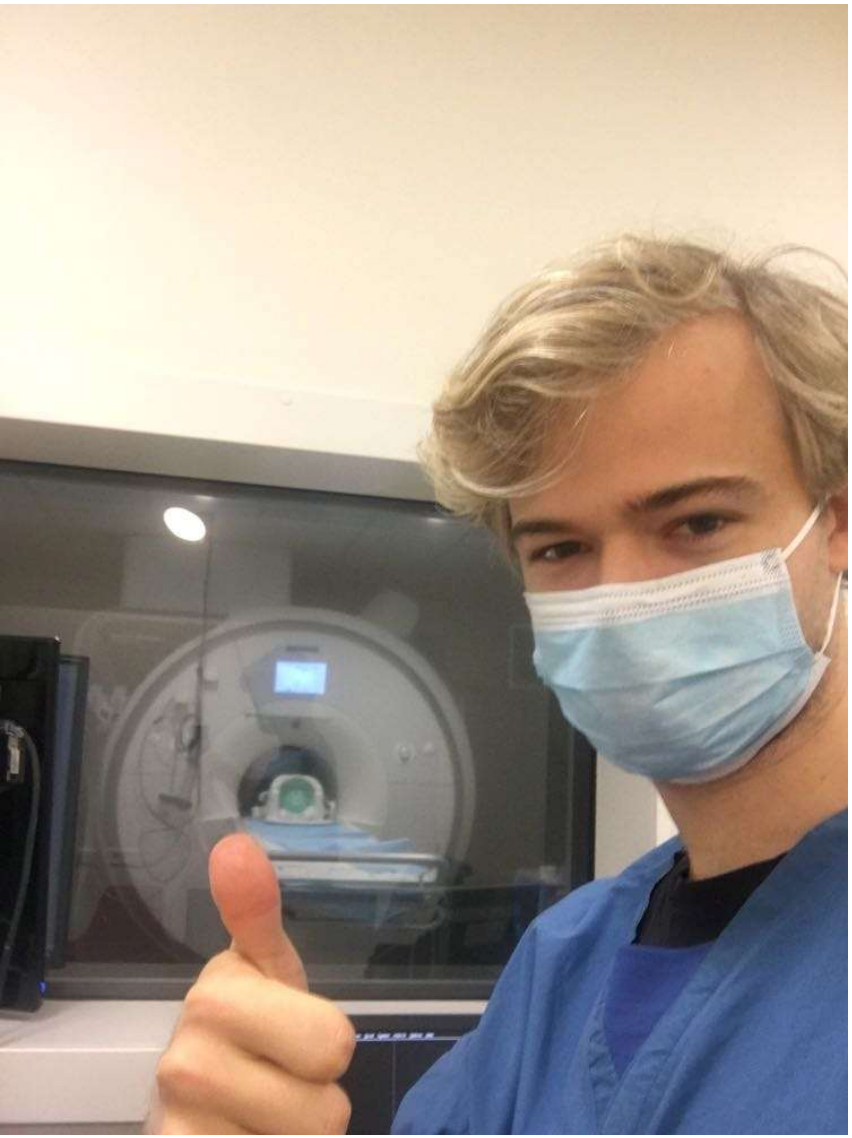
Class 1

A NEW CODE

The intro week was great but…



A bit about me:

- Did some nice AI and Data Science at KU's (Artificial Intelligence) degree.
  - Some NLP, P300 EEG Classification, tried to build a song recommender (failed horribly).
- Research Assistant (RA) on a couple of music fMRI projects last year.
- Now I am a RA for Daina on a fNRIS horror project
  - With Mina who you met in the intro week.
- I can help you with coding, neuroscience and psychology.
  - Don't ask me questions about linguistics (ask Fabio).
- Contact me on:
  - Mail: 202006317@post.au.dk
  - Facebook/LinkedIn: Sigurd Fyhn Sørensen

WHAT DO YOU EXPECT

”Rules” for the Class

1. No such thing as a stupid question!
2. AND… NO such thing as stupid answers!
3. Always alright to shout ”WAAAAIT!”
4. Help each other out!
5. Have fun ☺

Why
prog

- Much more
- Gives bette
- H4CK3R-M
- Allows us t
- Did someb

A couple of key terms:

1. Controlled Observations:
   - Lab experiment, standardized procedure.
   - keep confounding variables constant/adjust (controlled setting).
   - structured (code/measure behavior on a previous agreed scale).

2. Naturalistic Observations:
   - Unstructured, record all relevant behavior without system (not on a point scale)
   - Good as a preliminary study to get an idea of variables of interest.
   - High ecological validity but lacks representativeness.

