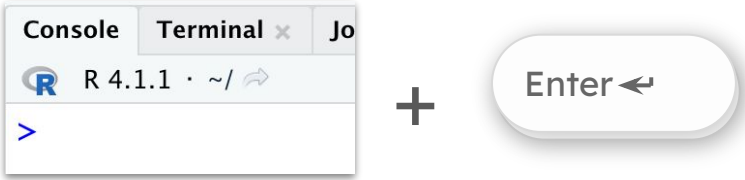
Slides:
https://docs.google.com/presentation/d/1tCmLc3myJKJL7wsn2sz5oUZi0sAkQxdwe6zSHASFELE/edit?usp=sharing
Files:
https://osf.io/x9zvy/files/osfstorage

Get Ready

1. **R**
   Wins: https://cran.r-project.org/bin/windows/base/
   Macs: https://cran.r-project.org/bin/macosx/
2. **RStudio**
   https://posit.co/download/rstudio-desktop/
3. **In RStudio, install packages**




```
> packages <- c("tidyverse","lme4", "lmerTest", "emmeans","ggeffects","modelbased")
> install.packages(packages)
```

```
> library(tidyverse); library(lme4); library(lmerTest); library(emmeans);
  library(ggeffects); library(modelbased)
```

조선대학교
CHOSUN UNIVERSITY

CHILD. LANGUAGE. LAB.

조선대학교
인문데이터과학연구소

Analyzing Mother-Child Interaction Data

Data Transformation and Linear Mixed Models

Chai Jun Ho

junhoc94@gmail.com

Multimodal Cue Usage in Child Development

- **Caregivers** use **multimodal cues** to **support** child **development** [1]
- For example, **eye gaze** in establishing **joint attention** and facilitating **language development** [2]
- **Children** use **non-verbal cues** to **communicate** [3]

[1] Abu-Zhaya et al., 2017; Ko et al., 2023; [2] Çetinçelik et al., 2021; Csibra, 2010; [3] Caselli et al., 2012

Objectives

Part I:

- Inspect the data
- Transform the data
- Model the data

Part II:

- Investigate interactions between variables
- Visualise the interactions

Think Faster, Talk Smarter with Matt Abrahams

Stanford Alumni
91.5K subscribers

Subscribe

Like | Share | ...

All | From Stanford Alumni | TED | Thoug

879. Think Fast, Talk Smart:

Part I

Objectives

1. Inspect the data
2. Conduct basic data transformation
3. Construct linear mixed effect model using *lme4* package
4. Inspect the model using *summary()* and car::*Anova()*

Get Ready

1. **R**

   Wins: https://cran.r-project.org/bin/windows/base/

   Macs: https://cran.r-project.org/bin/macosx/

2. **RStudio**

   https://posit.co/download/rstudio-desktop/

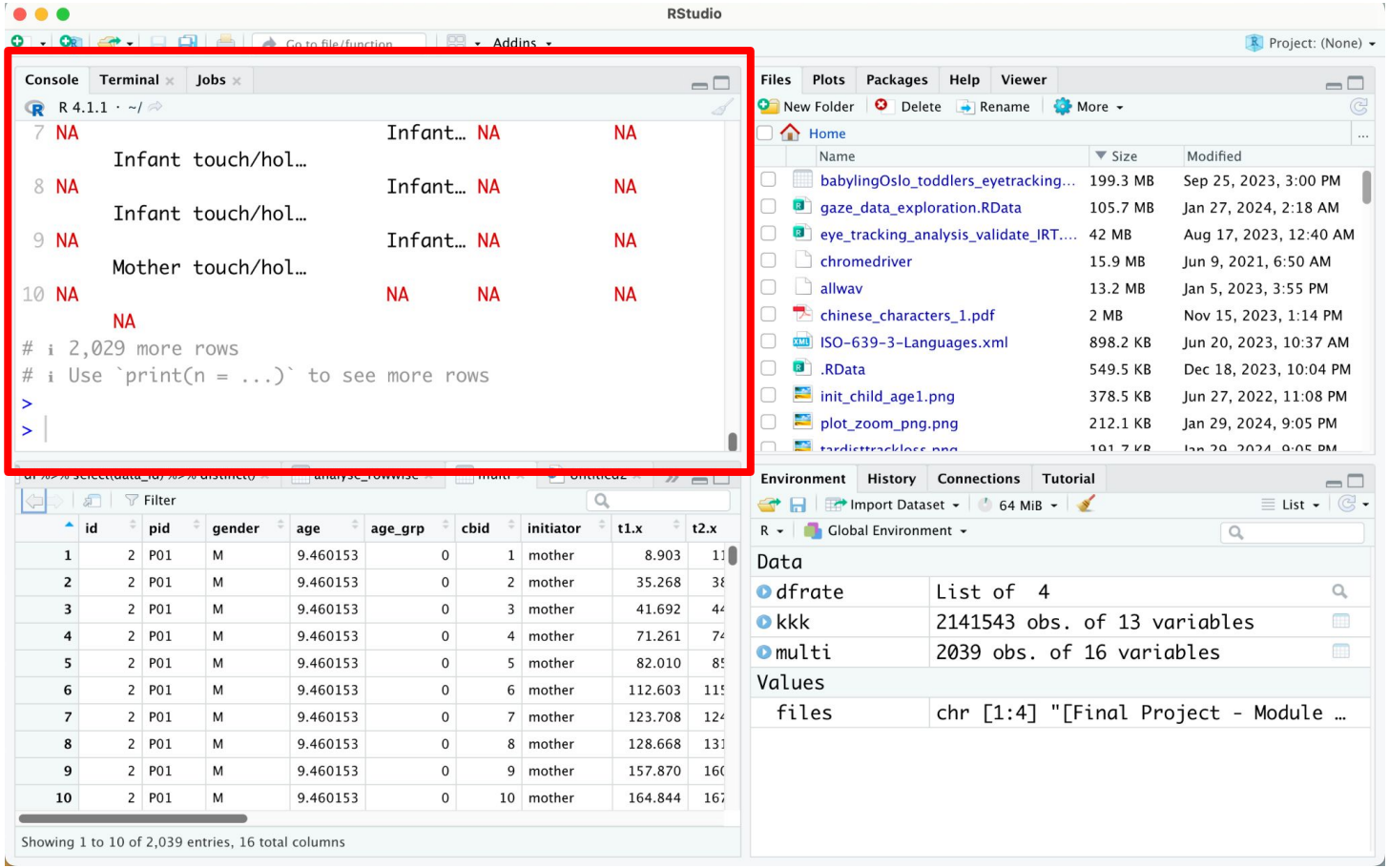
3. **In RStudio, install packages**


```
> packages <- c("tidyverse","lme4", "lmerTest", "emmeans","ggeffects","modelbased")
> install.packages(packages)
```

Get Ready

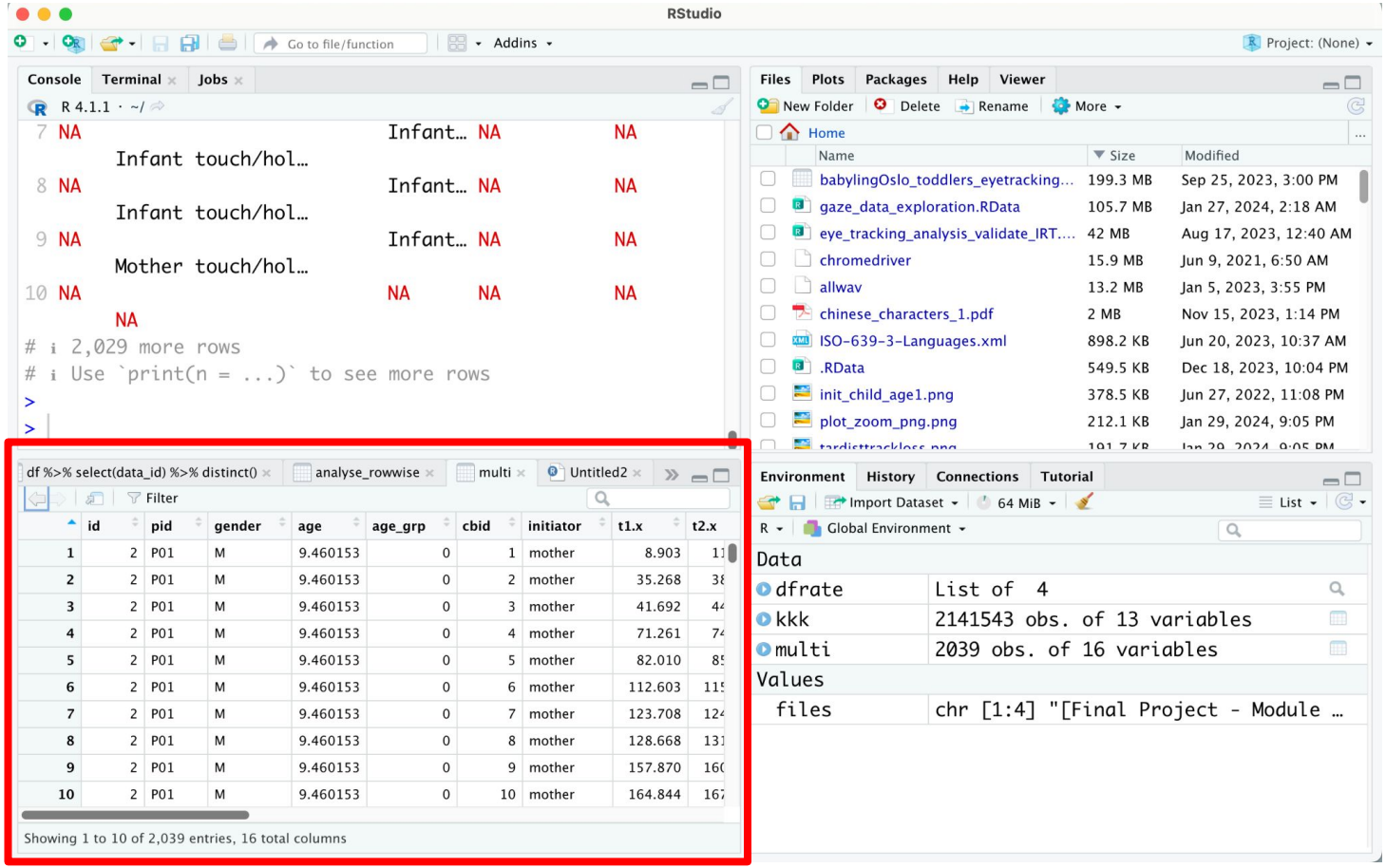
```
> library(tidyverse); library(lme4); library(lmerTest); library(emmeans);
  library(ggeffects); library(modelbased)
```

