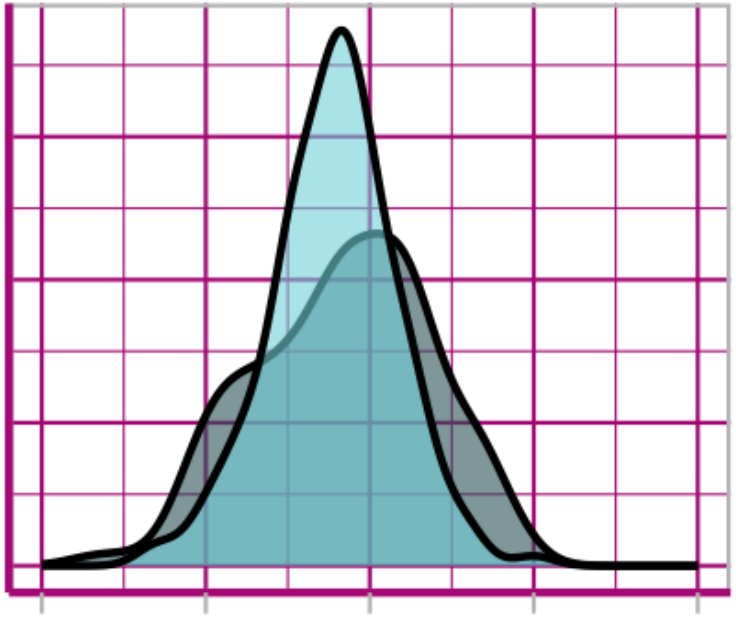
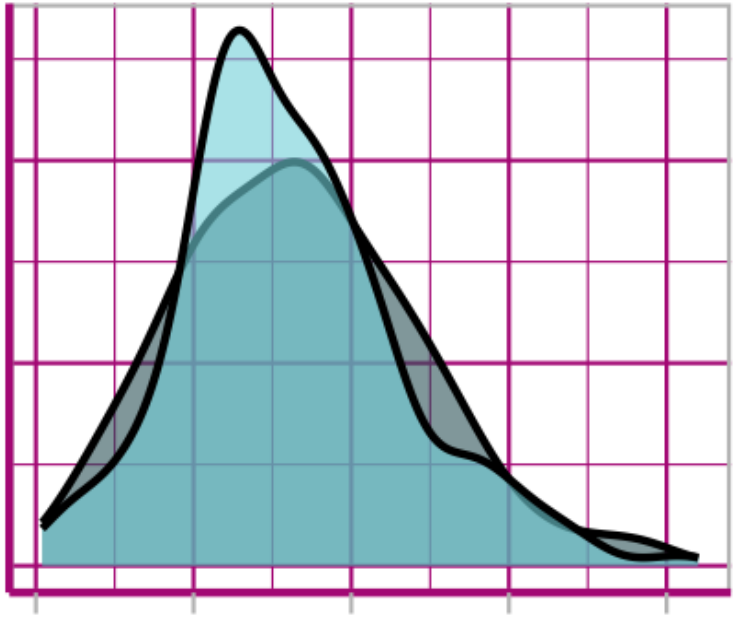
Data visualisation with ggplot2

Dr. Ryssa Moffat

R-Ladies Zurich

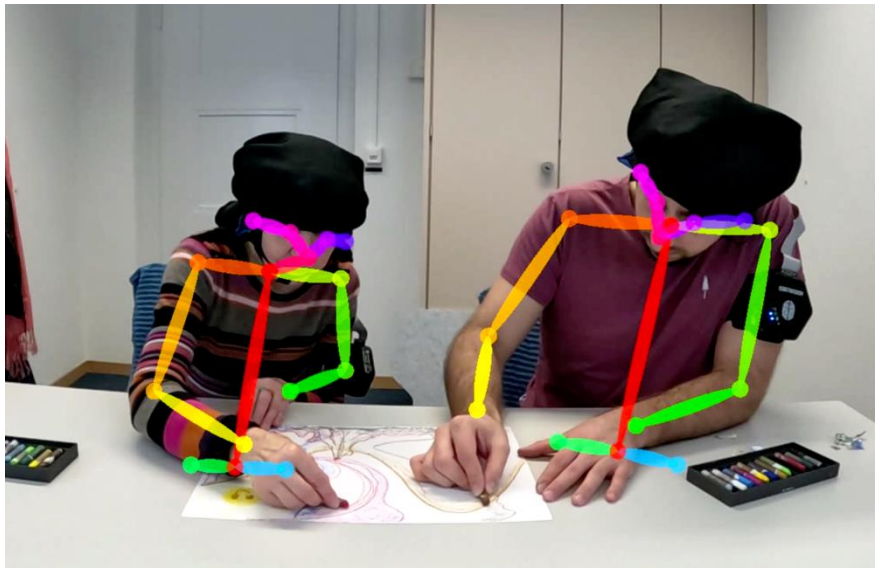
14.11.2024





About me

- Postdoctoral Researcher at ETH Zurich
- PhD in Cognitive Science
- R devotee since 2016