3. Participant Observations:
   - Same as naturalistic + researcher participation.
   - A zoologist living with and studying monkeys in their natural habitat for an immersive understanding.

Experiment:

There exist many types of experiments in the field of cognitive science, but they all have one goal:

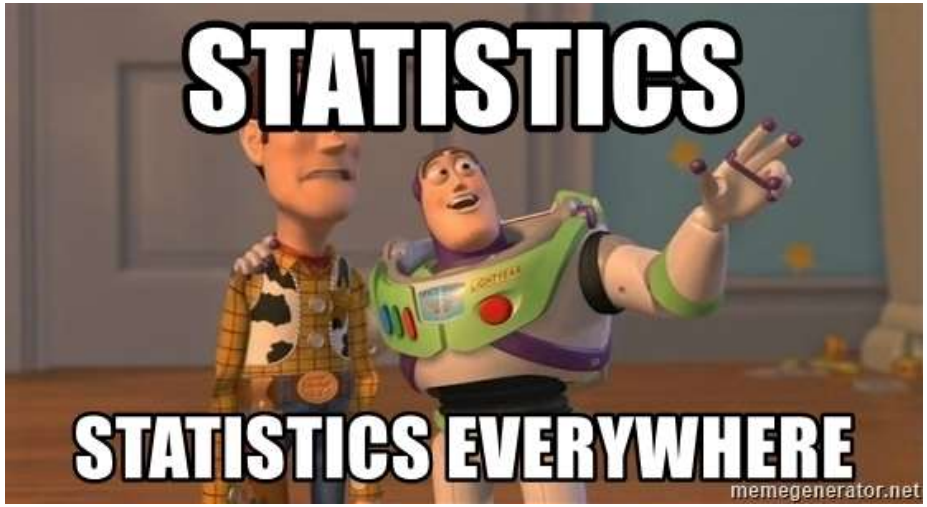
What is the point?!?

- Investigating the influence (**X**) holds on (**Y**).
  - Alcohols (X) effect on attention span (Y).
  - Would listening to music while reading (X) improve your recollection capabilities of said text (Y).
  - Will drug A compared to drug B (X) significantly increase your chances of getting better (Y)

Statistics are all around us:

Things based on statistics:
- Prediction of political elections
- The economy
- Medicine
- Gender equality studies
- And so forth…

At face value statistics can easily look like deterministic truth.
- However,…….



The fake effect:

Sometimes statistics are just too ludicrous!

Video of Simpson's paradox:
https://www.ted.com/talks/mark_liddell_how_statistics_can_be_misleading#t-151235

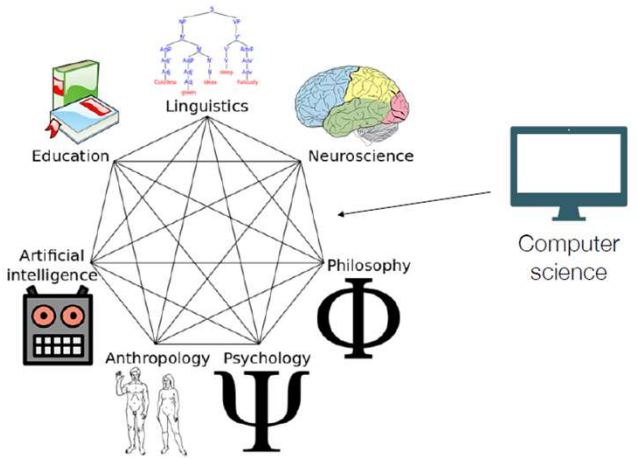
- Simpson's paradox: Block randomization in a controlled setting would avoid this.
- Another instance: A study in the UK showed that eating ice cream would significantly increase your chance of developing cancer.

  - A causality chain: Sunny -> Ice cream -> Cancer.
    - Weather is highly correlated with ice cream eating.
    - Seems as if ice cream eating can explain cancer.

Therefore, we need theory.

Without theory we would have accepted that eating ice cream would result in cancer.

Methods 1:
A course about methods in the cognitive sciences



Theory

Methods

Multidisciplinary

A little game: (of introspection) <3

You're all aspiring scientist and started planning a study:

1. Specify a **research** inquiry. (This is a question/pondering)
2. Write a specific, testable **hypothesis.** (this is a statement)
3. Consider which **variables** to collect and how they are related.
4. Plan how you will measure your **dependent variable. (1.person, 2. person , 3.person)**
5. Design experimental setup to manipulate your **independent variable** (explanatory variable) for an effect on **dependent variable** (response variable).
6. Assign subjects to groups, either **between-subjects** or **within-subjects.**
7. Consider to which degree you want to **blind** your study.