Learn about your tool



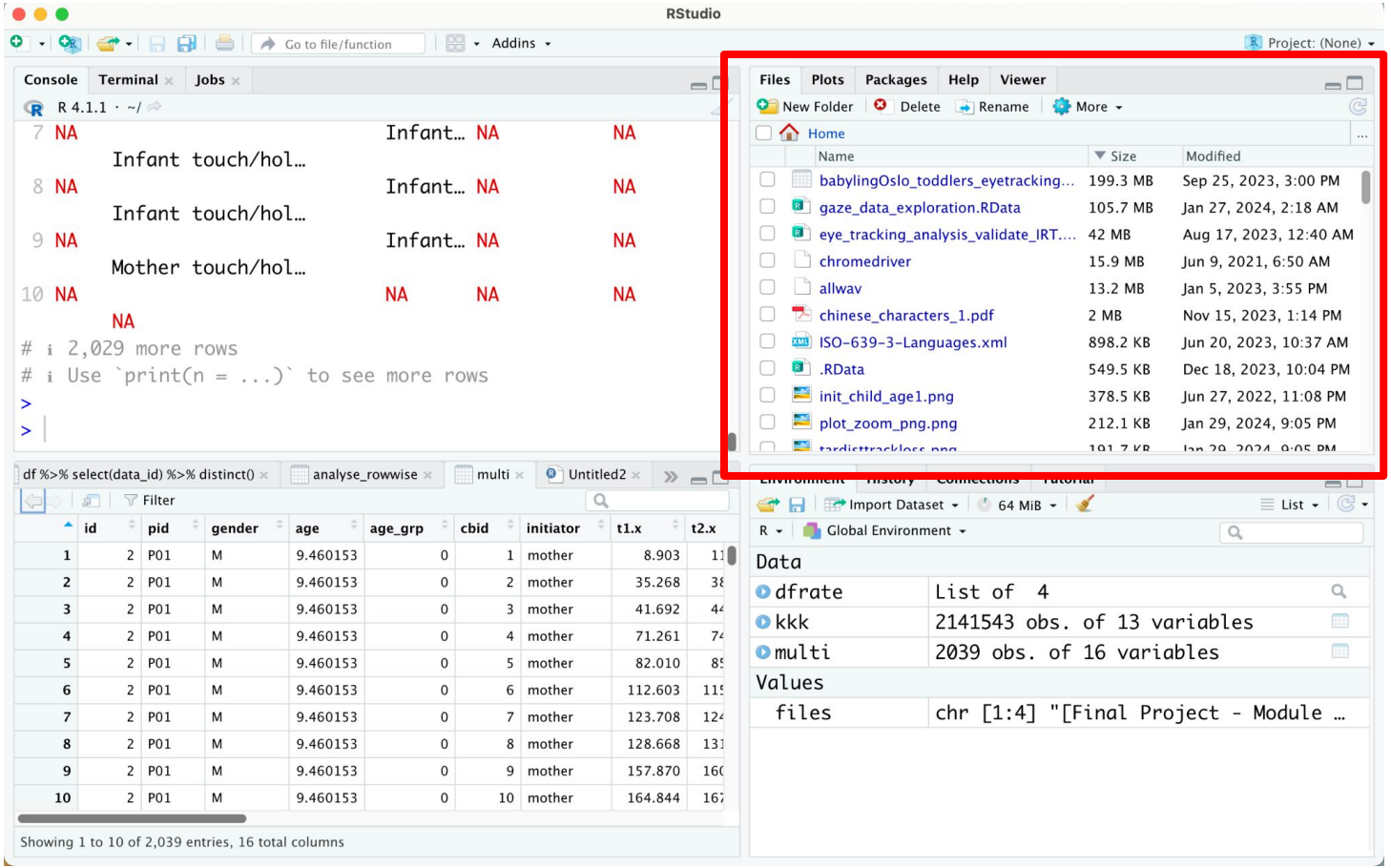
- Console
- Terminal
- Jobs
- *Render*

Learn about your tool



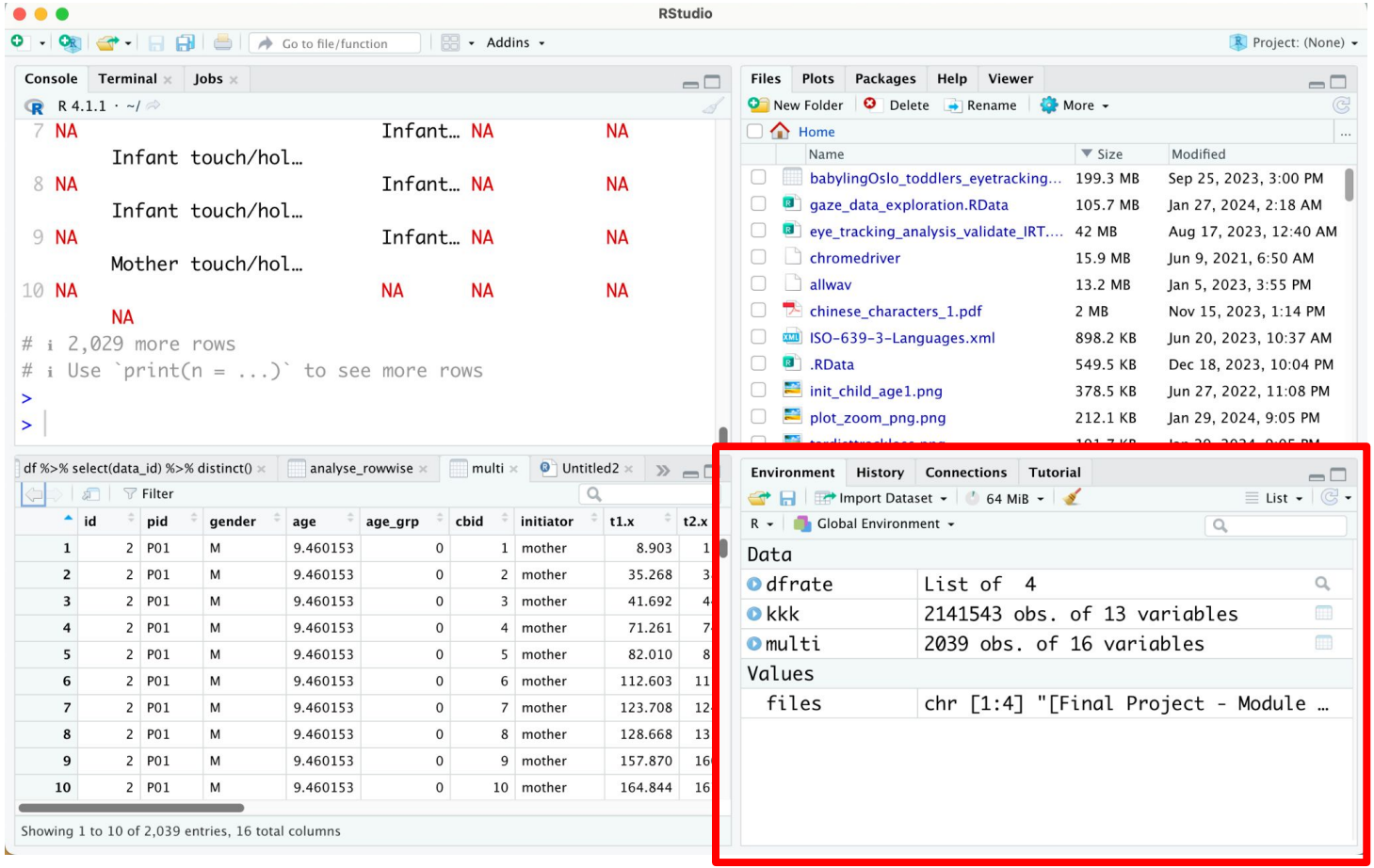
- Dataframe
- Scripts

Learn about your tool



- Files
- Plots
- Packages
- Help
- Viewer

Learn about your tool



- Environment
- History
- Connections
- Tutorial

Type something…

- Now, try:

```
> 1 * 2
> 1 / 2
> 2 ^ 2
> sqrt(1234)
```

Type something…

```
> x <- c(12, 34, 66)
> x / 4
> mean(x)
> sd(x)
```

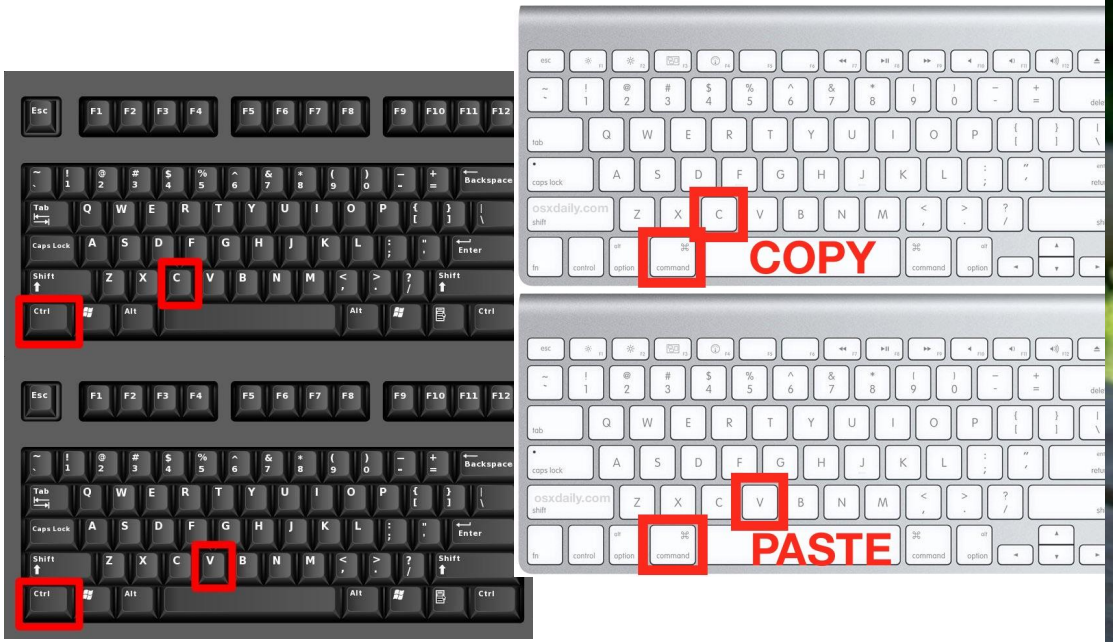
Type something…

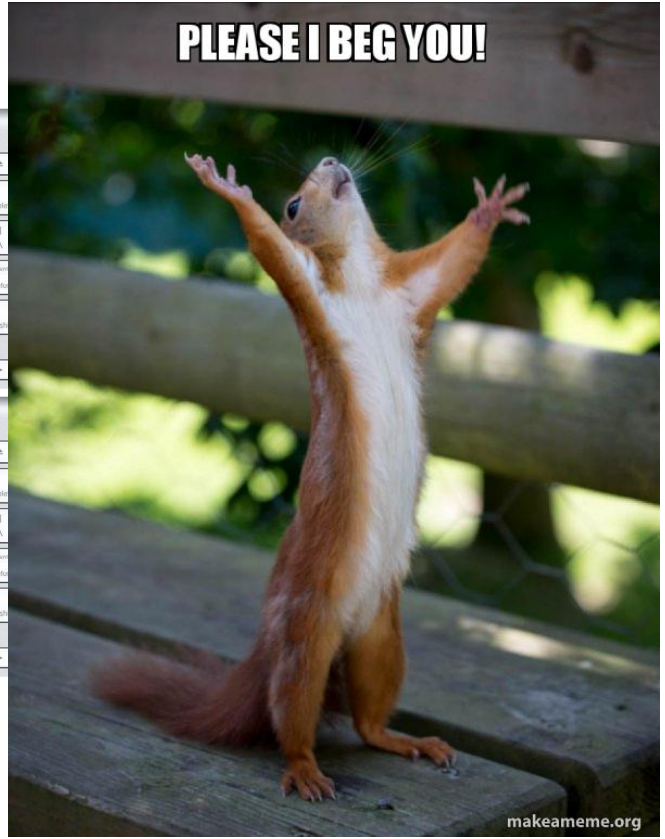
```
> head(iris)