Social Brain Sciences

ETH zürich

R Ladies

www.ryssamoffat.github.io

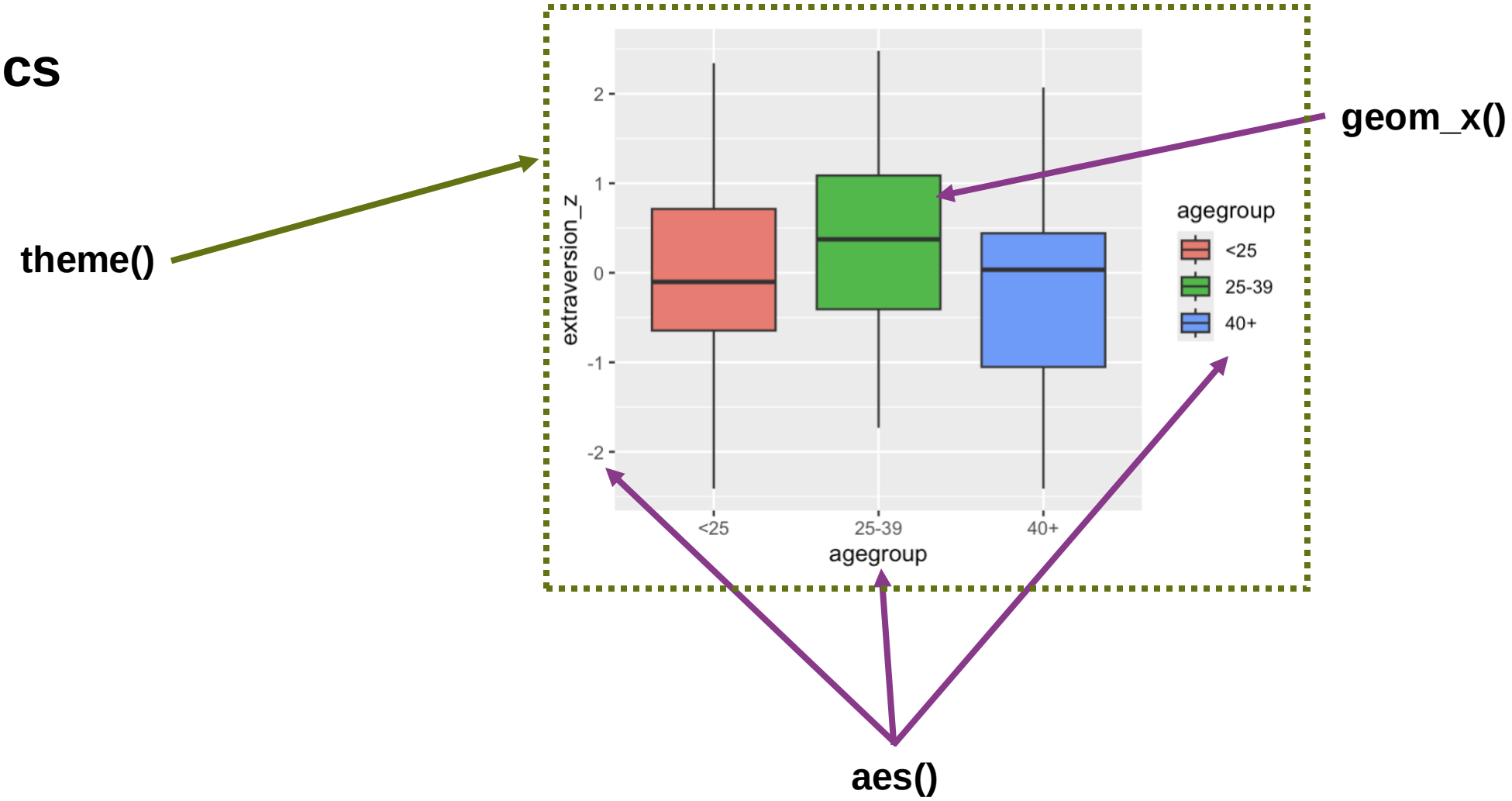
OSF

Plan

1. **Basics**
   - Syntax, Colors, Theme, Combining, Saving
2. **Important steps**
   - Pinning down relevant comparisons
   - Reducing redundancy
3. **Cool things**
   - Models into plots
   - 3D plots

BASICS

Syntax basics

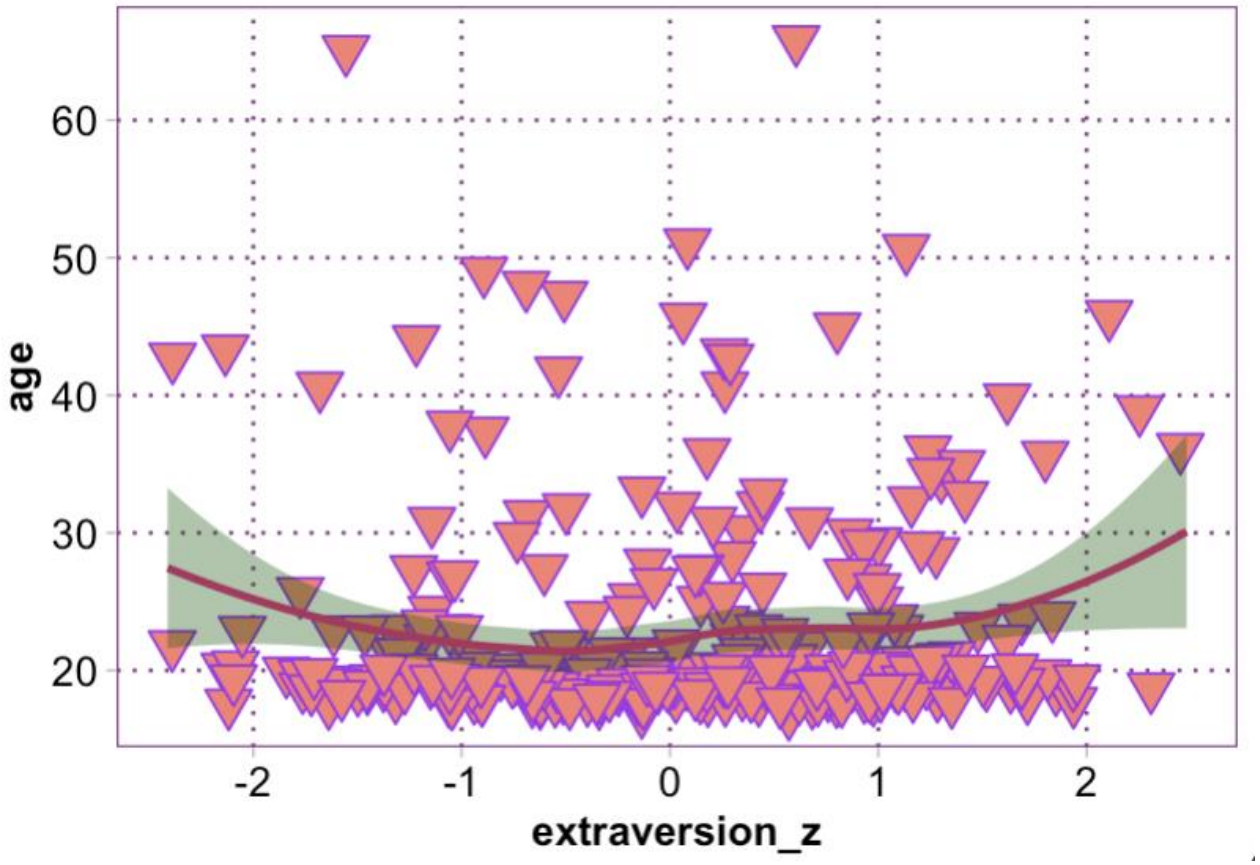



```
# boxplots (group with continuous, base ggplot colors)
personality_traits1 %>%
  ggplot(aes(x = agegroup, y = extraversion_z, fill = agegroup))+
  geom_boxplot()
```

Syntax basics

Familiarization with syntax:

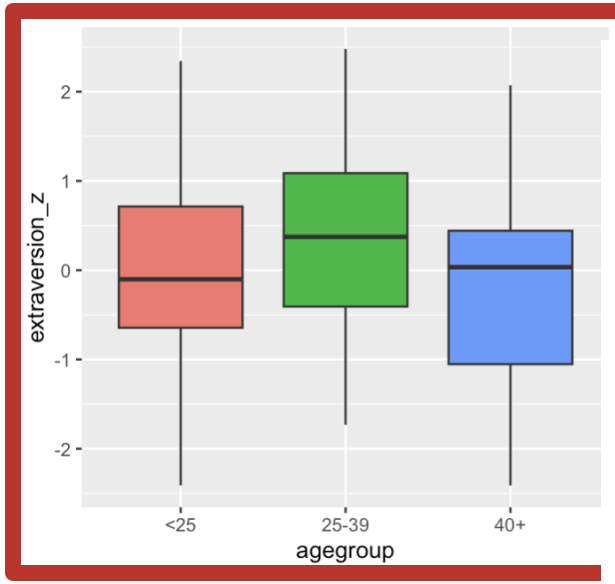
1. aes()
2. geom_x()
3. theme()