This is my suggestion:

1.: I wonder if I am going to be a good teacher for the methods classes.

2. H0 = I am not going to be better than the avg teacher. H1 = I am going to be better than the avg teacher.

3. **Binary Variable**: X1 = My Students , X2 = Not my student. **Method level:** Y = Score in a method test.

4. 2. Person methods as I am conducting multiple choice test.

5. **Test score** (Dependent Variable), **Teacher** (Independent Variable). Students will be assigned me or another teacher in a block-randomization based on prior skills in methods. (to ensure equal distribution and avoiding Simpson's paradox.)

6. Between subject design as participants will only participate in one condition. (X1 or X2)

7. Knowing that your teacher is going to be rated based on your performance might create self-awareness and skew the results. Therefore, we single-blind the study.

Let's get ready to rumble! Almost…

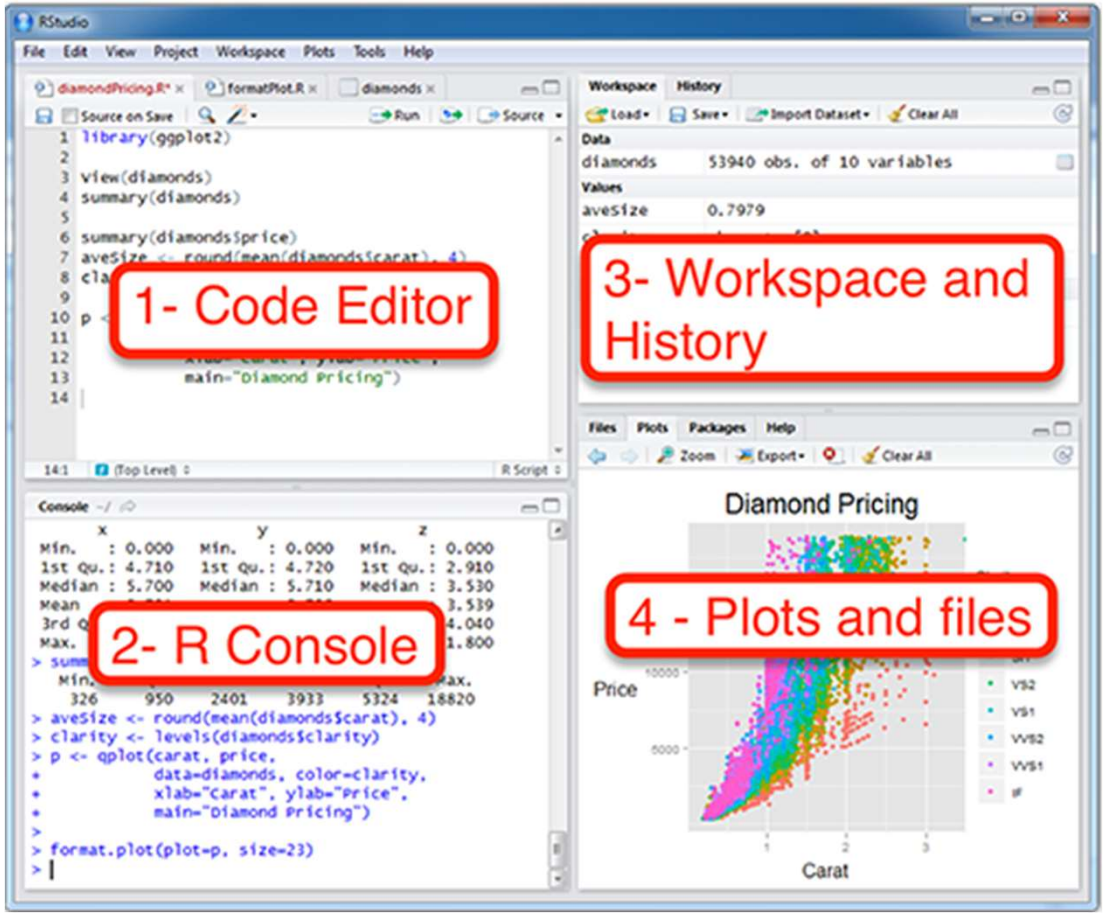
- Google is your new best friend
- No shame in reusing code (but try to understand it)
- Stackoverflow.com is where it's at

When you can't find an example of R code to steal from stackoverflow

R and Rstudio: What's the diff?