> View(iris)
> summary(iris)
> mean(iris$Sepal.Width)
> sd(iris$Sepal.Width)
> t.test(iris$Sepal.Width, iris$Petal.Width)
```

Arrows

- Fall in love with the **up** and **down** arrow
- Save you time rewriting code 🚀

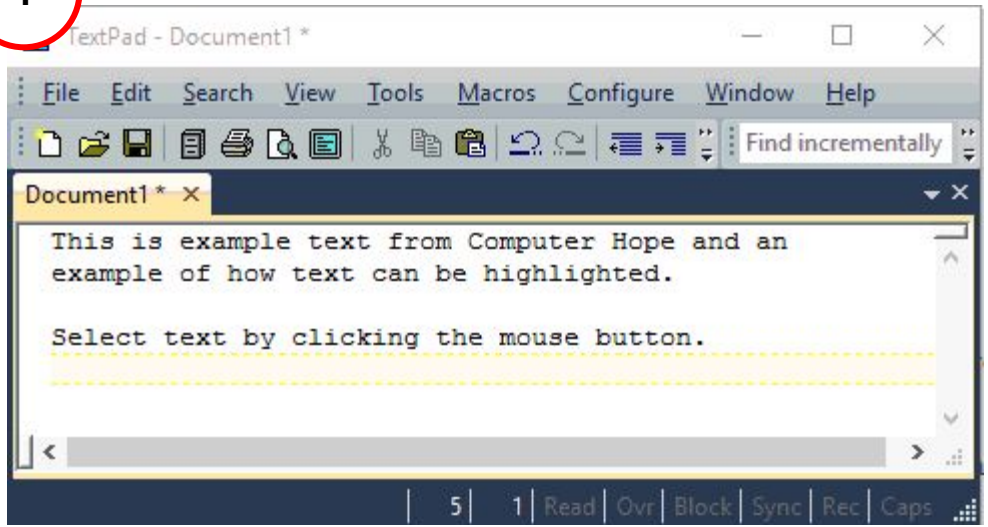
Copy and Paste



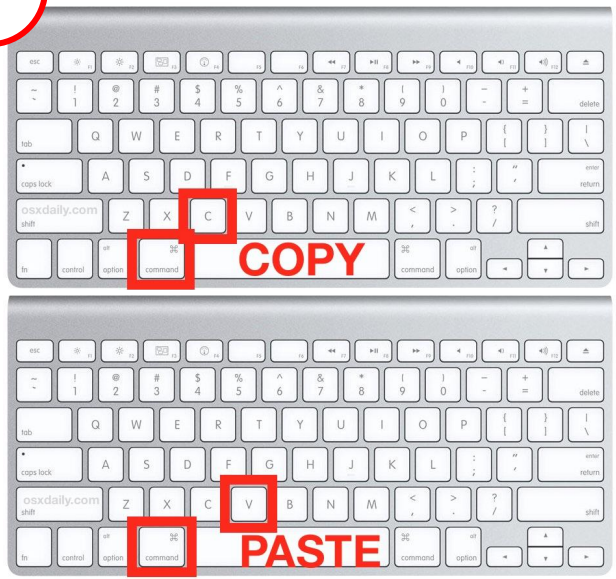


Copy and Paste

1



2



Set directory

For windows:

```
> setwd(normalizePath("~/Downloads", winval = "USERPROFILE"))
```

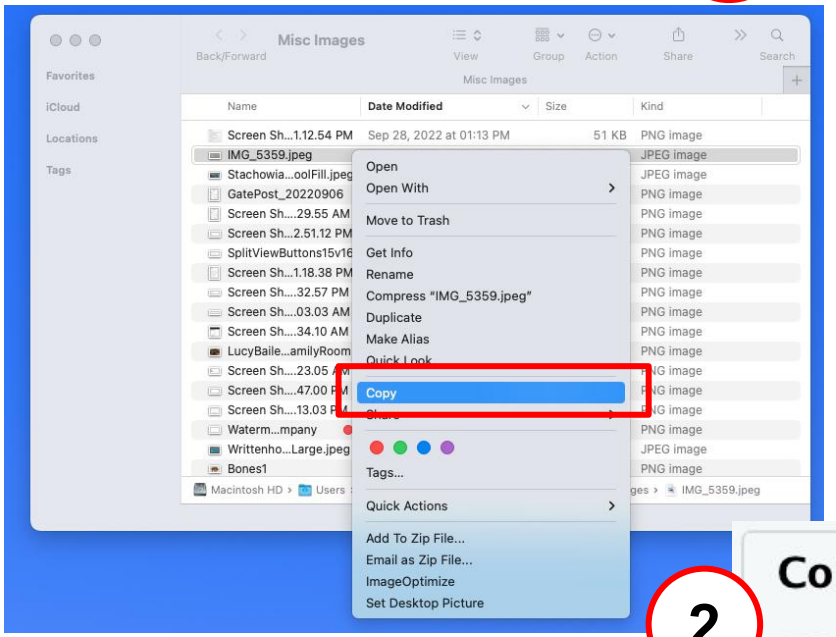
For macs:

```
> setwd("~/Downloads")
```

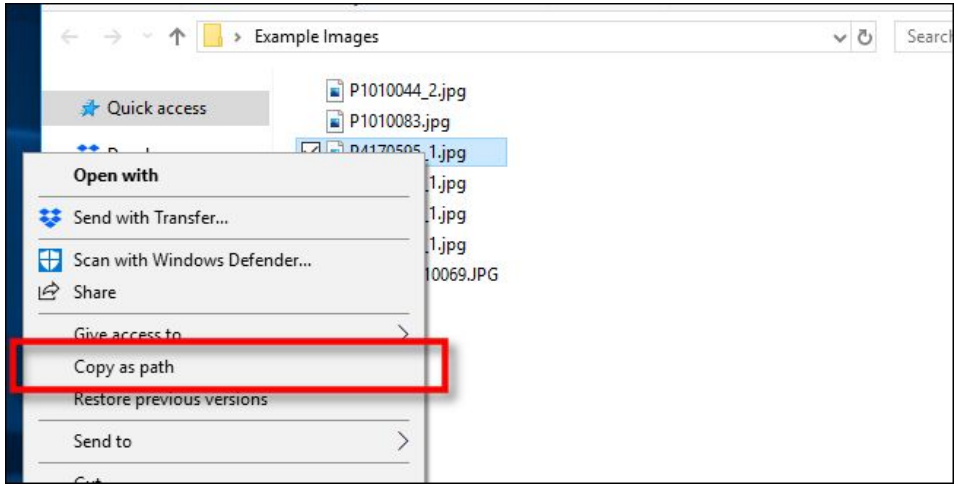
Obtain directory of file

For macs:

1

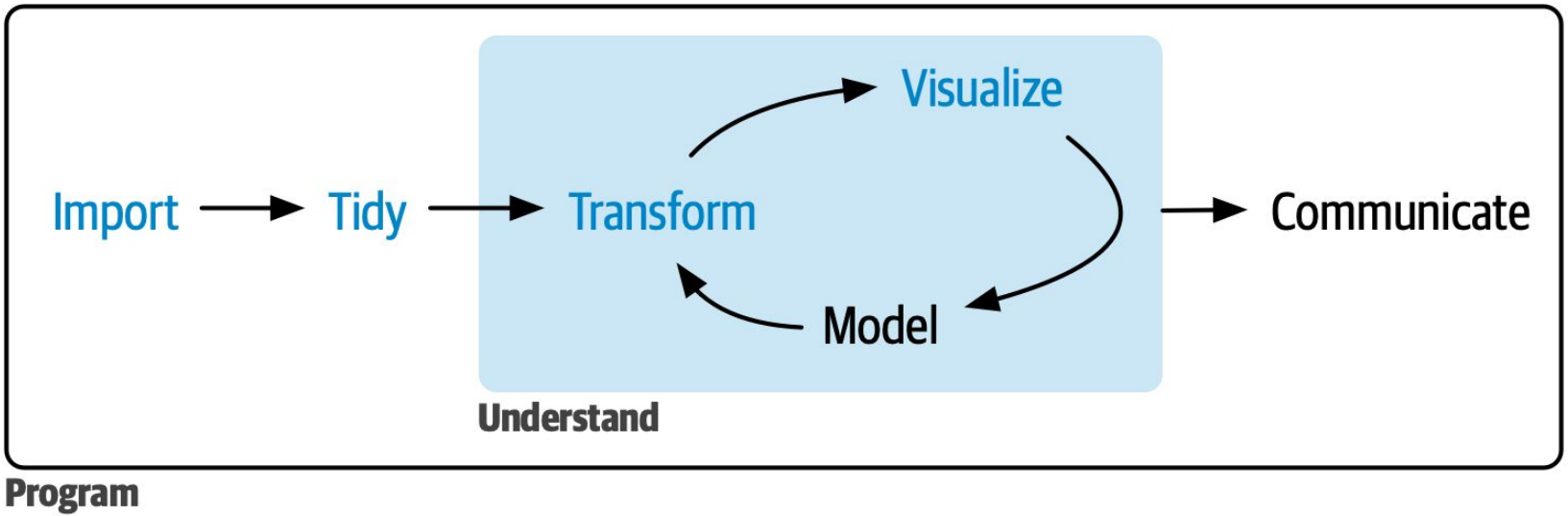


For windows:



2

Data Analysis Pipeline



Inspect data

1. Load the dataset

```
> multi <- read.csv("multi.csv")
```

Inspect data

1. Load the dataset

```
> multi <- read.csv("multi.csv")
```

2. Inspect the data

```
> head(multi)
> str(multi)
> summary(multi)
```

The data

```
> head(multi)
# A tibble: 6 × 16
     id pid   gender   age age_grp  cbid initiator   t1.x  t2.x GesturalCue     LookCue
  <dbl> <chr> <chr>  <dbl>   <dbl> <dbl> <chr>      <dbl> <dbl> <chr>           <chr>
1     2 P01   M       9.46       0     1 mother      8.90  11.9 NA              Infant…
2     2 P01   M       9.46       0     2 mother     35.3   38.3 NA              Mother…
3     2 P01   M       9.46       0     3 mother     41.7   44.7 Mother gestur… Mother…
4     2 P01   M       9.46       0     4 mother     71.3   74.3 NA              NA
5     2 P01   M       9.46       0     5 mother     82.0   85.0 NA              Infant…
6     2 P01   M       9.46       0     6 mother    113.   116.  NA              NA
# ℹ 5 more variables: PointingCue <chr>, TactileCue_Human <chr>,
#   TactileCue_Object <chr>, t1.y <dbl>, t2.y <dbl>
```

Try:

```
> colnames(multi)
```

```
> summary(multi)
       id              pid                gender               age
 Min.   : 2.00   Length:2039        Length:2039        Min.   : 6.066
 1st Qu.:12.00   Class :character   Class :character   1st Qu.: 9.460
 Median :22.00   Mode  :character   Mode  :character   Median :13.132
 Mean   :20.28                                         Mean   :15.116
 3rd Qu.:28.00                                         3rd Qu.:25.460
 Max.   :37.00                                         Max.   :30.789
    age_grp            cbid         initiator             t1.x
 Min.   :0.0000   Min.   :  1.0   Length:2039        Min.   :   3.988
 1st Qu.:0.0000   1st Qu.: 18.0   Class :character   1st Qu.: 748.777
 Median :1.0000   Median : 35.0   Mode  :character   Median :1432.124
 Mean   :0.8823   Mean   : 40.6                      Mean   :1439.758
 3rd Qu.:2.0000   3rd Qu.: 59.0                      3rd Qu.:2076.384
 Max.   :2.0000   Max.   :133.0                      Max.   :4504.613
      t2.x          GesturalCue          LookCue          PointingCue
 Min.   :   6.988   Length:2039        Length:2039        Length:2039
 1st Qu.: 752.999   Class :character   Class :character   Class :character
 Median :1432.792   Mode  :character   Mode  :character   Mode  :character
 Mean   :1443.507
 3rd Qu.:2079.588
 Max.   :4505.657
 TactileCue_Human   TactileCue_Object       t1.y              t2.y
 Length:2039        Length:2039        Min.   :   6.977   Min.   :   9.31
 Class :character   Class :character   1st Qu.: 749.186   1st Qu.: 750.16
 Mode  :character   Mode  :character   Median :1430.534   Median :1431.29
                                       Mean   :1437.136   Mean   :1438.64
                                       3rd Qu.:2075.214   3rd Qu.:2077.76
                                       Max.   :4397.769   Max.   :4398.98
```

Inspect data

- We have 5 categories of cues consisting the string “Cue”:

```
> grepl("Cue", colnames(multi))
```

```
 [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  TRUE  TRUE  TRUE  TRUE
[14]  TRUE FALSE FALSE
```

Inspect data

- We have 5 categories of cues consisting the string “Cue”:

```
> multi[,grepl("Cue", colnames(multi))]
```

```
# A tibble: 2,039 × 5
   GesturalCue                    LookCue  PointingCue TactileCue_Human TactileCue_Object
   <chr>                          <chr>    <chr>       <chr>             <chr>
 1 NA                             Infant…  NA          Mother touch in…  NA
 2 NA                             Mother…  NA          NA                Mother touch/hol…
 3 Mother gesture to object/ac…   Mother…  NA          NA                NA
 4 NA                             NA       NA          NA                NA
 5 NA                             Infant…  NA          NA                Infant touch/hol…
 6 NA                             NA       NA          NA                NA
 7 NA                             Infant…  NA          NA                Infant touch/hol…
 8 NA                             Infant…  NA          NA                Infant touch/hol…
 9 NA                             Infant…  NA          NA                Mother touch/hol…
```

dataframe[row,column]

```
> multi[,grepl("Cue", colnames(multi))]
```

column

row

Filter

	id	pid	gender	age	age_grp	cbid	initiator	t1.x	t2.x	GesturalCue	LookCu
1	2	P01	M	9.460153	0	1	mother	8.903	11.903	*NA*	Infant a
2	2	P01	M	9.460153	0	2	mother	35.268	38.268	*NA*	Mother
3	2	P01	M	9.460153	0	3	mother	41.692	44.692	Mother gesture to object/action	Mother
4	2	P01	M	9.460153	0	4	mother	71.261	74.261	*NA*	*NA*
5	2	P01	M	9.460153	0	5	mother	82.010	85.010	*NA*	Infant a
6	2	P01	M	9.460153	0	6	mother	112.603	115.603	*NA*	*NA*
7	2	P01	M	9.460153	0	7	mother	123.708	124.245	*NA*	Infant a
8	2	P01	M	9.460153	0	8	mother	128.668	131.668	*NA*	Infant a
9	2	P01	M	9.460153	0	9	mother	157.870	160.870	*NA*	Infant a
10	2	P01	M	9.460153	0	10	mother	164.844	167.844	*NA*	*NA*
11	2	P01	M	9.460153	0	11	mother	169.582	172.582	*NA*	*NA*

Inspect data

- And each category have their own sub-categories:

```
> lapply(multi[,grepl("Cue", colnames(multi))], unique)
```

```
$GesturalCue
[1] NA                                "Mother gesture to object/action"
[3] "Infant gesture to object/action"

$LookCue
[1] "Infant and mother look object/action/face/body part(s)"
[2] "Mother look object/action/face/body part(s) (initiator of CB)"
[3] "Mother following infant look object/action/face/body part(s)"
[4] NA
[5] "Infant look object/action/face/body part(s) (initiator of CB)"
[6] "Infant following mother look object/action/face/body part(s)"

$PointingCue
[1] NA                               "Infant point to object/action" "Mother point to object/action"

$TactileCue_Human
[1] "Mother touch infant"                NA
[3] "Infant and mother touch each other" "Infant touch mother"

$TactileCue_Object
[1] NA
[2] "Mother touch/hold object"
[3] "Infant touch/hold object"
[4] "Infant and mother touch/hold object"
[5] "Infant and mother touch/hold object Infant touch/hold object"
```

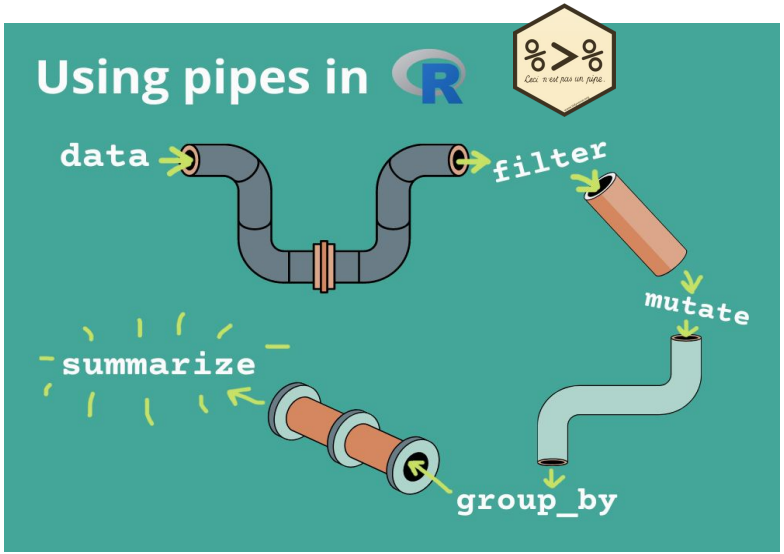
The way of tidyverse/magrittr

```
> multi %>%
  .[, grepl("Cue", colnames(.))] %>%
  lapply(unique)
```


```
> lapply(multi[,grepl("Cue", colnames(multi))], unique)
```