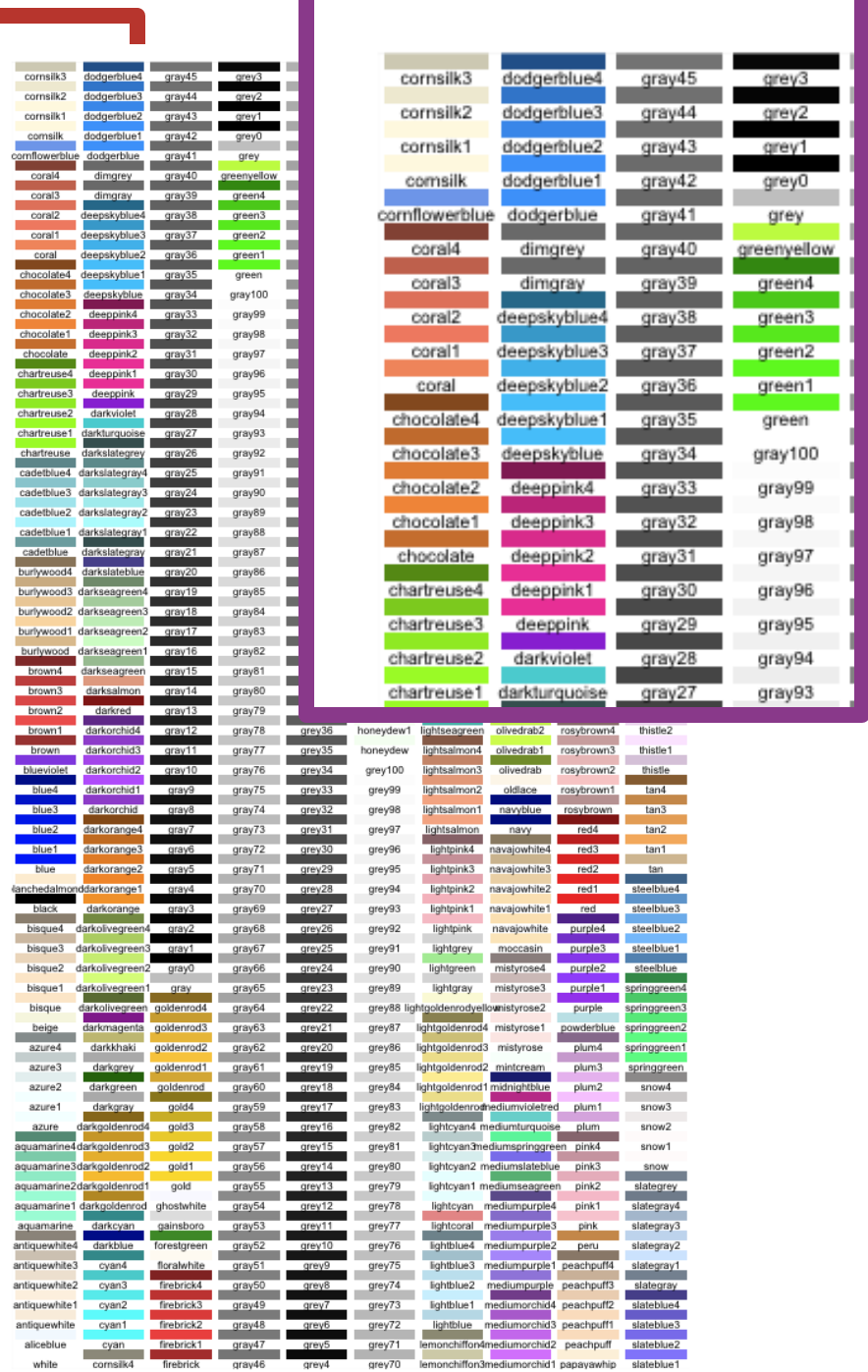



```
# scatter plot with trend line
# try changing color, fill, method (loess, lm)
personality_traits %>%
  ggplot(aes(x = extraversion_z, y = age)) +
  geom_jitter(color = "purple", fill = "salmon", shape = 25, size = 4)+
  geom_smooth(method = "loess", color = "maroon", fill = "darkgreen")
```