- R: The coding language
  - Free and Open Source (wuhu)
  - Powerful and flexible
  - Really nice syntax ☺
- Rstudio: The coding environment
  - The "Word" for programming
  - Makes the development experience a ton nicer

Let's  ROCK!


RStudio
File Edit View Project Workspace Plots Tools Help
diamondPricing.R* formatPlot.R diamonds
Source on Save Run Source
1 library(ggplot2)
2
3 view(diamonds)
4 summary(diamonds)
5
6 summary(diamonds$price)
7 aveSize <- round(mean(diamonds$carat), 4)
8 cla
9
10 p <
11
12
13 main="Diamond Pricing")
14
1- Code Editor
14:1 (Top Level) R Script
Console ~/
x y z
Min. : 0.000 Min. : 0.000 Min. : 0.000
1st Qu.: 4.710 1st Qu.: 4.720 1st Qu.: 2.910
Median : 5.700 Median : 5.710 Median : 3.530
Mean 3.539
3rd Q 4.040
Max. 1.800
> summ
Min. ax.
326 950 2401 3933 5324 18820
> aveSize <- round(mean(diamonds$carat), 4)
> clarity <- levels(diamonds$clarity)
> p <- qplot(carat, price,
+ data=diamonds, color=clarity,
+ xlab="Carat", ylab="Price",
+ main="Diamond Pricing")
>
> format.plot(plot=p, size=23)
>
2- R Console
Workspace History
Load Save Import Dataset Clear All
Data
diamonds 53940 obs. of 10 variables
Values
aveSize 0.7979
3- Workspace and History
Files Plots Packages Help
Zoom Export Clear All
Diamond Pricing
Price
10000
5000
1
2
3
Carat
VS2
VS1
VVS2
VVS1
IF
4 - Plots and files

Variables

- ”Boxes” to use for calculations
- Assigned with the ”<-” operator
  - E.g. `favorite_instructor <- "Jonathan"`
- Super-duper practical for writing gr8 code

Power tip: Use shortcuts! For fast ‘<–‘ try:

alt + ‘-‘ (windows)

⌥ + ‘-‘ (mac)

Data Types: The Atoms of R

- Numerical: Integers and floating points
  - Doing them maths!
- Character: ”symbols” and text
  - Remember the quotations
- Factors: Making text computable
  - For doing analysis with categories
- Boolean: TRUE or FALSE
  - Logic operations

```
me <- "Jonathan"

my_age <- 21

is_awesome <- TRUE
```

Vectors and Data Frames

- Vector
  - Ordered list of values
  - NB: Same data type!
  - Created with the `c()`-function
  - Accessed with indeces
    - `> my_friends[1]`
      `[1] "Fabio"`
- Data frame
  - Collection of vectors
  - Access vectors with $

    `df$age`

```
my_friends <- c("Fabio", "Sebber")
```

```
df <- data.frame(name=me, age=my_age,
                 awesomeness=is_awesome)
```

Exercise 1

1. Create a vector of the names of your study group, and assign it to a variable
2. Similarly make another vector with your guess on how many siblings they have
3. Add 2 to both vectors – what happens?
4. Check the class of both vectors and report the output
5. How many siblings in total?
6. What is the product of person 1 and person 2's number of siblings

Useful commands:

```
c()
sum()
class()
```

Extra:

- Why does this throw an error: `name <- Peter`
- Append a word to the number-vector
- What happens when you multiply by 2 now?
- Remove the word and try again!

Exercise 2

1. Create a dataframe with the previous vectors
2. Add gender to the dataframe
3. Add a new person to the dataframe
4. What is the mean number of siblings?
5. Ask people how many siblings they have and put the actual numbers as a separate column in your data frame
6. Make a column with numbers showing how much you were ‘off’
7. Comment your code

New functions you will need:

data.frame()

factor()

mean()

rbind()

Exercise three: extreme-edition!

1. How many people have 3 siblings?

2. Who has more than 3 siblings?

3. What does function length() do? **hint: use '?'*

4. What is the mean of c(2, 7, 'Cat')? What is the problem? ** hint:*

use class()

5. What is the problem in the following code?

```
105  names <- c("Peter", "Natalie", "Maya")
106  n_pets <- c(1,3,8)
107  pet_frame <- data.frame(names=names n_pets=n_pets)
```

6. Create a subset of the data for each person where you

guessed right

7. Try out functions round(); length(); unique(); mean()

New useful functions

```
subset()
round()
length()
unique()
mean()
```

What have we learned?

- Programming is awesome and fun! (right?)
- Programming is frustrating…
- The building blocks of R (data types, vectors, data frames)
- How to get help