The way of tidyverse/magrittr

```
> multi %>%
  .[, grepl("Cue", colnames(.))] %>%
  lapply(unique)
```




```
> lapply(multi[,grepl("Cue", colnames(multi))], unique)
```

The way of tidyverse/magrittr

```
> multi %>%
  .[, grepl("Cue", colnames(.))] %>%
  lapply(unique)
```

```
> lapply(multi[,grepl("Cue",
  colnames(multi))], unique)
```

Research Question: What kind of cues precede conversation islands?

```
$GesturalCue
[1] NA                                  "Mother gesture to object/action"
[3] "Infant gesture to object/action"

$LookCue
[1] "Infant and mother look object/action/face/body part(s)"
[2] "Mother look object/action/face/body part(s) (initiator of CB)"
[3] "Mother following infant look object/action/face/body part(s)"
[4] NA
[5] "Infant look object/action/face/body part(s) (initiator of CB)"
[6] "Infant following mother look object/action/face/body part(s)"

$PointingCue
[1] NA                                "Infant point to object/action" "Mother point to object/action"

$TactileCue_Human
[1] "Mother touch infant"                NA
[3] "Infant and mother touch each other" "Infant touch mother"

$TactileCue_Object
[1] NA
[2] "Mother touch/hold object"
[3] "Infant touch/hold object"
[4] "Infant and mother touch/hold object"
[5] "Infant and mother touch/hold object Infant touch/hold object"
```

What do we need?

Research Question: What kind of cues precede conversation islands?

- Count things
- Things = Cues
- Let’s focus on the categories of cues

How can we do it?

Research Question: What kind of cues precede conversation islands?

```
> multi %>%
  .[, grepl("Cue", colnames(.))] %>%
  lapply(function(x) table(x, multi$id, multi$age_grp)) %>%
  lapply(as.data.frame) %>%
  bind_rows(.id = "Cue")
```

How can we do it?

Research Question: What kind of cues precede conversation islands?

```
> multi %>%
  .[, grepl("Cue", colnames(.))]
```

```
# A tibble: 2,039 × 5
   GesturalCue                    LookCue  PointingCue TactileCue_Human TactileCue_Object
   <chr>                          <chr>    <chr>       <chr>             <chr>
 1 NA                             Infant…  NA          Mother touch in…  NA
 2 NA                             Mother…  NA          NA                Mother touch/hol…
 3 Mother gesture to object/ac…   Mother…  NA          NA                NA
 4 NA                             NA       NA          NA                NA
 5 NA                             Infant…  NA          NA                Infant touch/hol…
 6 NA                             NA       NA          NA                NA
 7 NA                             Infant…  NA          NA                Infant touch/hol…
 8 NA                             Infant…  NA          NA                Infant touch/hol…
 9 NA                             Infant…  NA          NA                Mother touch/hol…
```

How can we do it?

Research Question: What kind of cues precede conversation islands?

```
> multi %>%
  .[, grepl("Cue", colnames(.))] %>%
  lapply(function(x) table(x, multi$id, multi$age_grp))
```

%>% View()

Name	Type	Value
.	list [5]	List of length 5
GesturalCue	integer [2 x 34 x 3] (S3: table	1 1 0 0 0 0 0 0 3 0 1 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
LookCue	integer [5 x 34 x 3] (S3: table	28 0 3 9 18 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 15 0 3 5 6 29
PointingCue	integer [2 x 34 x 3] (S3: table	2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
TactileCue_Human	integer [3 x 34 x 3] (S3: table	0 0 35 0 0 0 0 0 0 0 0 0 0 0 11 0 0 5 0 0 0 0 0 0 0 0 0
TactileCue_Object	integer [4 x 34 x 3] (S3: table	11 0 22 9 0 0 0 0 0 0 0 0 0 0 0 4 0 20 1 9 0 26 7 0 0

How can we do it?

Research Question: What kind of cues precede conversation islands?

```
> multi %>%
  .[, grepl("Cue", colnames(.))] %>%
  lapply(function(x) table(x, multi$id, multi$age_grp)) %>%
  .[["GesturalCue"]]
```

	x	Var2	Var3	Freq
1	Infant gesture to object/action	2	0	1
2	Mother gesture to object/action	2	0	1
3	Infant gesture to object/action	3	0	0

lapply() & function()

```
>    lapply(function(x) table(x, multi$id, multi$age_grp))
```

function()

lapply()

Many Things

Things you want

table()

unique()

sum()

na.omit()

function()

tal)

t()

Google

https://github.com/rstudio/cheatsheets/blob/main/base-r.pdf

ChatGPT

perplexity

How can we do it?

Research Question: What kind of cues precede conversation islands?

```
> multi %>%
  .[, grepl("Cue", colnames(.))] %>%
  lapply(function(x) table(x, multi$id, multi$age_grp))
```

	x	Var2	Var3	Freq
1	Infant gesture to object/action	2	0	1
2	Mother gesture to object/action	2	0	1
3	Infant gesture to object/action	3	0	0

How can we do it?

Research Question: What kind of cues precede conversation islands?

```
> multi %>%
  .[, grepl("Cue", colnames(.))] %>%
  lapply(function(cues) table(cues, id = multi$id, age_grp =
  multi$age_grp))
```

	cues	id	age_grp	Freq
1	Infant gesture to object/action	2	0	1
2	Mother gesture to object/action	2	0	1
3	Infant gesture to object/action	3	0	0

How can we do it?

Research Question: What kind of cues precede conversation islands?

```
> multi %>%
  .[, grepl("Cue", colnames(.))] %>%
  lapply(function(cues) table(cues, id = multi$id, age_grp =
  multi$age_grp)) %>%
  lapply(as.data.frame)
```

.	list [5]	List of length 5
GesturalCue	list [24 x 3] (S3: data.frame)	A data.frame with 24 rows and 3 columns
LookCue	list [60 x 3] (S3: data.frame)	A data.frame with 60 rows and 3 columns
PointingCue	list [24 x 3] (S3: data.frame)	A data.frame with 24 rows and 3 columns
TactileCue_Human	list [36 x 3] (S3: data.frame)	A data.frame with 36 rows and 3 columns
TactileCue_Object	list [48 x 3] (S3: data.frame)	A data.frame with 48 rows and 3 columns

How can we do it?

Research Question: What kind of cues precede conversation islands?

```
> multi %>%
  .[, grepl("Cue", colnames(.))] %>%
  lapply(function(cues) table(cues, id = multi$id, age_grp =
  multi$age_grp)) %>%
  lapply(as.data.frame)  %>%
  bind_rows(.id = "Cue")
```

	Cue	cues	id	age_grp	Freq
1	GesturalCue	Infant gesture to object/action	2	0	1
2	GesturalCue	Mother gesture to object/action	2	0	1

Data for Analysis

Research Question: What kind of cues precede conversation islands?

```
> multi %>%
  .[, grepl("Cue", colnames(.))] %>%
  lapply(function(cues) table(cues, id = multi$id, age_grp =
  multi$age_grp)) %>%
  lapply(as.data.frame)  %>%
  bind_rows(.id = "Cue") -> dfmod
```

	Cue	cues	id	age_grp	Freq
1	GesturalCue	Infant gesture to object/action	2	0	1
2	GesturalCue	Mother gesture to object/action	2	0	1

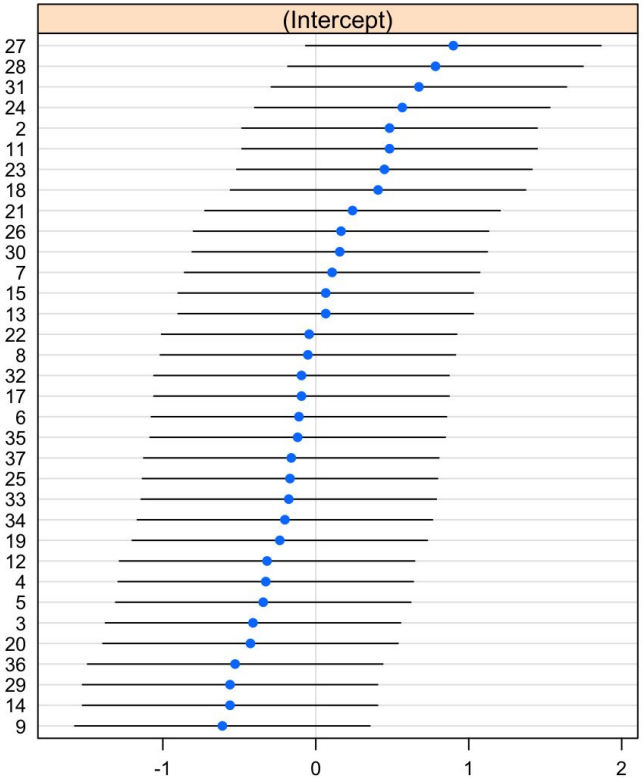
Linear Mixed-effects Model

Research Question: What kind of cues precede conversation islands?