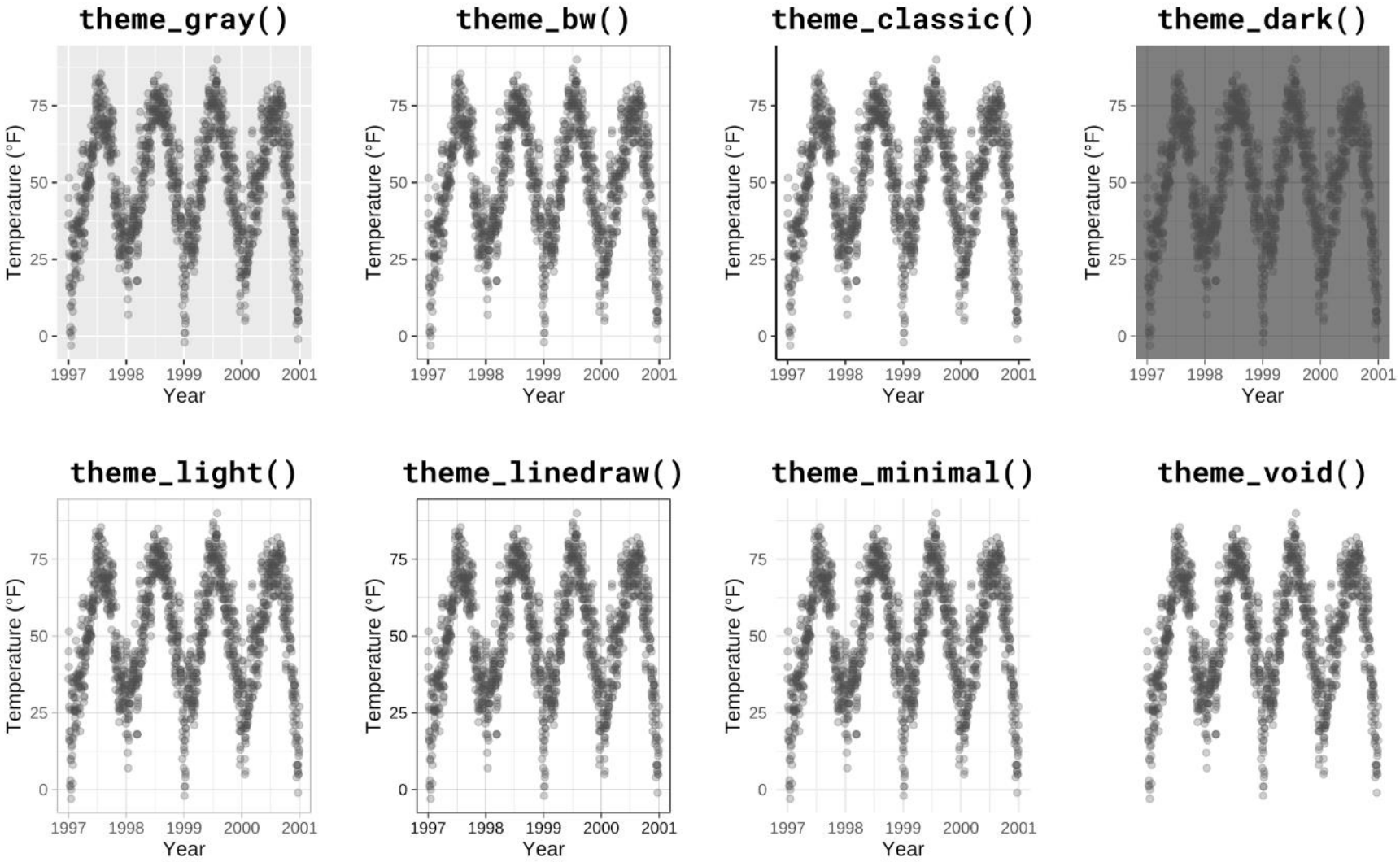
Colors

1. Color inspiration
2. Match existing colors
3. Coblis — Color Blindness Simulator

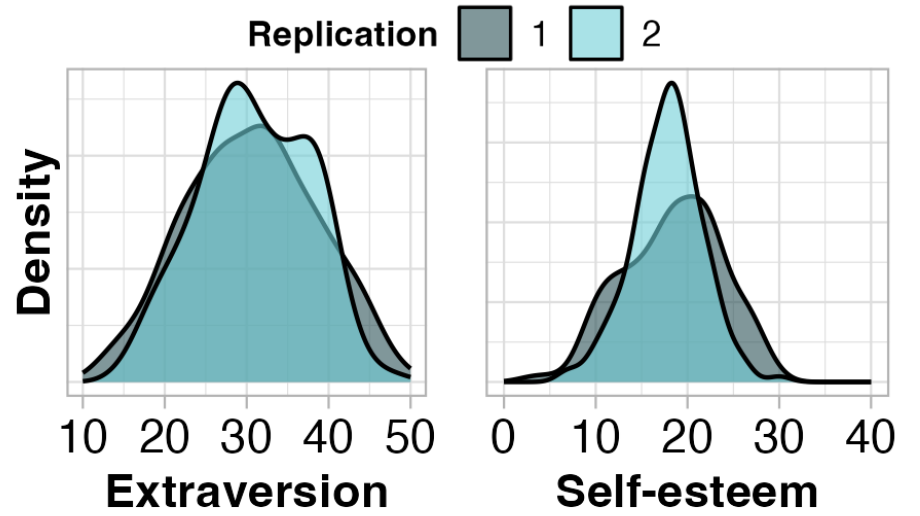


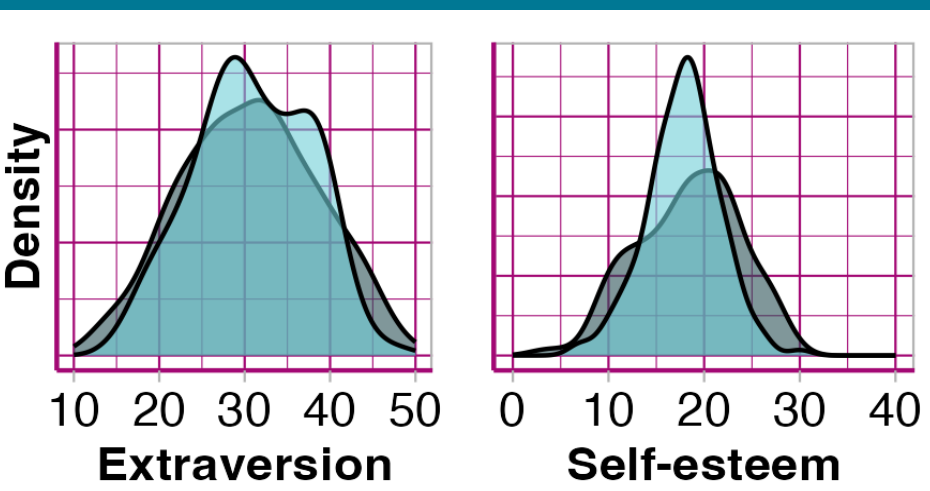


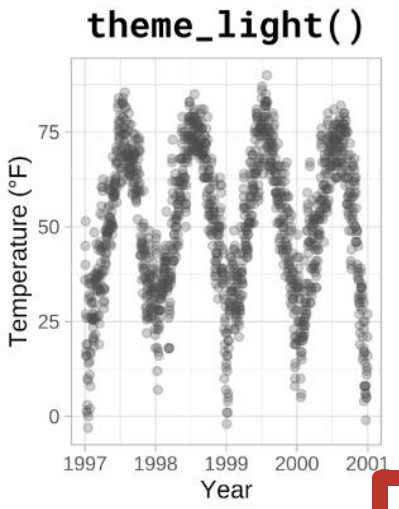
Setting your theme

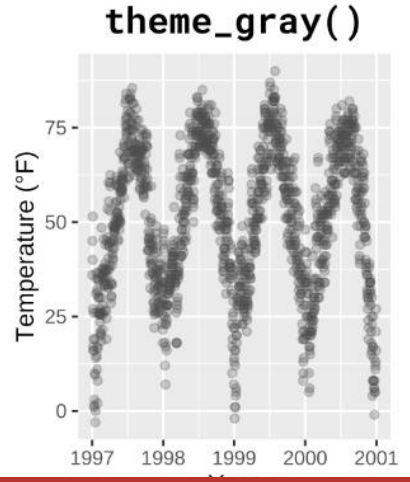


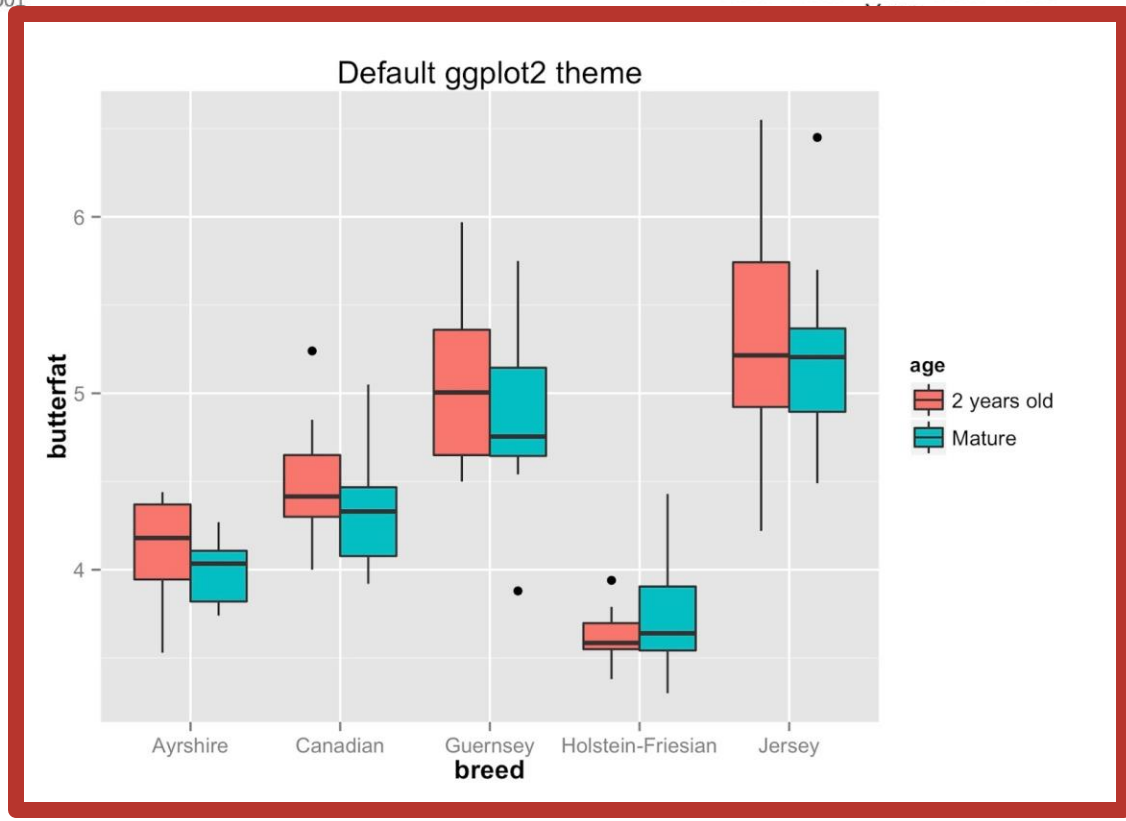
Setting your theme











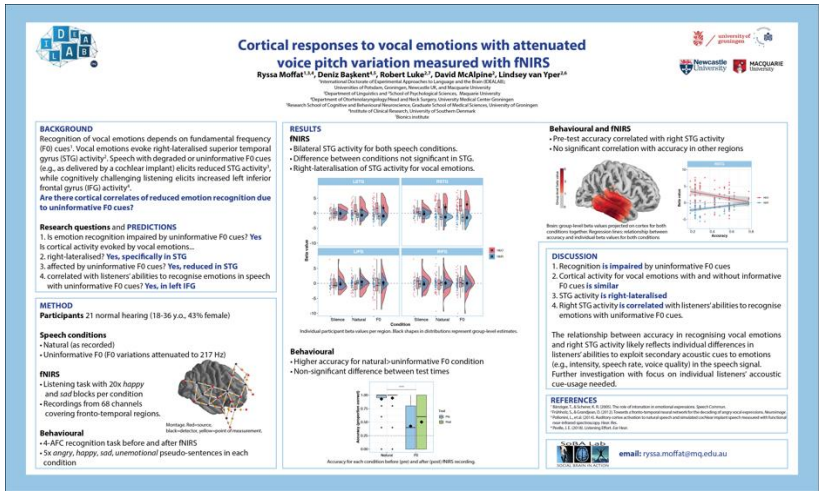