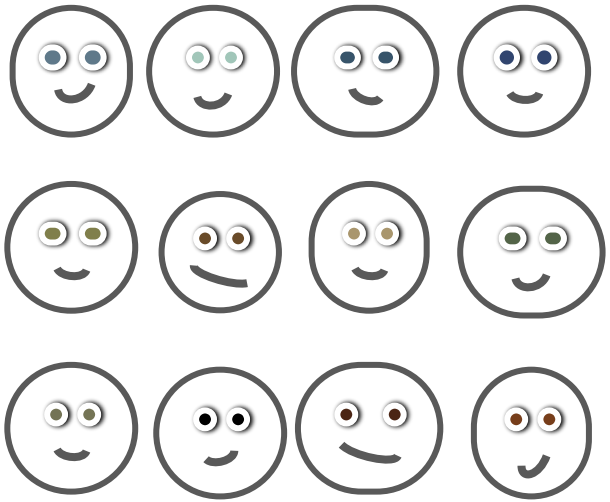
lmer(outcome ~ predictors +

(1 | random effect), data = data)

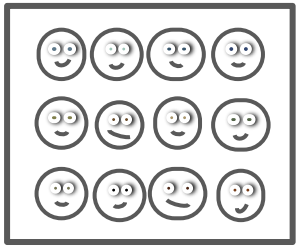
by-intercept random effect

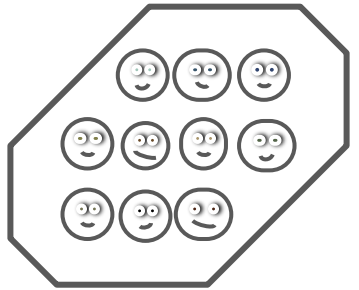


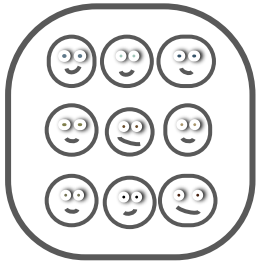
random effect



and/or

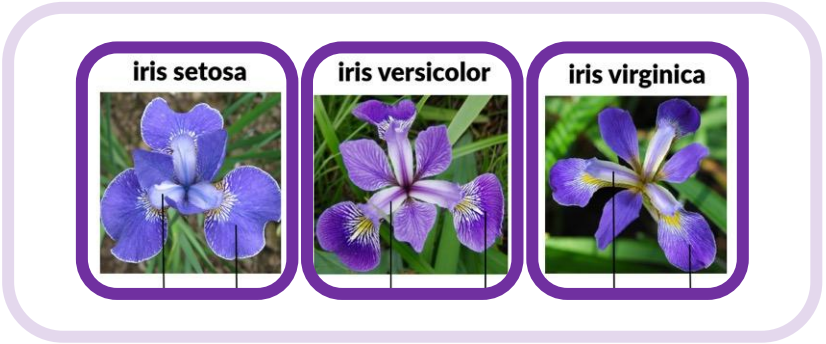






by-individual id
subject…etc

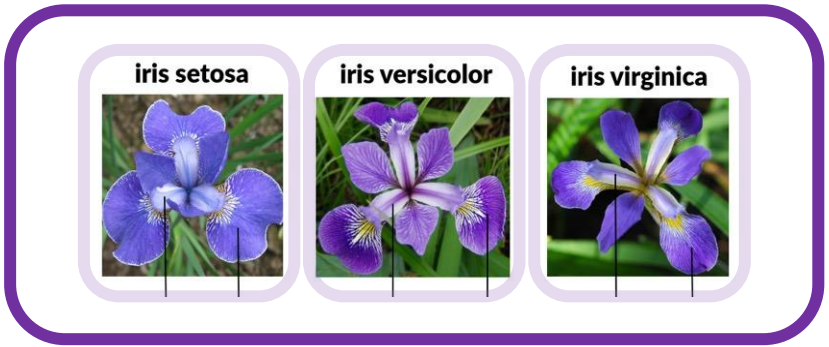
by-group site
school…etc



Species as

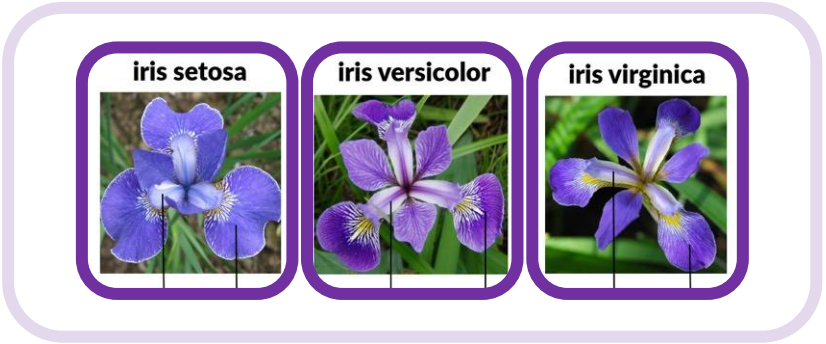
Fixed

Effects

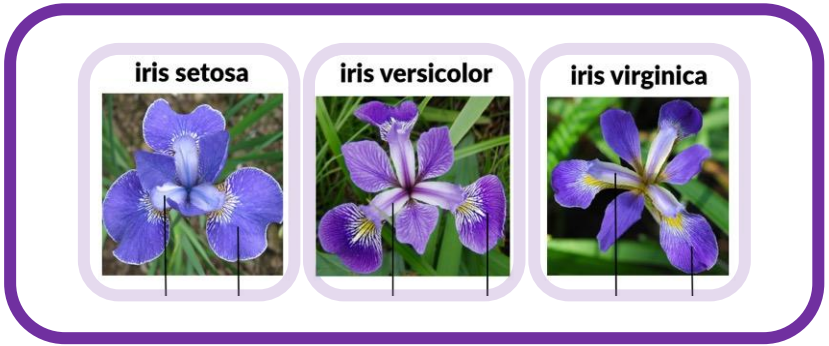


Species as

Random Effects



$y \sim x + \text{Species}$



$y \sim x + (1 \mid \text{Species})$

Random effects are variables based on grouping definition:

- Plant and plant fertilizer across areas
- Exam scores and SES across schools
- RT and condition across participants

They are mostly for repeated measures:

- Repeatedly measure plant growth within each area
- Repeatedly measure exam scores within each school
- Repeatedly measure RT within each participant

Modelling

Research Question: What kind of cues precede conversation islands?

lmer(outcome ~ predictors +

(1 | random effect), data = data)

	Cue	cues	id	age_grp	Freq
1	GesturalC…	Infant gesture to…	2	0	1
2	GesturalC…	Mother gesture t…	2	0	1
3	GesturalC…	Infant gesture to…	3	0	0
4	GesturalC…	Mother gesture t…	3	0	0
5	GesturalC…	Infant gesture to…	4	0	0
6	GesturalC…	Mother gesture t…	4	0	0
7	GesturalC…	Infant gesture to…	5	0	0
8	GesturalC…	Mother gesture t…	5	0	0
9	GesturalC…	Infant gesture to…	6	0	3

```
> library(lme4);
  library(lmerTest)
```

Modelling

Research Question: What kind of cues precede conversation islands?

```
> library(lme4); library(lmerTest)
```

For example,

```
> lmer(Freq ~ Cue + (1 | id), data = dfmod) -> mod_freq
```

Results

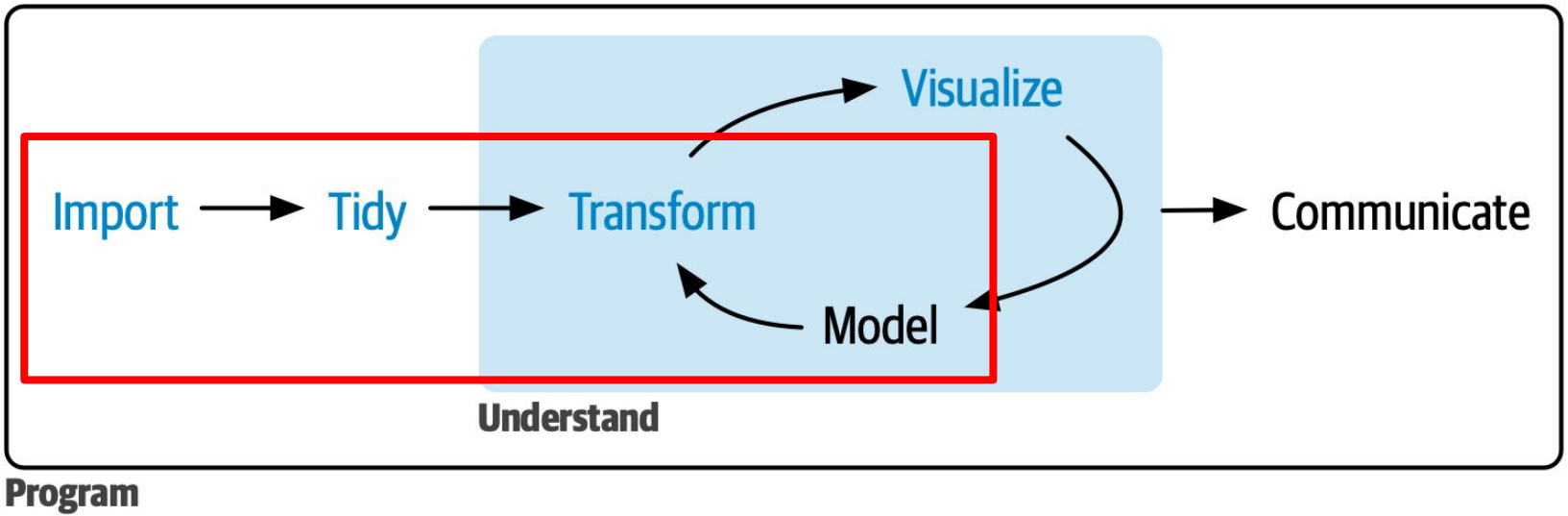
Research Question: What kind of cues precede conversation islands?

```
> summary(mod_freq)
```

```
> colnames(multi[,grepl("Cue", colnames(multi))])
[1] "GesturalCue"       "LookCue"
[3] "PointingCue"       "TactileCue_Human"
[5] "TactileCue_Object"
```

```
> summary(mod_freq)
Linear mixed model fit by REML. t-tests use Satterthwaite's method [
lmerModLmerTest]
Formula: Freq ~ Cue + (1 | id)
   Data: dfmod

REML criterion at convergence: 10144.8

Scaled residuals:
    Min      1Q  Median      3Q     Max
-0.7078 -0.4773 -0.1598  0.0068  9.3804

Random effects:
 Groups   Name        Variance Std.Dev.
 id       (Intercept)  0.4044  0.6359
 Residual             29.0623  5.3909
Number of obs: 1632, groups:  id, 34

Fixed effects:
                    Estimate Std. Error       df t value Pr(>|t|)
(Intercept)           0.2059     0.3929 653.0957   0.524    0.600
CueLookCue            2.3255     0.4466 1594.0000   5.207 2.17e-07 ***
CuePointingCue       -0.1716     0.5338 1594.0000  -0.321    0.748
CueTactileCue_Human   0.5490     0.4873 1594.0000   1.127    0.260
CueTactileCue_Object  2.7108     0.4623 1594.0000   5.864 5.48e-09 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Data Analysis Pipeline



Research Question: Does the use of non-verbal cues changes as children age?