Setting your theme

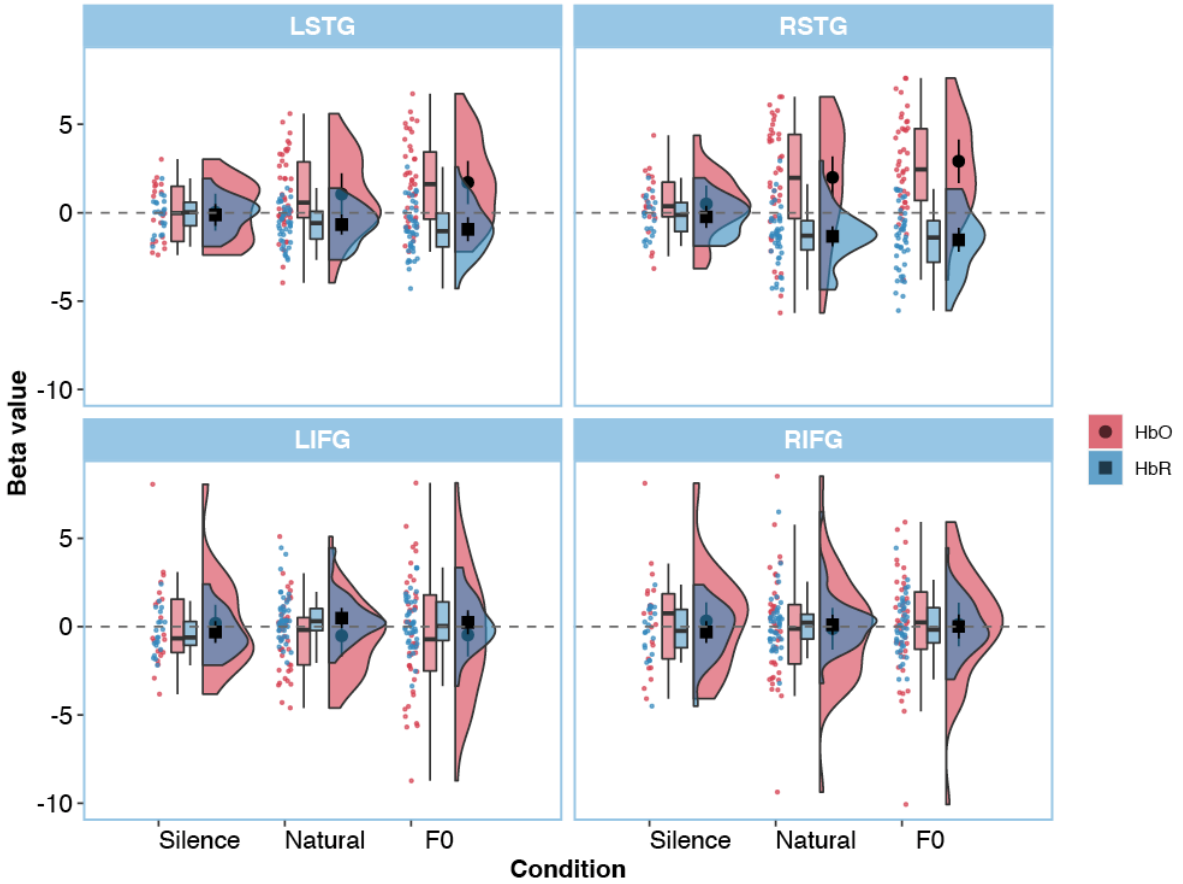
Why?

- Aesthetically pleasing posters
- A unique ”fingerprint”

What kinds of components?

1. Font (face, size, color)
2. Background, panels
3. Axis labels + ticks
4. So much more!

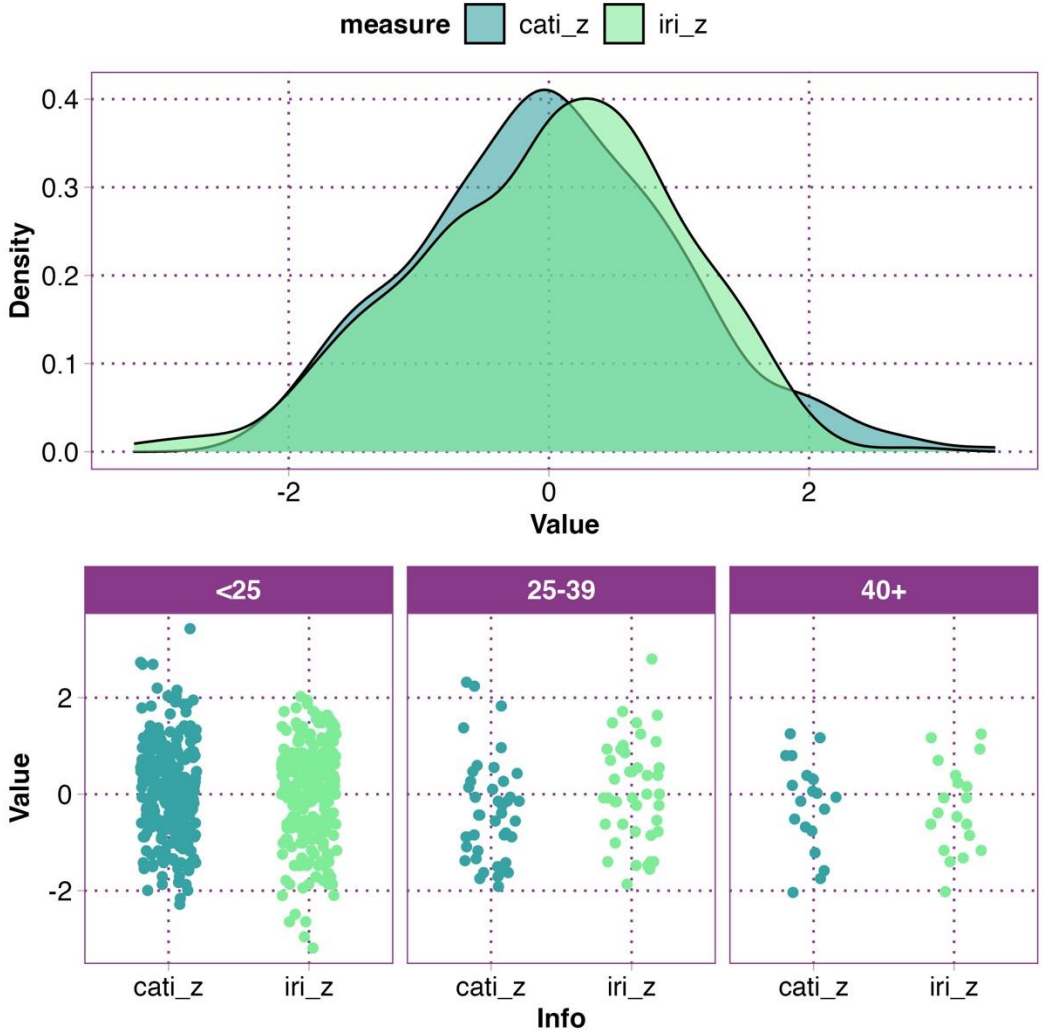




Composite figures

ggarrange()

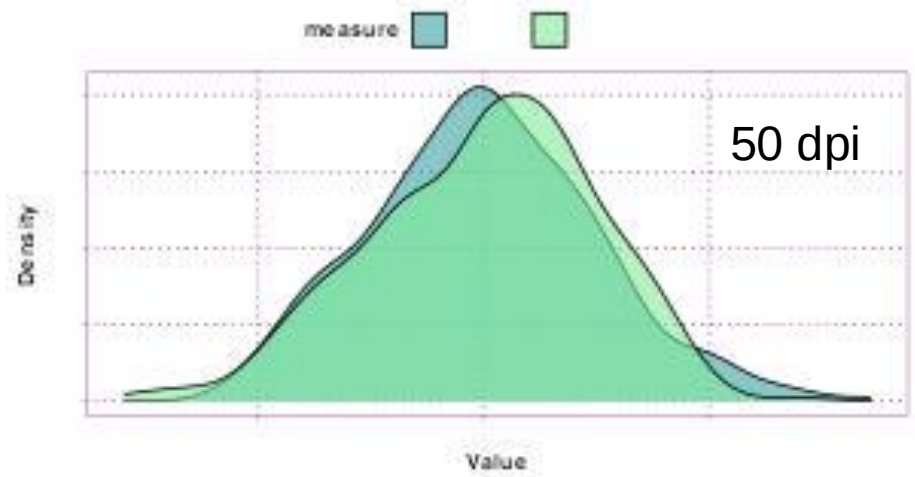
1. nrow/ncol
2. widths/heights
3. common legends

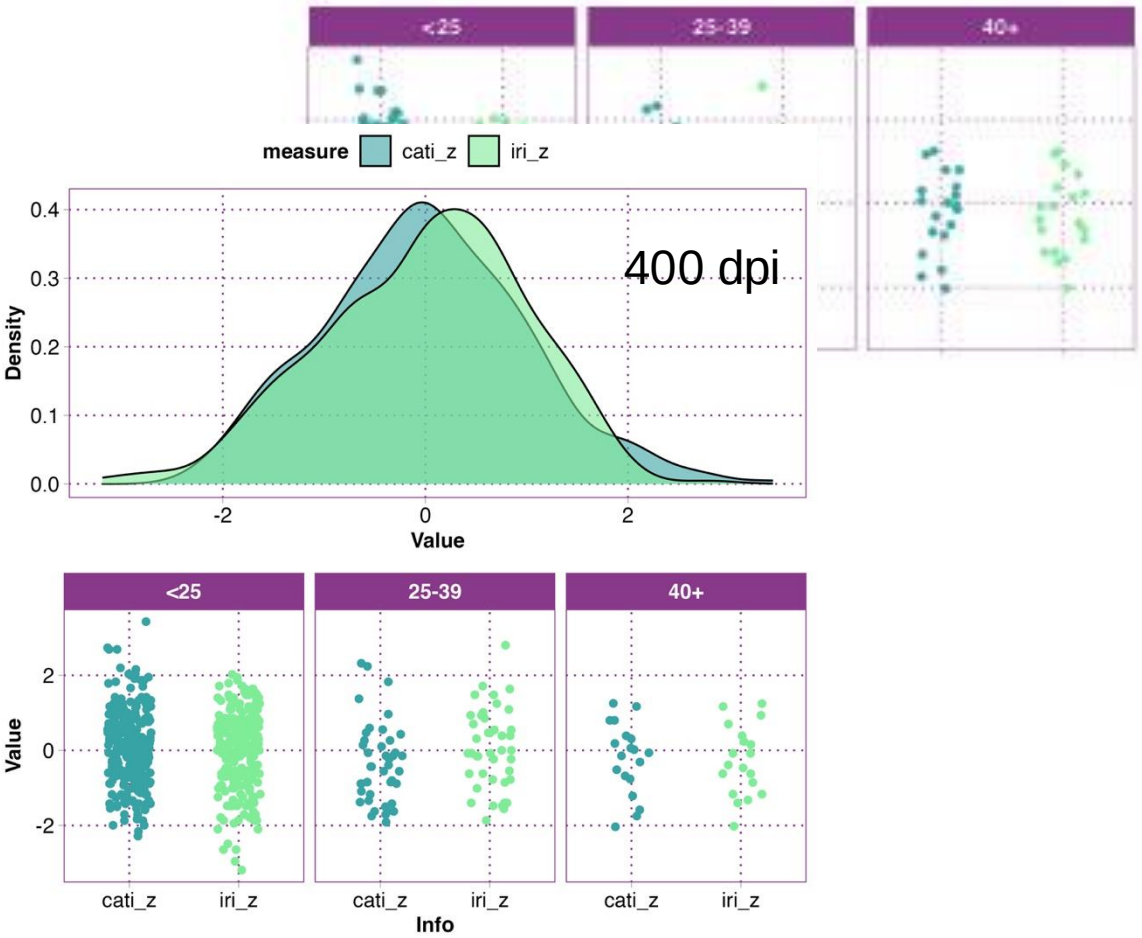


Saving figures

ggsave()

- Saves the last figure you ran
- Set resolution in DPI (dots per inch)
  - 300 minimum for print, more = better