```
> dfmod %>% View()
```

	Cue	cues	id	age_grp	Freq
1	GesturalC…	Infant gesture to…	2	0	1
2	GesturalC…	Mother gesture t…	2	0	1
3	GesturalC…	Infant gesture to…	3	0	0
4	GesturalC…	Mother gesture t…	3	0	0
5	GesturalC…	Infant gesture to…	4	0	0
6	GesturalC…	Mother gesture t…	4	0	0
7	GesturalC…	Infant gesture to…	5	0	0
8	GesturalC…	Mother gesture t…	5	0	0
9	GesturalC…	Infant gesture to…	6	0	3

What do we need?

Research Question: Does the use of non-verbal cues changes as children age?

- Detect presence of cues
- Interaction levels: might differ across age & gender,
- Random effect: differences accounted at the conversational block and subject level

How can we do it?

Research Question: Does the use of non-verbal cues changes as children age?

```
> multi %>%
    select("id", "age_grp", "age", "gender", "cbid",
           grep("Cue", colnames(multi), value = TRUE)) %>%
    group_by(id, age_grp, age, gender, cbid) %>%
    summarise(cue_presence = ifelse(any(!is.na(across(everything()))), 1,
0))
```

How can we do it?

Research Question: Does the use of non-verbal cues changes as children age?

```
> multi %>%
    select("id", "age_grp", "age", "gender", "cbid",
           grep("Cue", colnames(multi), value = TRUE))
```

	id	age_grp	age	gender	cbid	GesturalCue	LookCue	PointingCue	TactileCue_Human	TactileCue_Object
1	2	0	9.460153	M	1	*NA*	Infant and mother look object/action/face/body part(s)	*NA*	Mother touch infant	*NA*
2	2	0	9.460153	M	2	*NA*	Mother look object/action/face/body part(s) (initiator...	*NA*	*NA*	Mother touch/hold object
3	2	0	9.460153	M	3	Mother gesture to object/action	Mother following infant look object/action/face/body...	*NA*	*NA*	*NA*
4	2	0	9.460153	M	4	*NA*	*NA*	*NA*	*NA*	*NA*
5	2	0	9.460153	M	5	*NA*	Infant and mother look object/action/face/body part(s)	*NA*	*NA*	Infant touch/hold object
6	2	0	9.460153	M	6	*NA*	*NA*	*NA*	*NA*	*NA*
7	2	0	9.460153	M	7	*NA*	Infant and mother look object/action/face/body part(s)	*NA*	*NA*	Infant touch/hold object
8	2	0	9.460153	M	8	*NA*	Infant and mother look object/action/face/body part(s)	*NA*	*NA*	Infant touch/hold object
9	2	0	9.460153	M	9	*NA*	Infant and mother look object/action/face/body part(s)	*NA*	*NA*	Mother touch/hold object
10	2	0	9.460153	M	10	*NA*	*NA*	*NA*	*NA*	*NA*
11	2	0	9.460153	M	11	*NA*	*NA*	*NA*	*NA*	*NA*
12	2	0	9.460153	M	12	*NA*	*NA*	*NA*	*NA*	*NA*
13	2	0	9.460153	M	13	*NA*	*NA*	*NA*	*NA*	*NA*

How can we do it?

Research Question: Does the use of non-verbal cues changes as children age?

```
> multi %>%
    select("id", "age_grp", "age", "gender", "cbid",
           grep("Cue", colnames(multi), value = TRUE)) %>%
    group_by(id, age_grp, age, gender, cbid)
```

	id	age_grp	age	gender	cbid	GesturalCue	LookCue	PointingCue	TactileCue_Human	TactileCue_Object
1	2	0	9.460153	M	1	*NA*	Infant and mother look object/action/face/body part(s)	*NA*	Mother touch infant	*NA*
2	2	0	9.460153	M	2	*NA*	Mother look object/action/face/body part(s) (initiator...	*NA*	*NA*	Mother touch/hold object
3	2	0	9.460153	M	3	Mother gesture to object/action	Mother following infant look object/action/face/body...	*NA*	*NA*	*NA*
4	2	0	9.460153	M	4	*NA*	*NA*	*NA*	*NA*	*NA*
5	2	0	9.460153	M	5	*NA*	Infant and mother look object/action/face/body part(s)	*NA*	*NA*	Infant touch/hold object
6	2	0	9.460153	M	6	*NA*	*NA*	*NA*	*NA*	*NA*
7	2	0	9.460153	M	7	*NA*	Infant and mother look object/action/face/body part(s)	*NA*	*NA*	Infant touch/hold object
8	2	0	9.460153	M	8	*NA*	Infant and mother look object/action/face/body part(s)	*NA*	*NA*	Infant touch/hold object
9	2	0	9.460153	M	9	*NA*	Infant and mother look object/action/face/body part(s)	*NA*	*NA*	Mother touch/hold object
10	2	0	9.460153	M	10	*NA*	*NA*	*NA*	*NA*	*NA*
11	2	0	9.460153	M	11	*NA*	*NA*	*NA*	*NA*	*NA*
12	2	0	9.460153	M	12	*NA*	*NA*	*NA*	*NA*	*NA*
13	2	0	9.460153	M	13	*NA*	*NA*	*NA*	*NA*	*NA*

How can we do it?

Research Question: Does the use of non-verbal cues changes as children age?

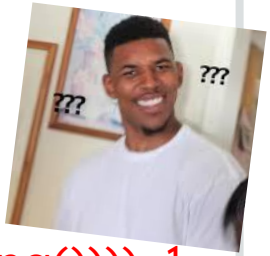
```
> multi %>%
    select("id", "age_grp", "age", "gender", "cbid",
           grep("Cue", colnames(multi), value = TRUE)) %>%
    group_by(id, age_grp, age, gender, cbid)
```

- Interaction levels: might differ across age & gender,
- Random effect: differences accounted at the conversational block and subject level

How can we do it?

Research Question: Does the use of non-verbal cues changes as children age?

```
> multi %>%
    select("id", "age_grp", "age", "gender", "cbid",
           grep("Cue", colnames(multi), value = TRUE)) %>%
    group_by(id, age_grp, age, gender, cbid) %>%
    summarise(cue_presence = ifelse(any(!is.na(across(everything()))), 1,
0))
```



	id	age_grp	age	gender	cbid	cue_presence
1	2	0	9.460153	M	1	1
2	2	0	9.460153	M	2	1

To summarise data

```
> multi %>% ... %>% summarise() %>% summarised_data
```

Data frame %>% **summarise()** -> Summarised data

To summarise data

```
> multi %>% ... %>% summarise(cue_presence = ) %>% summarised_data
```

Data frame %>% **summarise()** -> Summarised data

cue_presence =

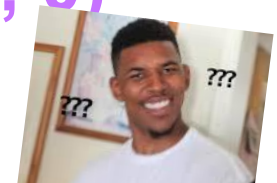
To summarise data

```
> multi %>% ... %>% summarise(cue_presence =
  ifelse(any(!is.na(across(everything()))), 1, 0)) %>% summarised_data
```

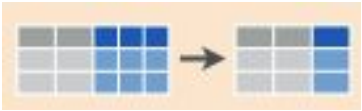
Data frame %>% **summarise()** -> Summarised data

cue_presence =

ifelse(any(!is.na(across(everything()))), 1, 0)



Work out the magic



```
ifelse(
    any(
        !is.na(
            across(everything())
        )
    ),
    1,
    0
)
```

if there are
    any
        non-missing values
            across all values,

returns 1,
else returns 0

Data for Analysis

Research Question: Does the use of non-verbal cues changes as children age?

```
> multi %>%
    select("id", "age_grp", "age", "gender", "cbid",
           grep("Cue", colnames(multi), value = TRUE)) %>%
    group_by(id, age_grp, age, gender, cbid) %>%
    summarise(cue_presence = ifelse(any(!is.na(across(everything()))), 1,
0)) -> dfmod_prop
```

	id	age_grp	age	gender	cbid	cue_presence
1	2	0	9.460153	M	1	1
2	2	0	9.460153	M	2	1

Data for Analysis

Research Question: Does the use of non-verbal cues changes as children age?

```
> multi %>%
    select("id", "age_grp", "age", "gender", "cbid",
           grep("Cue", colnames(multi), value = TRUE)) %>%
    group_by(id, age_grp, age, gender, cbid) %>%
    summarise(cue_presence = ifelse(any(!is.na(across(everything()))), 1,
0)) -> dfmod_prop
```

	id	age_grp	age	gender	cbid	cue_presence
1	2	0	9.460153	M	1	1
2	2	0	9.460153	M	2	1

Modelling

Research Question: Does the use of non-verbal cues changes as children age?

- Modelling the effect of **age (and gender)** on the **presence of non-verbal cues**
  - Account for the random effect of **id and cbid**

	id	age_grp	age	gender	cbid	cue_presence
1	2	0	9.460153	M	1	1
2	2	0	9.460153	M	2	1

Modelling

Research Question: Does the use of non-verbal cues changes as children age?

```
> glmer(cue_presence ~ age_grp * gender + (1 + cbid | id),
        family = binomial,
        data = dfmod_prop) -> mod_prop
```

	id	age_grp	age	gender	cbid	cue_presence
1	2	0	9.460153	M	1	1
2	2	0	9.460153	M	2	1

Modelling

Research Question: Does the use of non-verbal cues changes as children age?