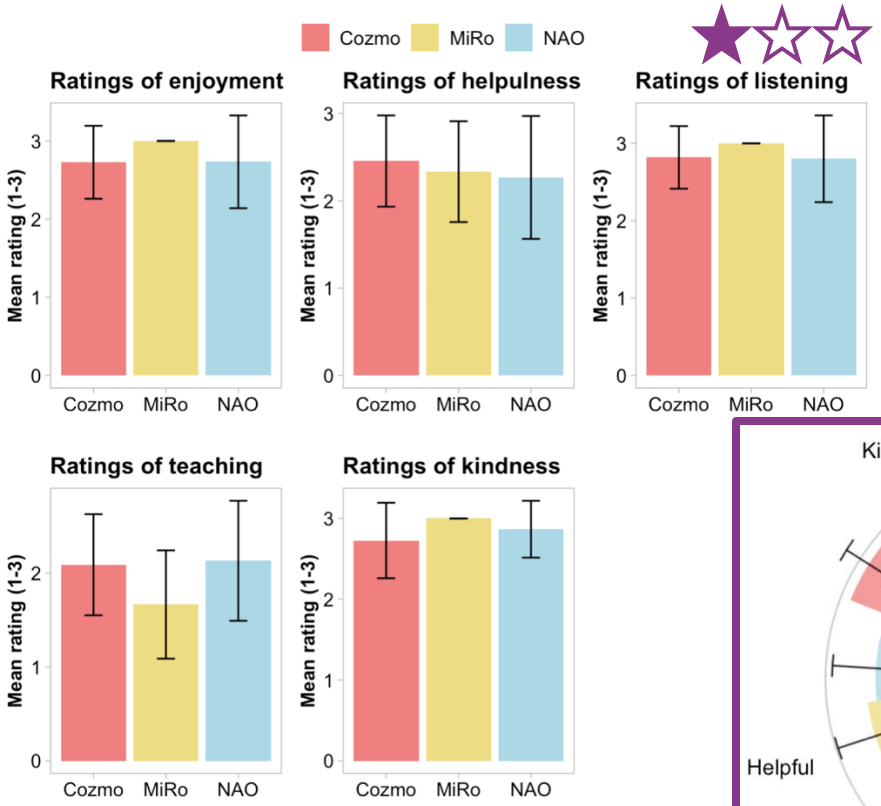


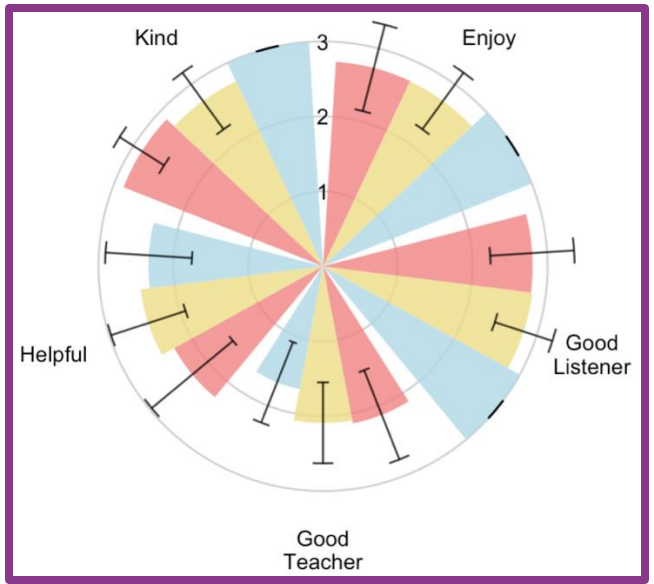


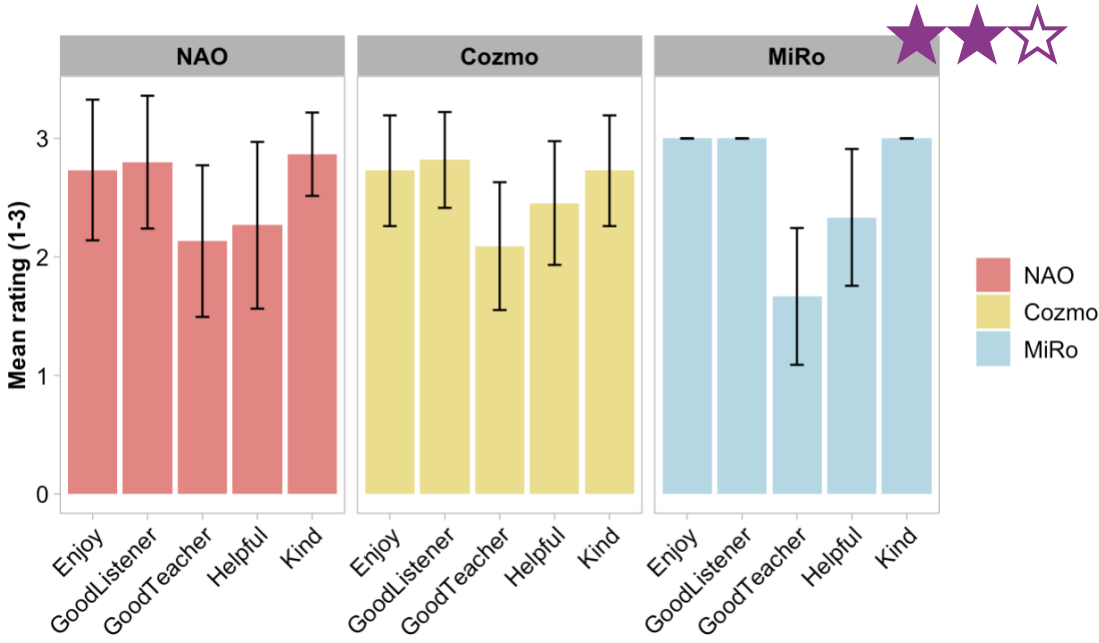
IMPORTANT STEPS

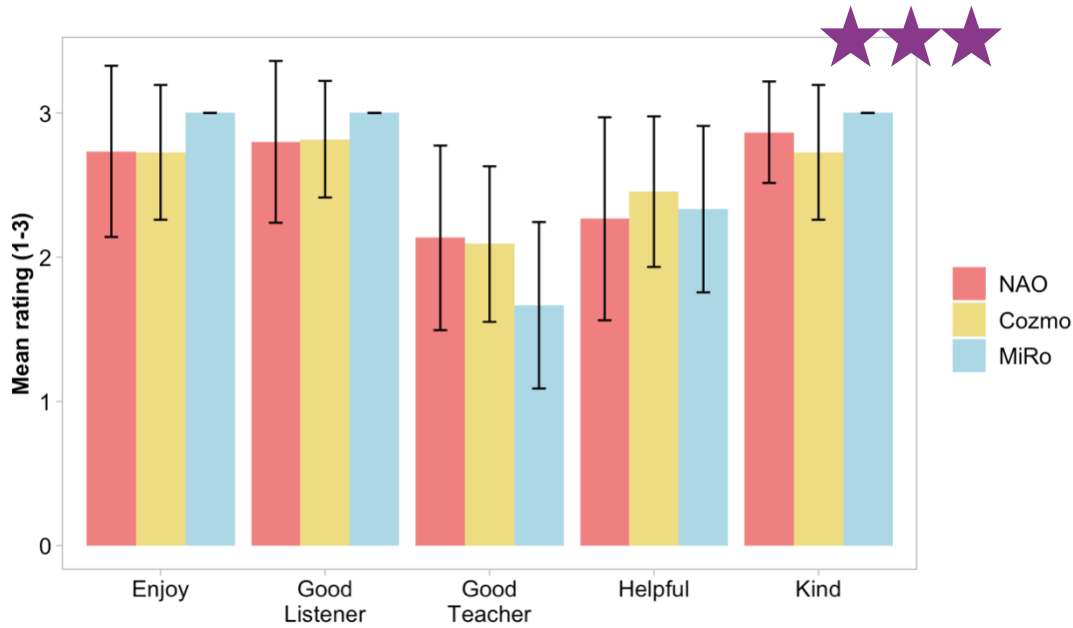
Relevant comparisons & Redundancy

- Put things to be compared side by side
- Try to only have each label once









Relevant comparisons & Redundancy

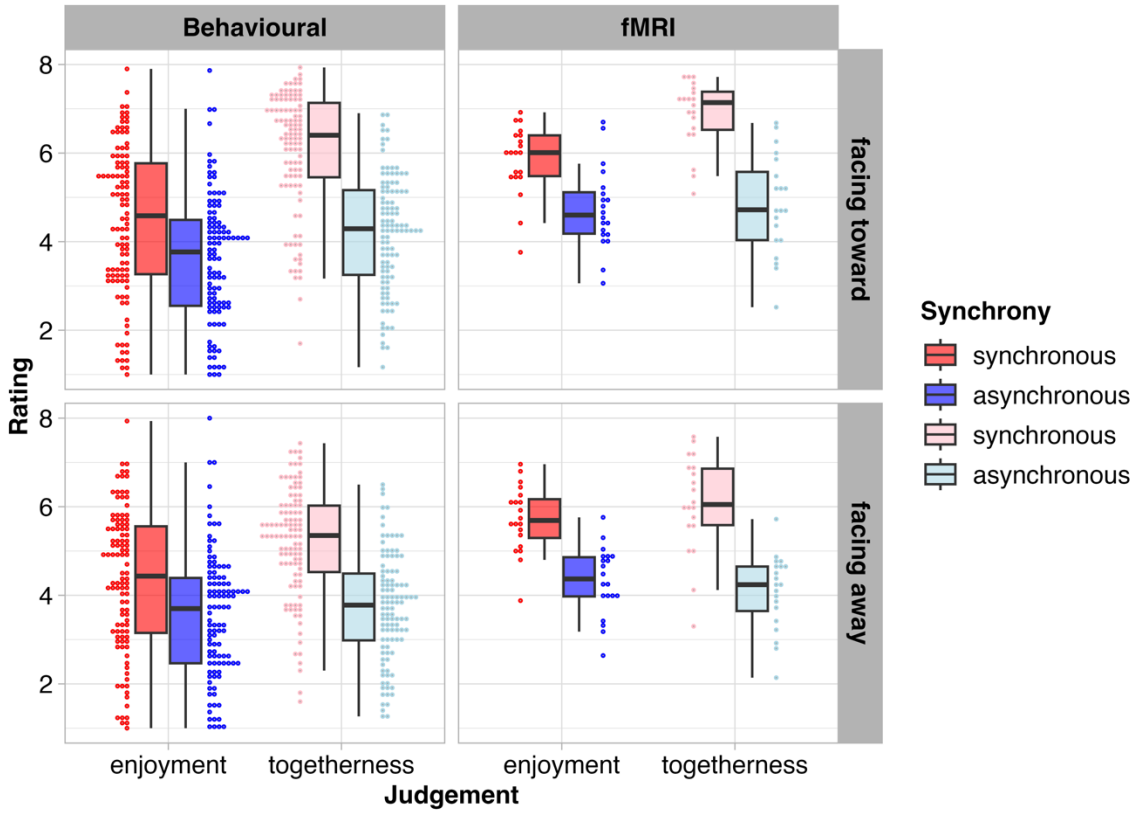
- Put things to be compared side by side
- Try to only have each label once

QUESTIONS TO CONSIDER:

- Are the conditions/items to be compared described in one column?
- Do any columns have enough in common to be combined and described in a new column?
- Should the numeric information be simplified and shared as a table elsewhere?

Relevant comparisons & Redundancy

- We have ratings on scale of 1-4 and ratings on scale of 1-8. How can we plot them together?



ID	synchrony	judgement	rating	experiment
Beh1	synchronous	enjoyment	4	behavioural
Beh2	asynchronous	togetherness	8	behavioural

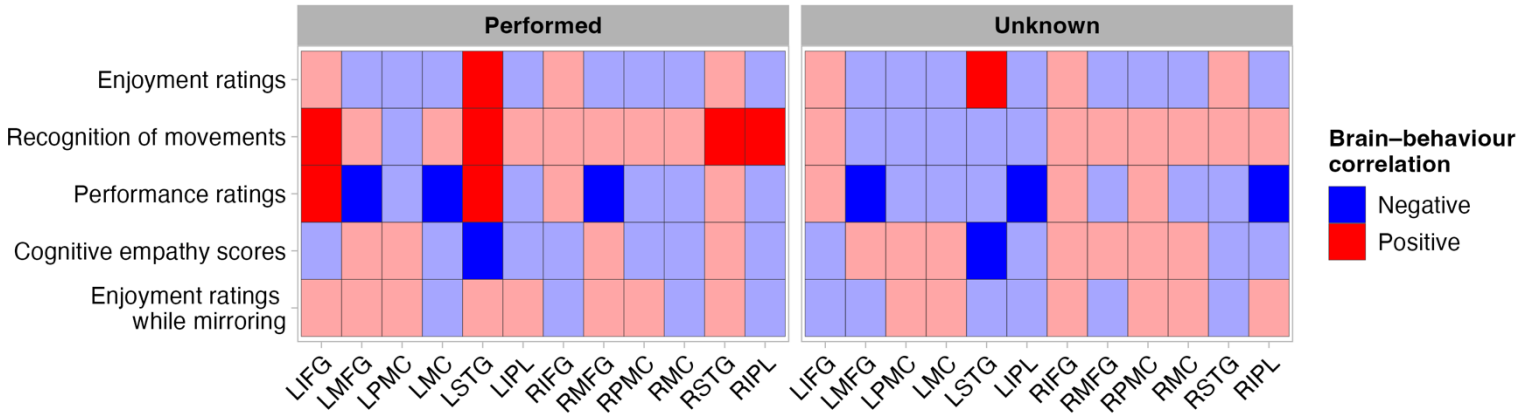
ID	synchrony	judgement	rating	experiment
Fmri1	synchronous	enjoyment	1	fMRI
Fmri2	asynchronous	togetherness	2	fMRI

ID	experiment	synchrony	judgement	rating
Beh1	behavioural	synchronous	enjoyment	4
Beh2	behavioural	asynchronous	togetherness	8
Fmri1	fMRI	synchronous	enjoyment	2
Fmri2	fMRI	asynchronous	togetherness	4

Relevant comparisons & Redundancy

%>%

- Simple heat maps can really help present multi-dimensonal data.



Judgement/Measure	Condition	Brain area	Value	Significant	Direction
Enjoyment	Performed	LIFG	0.04	no	postive
Recognition	Performed	LIFG	0.6	yes	positive
Performance	Performed	LIFG	0.4	yes	positive
Empathy	Performed	LIFG	-0.01	no	positive
Miroring enjoyment	Performed	LIFG	0.03	no	negative

Relevant comparisons & Redundancy

- Pivoting longer and wider are almost always the solution.
- Typically long data is more useful for plotting.
- But sometimes, the data must get wider before it gets longer.

ID	Item	Value
S1	Group	A
S1	Pic1	0.5
S1	Pic2	0.2
S1	Pic3	0.1
S2	Group	B
S2	Pic1	0.8
S2	Pic2	0.3
S2	Pic3	0.1

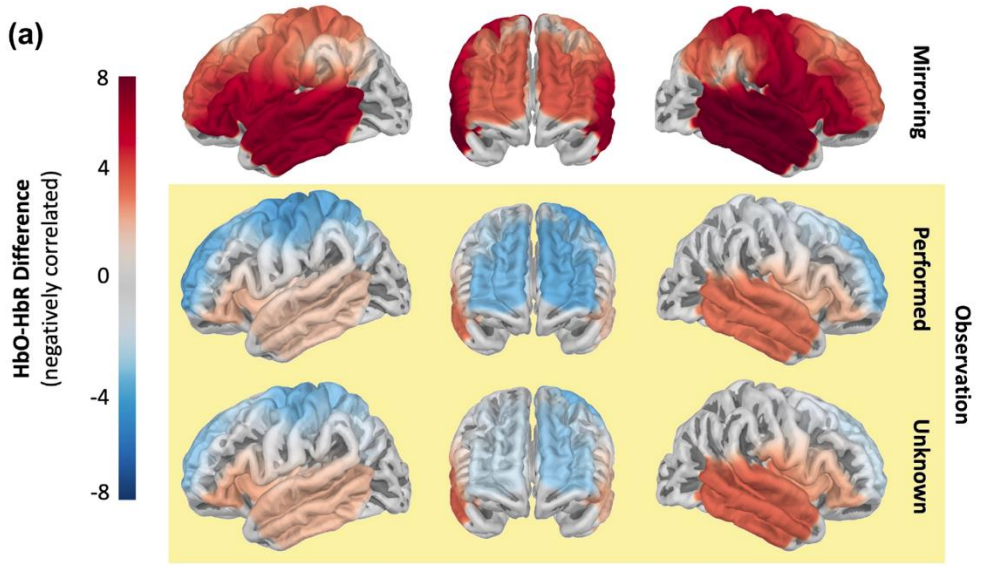
ID	Group	Pic1	Pic2	Pic3
S1	A	0.5	0.2	0.4
S2	B	0.8	0.3	0.1

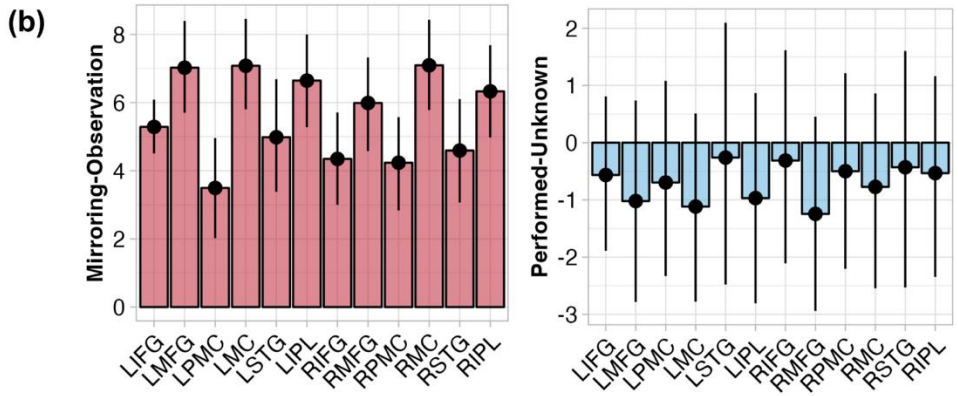
ID	Group	Picture	Rating	Price	Annimate
S1	A	Pic1	0.5	100	annimate
S1	A	Pic2	0.2	100	annimate
S1	A	Pic3	0.3	100	inannimate
S1	B	Pic1	0.8	20	annimate
S2	B	Pic2	0.3	20	annimate
S2	B	Pic3	0.1	20	inannimate

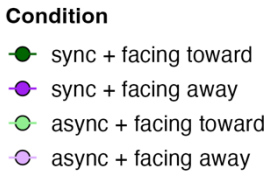
EXCITING THINGS

From models to plots

- Plotting raw data is **ALWAYS** the first step.
- Showing model parameter estimates and contrast estimates can also be helpful.







3D plots

Plotly 3D plots!