```
> glmer(cue_presence ~
        age_grp * gender +
        (1 + cbid | id),
```

- **outcome** : cue_presence
- **predictors**: age_grp, gender and interaction term
- **random effects**: conversation block index by-subject intercept

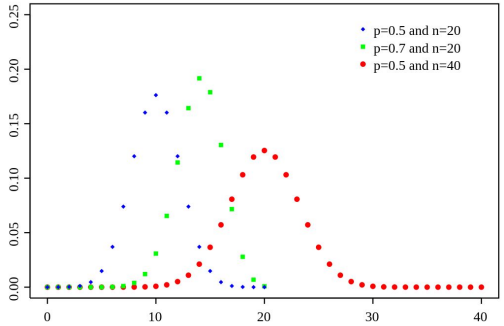
	id	age_grp	age	gender	cbid	cue_presence
1	2	0	9.460153	M	1	1
2	2	0	9.460153	M	2	1

Modelling

Research Question: Does the use of non-verbal cues changes as children age?

```
> glmer(cue_presence ~
        age_grp * gender +
        (1 + cbid | id),
        family = binomial,
```

- **Data distribution**: binomial (0, 1)



	id	age_grp	age	gender	cbid	cue_presence
1	2	0	9.460153	M	1	1
2	2	0	9.460153	M	2	1

Modelling

Research Question: Does the use of non-verbal cues changes as children age?

```
> glmer(cue_presence ~
        age_grp * gender +
        (1 + cbid | id),
        family = binomial,
        data = dfmod_prop) -> mod_prop
```

	id	age_grp	age	gender	cbid	cue_presence
1	2	0	9.460153	M	1	1
2	2	0	9.460153	M	2	1

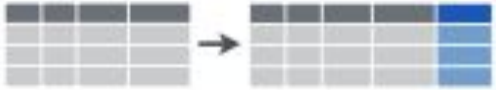
Results

Research Question: Does the use of non-verbal cues changes as children age?

```
> summary(mod_prop)
```

```
Random effects:
 Groups Name        Variance Std.Dev. Corr
 id     (Intercept) 0.868108 0.9317
        cbid        0.000697 0.0264   -0.92
Number of obs: 2039, groups:  id, 34

Fixed effects:
               Estimate Std. Error z value Pr(>|z|)
(Intercept)     0.98952    0.23284   4.250 2.14e-05 ***
age_grp        -0.49550    0.18852  -2.628  0.00858 **
genderM        -0.03819    0.27899  -0.137  0.89111
age_grp:genderM 0.23086    0.23041   1.002  0.31636
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Make a factor variable



```
> dfmod_prop <- dfmod_prop %>%
                mutate(age_grp = factor(age_grp,
                levels = c(0, 1, 2)))
```

Data frame %>% **mutate()** -> transformed data

age_grp =

factor(age_grp, levels = c(0, 1, 2))

Work out the magic

```
factor(
    age_grp,
        levels = c(0, 1, 2)
)
```

create a factor
with column age_grp
with ordered levels: 0, 1, 2

age_grp column should contain these values

Refit the model

```
> glmer(cue_presence ~ age_grp * gender + (1 + cbid | id),  family =
  binomial, data = dfmod_prop) -> mod_prop
```

```
> summary(mod_prop)
```

```
Random effects:
 Groups Name        Variance  Std.Dev. Corr
 id     (Intercept) 0.8643747 0.9297
        cbid        0.0006972 0.0264   -0.92
Number of obs: 2039, groups:  id, 34

Fixed effects:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)       0.91673    0.25206   3.637 0.000276 ***
age_grp1         -0.30447    0.31846  -0.956 0.339025
age_grp2         -1.03002    0.37912  -2.717 0.006590 **
genderM           0.03527    0.30358   0.116 0.907519
age_grp1:genderM  0.04103    0.44052   0.093 0.925784
age_grp2:genderM  0.50109    0.46124   1.086 0.277303
```

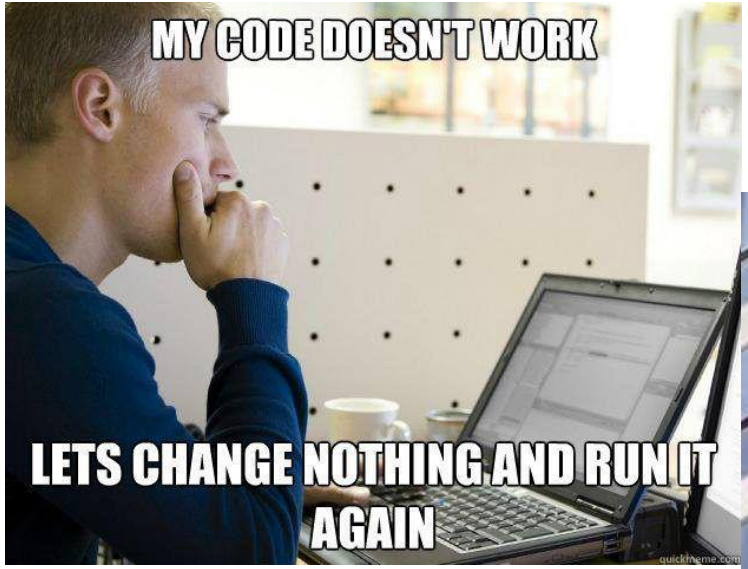
Refit the model

```
> glmer(cue_presence ~ age_grp * gender + (1 + cbid | id),  family =
  binomial, data = dfmod_prop) -> mod_prop
```

```
> summary(mod_prop)
```

```
Random effects:
 Groups Name        Variance  Std.Dev. Corr
 id     (Intercept) 0.8643747 0.9297
        cbid        0.0006972 0.0264   -0.92
Number of obs: 2039, groups:  id, 34

Fixed effects:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)        0.       .        .37 0.000276 ***
age_grp1          -0.       .        .56 0.339025
age_grp2          -1.03002  0.37912  -2.717 0.006590 **
genderM            0.03527  0.30358   0.116 0.907519
age_grp1:genderM   0.04103  0.44052   0.093 0.925784
age_grp2:genderM   0.50109  0.46124   1.086 0.277303
```

age_grp0:genderF


MY CODE DOESN'T WORK
LETS CHANGE NOTHING AND RUN IT AGAIN


It doesn't work..... why?
It works..... why?

RStudio + GitHub Copilot & Web

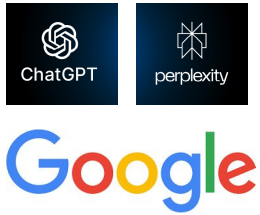
https://github.com/rstudio/cheatsheets/blob/main/base-r.pdf

Apply to GitHub Education as a student:

https://docs.github.com/en/education/explore-the-benefits-of-teaching-and-learning-with-github-education/github-education-for-students/apply-to-github-education-as-a-student

Setup GitHub Copilot in RStudio:

https://docs.posit.co/ide/user/ide/guide/tools/copilot.html#setup



Part II

Objectives

1. Investigate interactions between categorical variables using *emmeans()*
2. Investigate interactions involving continuous variables using *modelbased* package
3. Visualising the effects using the *ggeffect* and *modelbased* package.

```
> library(emmeans); library(modelbased); library(ggeffects)
```

Revisiting Research Question: Does the use of non-verbal cues changes as children age?

- Does **initiator** of conversation have an effect on the presence of **non-verbal cues**,
- along with **age** and **gender**?

	id	pid	gender	age	age_grp	cbid	initiator	t1.x
1	2	P01	M	9.460153	0	1	mother	8.
2	2	P01	M	9.460153	0	2	mother	35.
3	2	P01	M	9.460153	0	3	mother	41.
4	2	P01	M	9.460153	0	4	mother	71.
5	2	P01	M	9.460153	0	5	mother	82.
6	2	P01	M	9.460153	0	6	mother	112.
7	2	P01	M	9.460153	0	7	mother	123.
8	2	P01	M	9.460153	0	8	mother	128.
9	2	P01	M	9.460153	0	9	mother	157.
10	2	P01	M	9.460153	0	10	mother	164.

Reshape the data

Research Question: Does the use of non-verbal cues changes as children age?

```
> multi %>%
    select("id", "age_grp", "age", "gender", "cbid", "initiator",
           grep("Cue", colnames(multi), value = TRUE)) %>%
    group_by(id, age_grp, age, gender, cbid, initiator) %>%
    summarise(cue_presence = ifelse(any(!is.na(across(everything()))), 1,
0)) -> dfmod_prop
```

	id	age_grp	age	gender	cbid	initiator	cue_presence
1	2	0	9.460153	M	1	mother	1
2	2	0	9.460153	M	2	mother	1
3	2	0	9.460153	M	3	mother	1

Refit the model

Research Question: Does the use of non-verbal cues changes as children age?

```
> glmer(cue_presence ~ age_grp * gender * initiator + (1 + cbid | id),
  family = binomial, data = dfmod_prop) -> mod_prop
```

```
> glmer(cue_presence ~ age_grp * gender * initiator + (1 + cbid | i
d),  family = binomial, data = dfmod_prop) -> mod_prop
Warning message:
In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkCon
v,  :
  Model failed to converge with max|grad| = 0.00701575 (tol = 0.00
2, component 1)
```

Convergence Issue (more common than you thought!)

1. Check appropriateness of model specification give data structure
2. Adjust convergence control parameters, for example:
   control=glmerControl(optCtrl=list(maxfun=2e5))
3. Try different optimizers, for example:
   control=glmerControl(optimizer="nloptwrap",
   optCtrl=list(algorithm="NLOPT_LN_BOBYQA"))
4. Simplify random effects structure, for example: (1 | id) + (0 + group | id)
5. Re-scale variables
6. Try alternate packages with different estimators (glmmTMB, brms, MCMCglmm)
7. Reconsider model and data

Refit the model

Research Question: Does the use of non-verbal cues changes as children age?

```
> glmmTMB::glmmTMB(cue_presence ~ age_grp * gender * initiator +
  (1 + cbid | id),  family = binomial, data = dfmod_prop) ->
  mod_prop_alt
```

```
> glmmTMB::glmmTMB(cue_presence ~ age_grp * gender * initiator + (1
+ cbid | id),  family = binomial, data = dfmod_prop) -> mod_prop
>
```

yay!

Compare the converged and non-converged models

glmmTMB::glmmTMB()

```
> car::Anova(mod_prop_alt, type = 3)
Analysis of Deviance Table (Type III Wald chisquare tes

Response: cue_presence
                         Chisq Df Pr(>Chisq)
(Intercept)              7.2597  1   0.007052 **
age_grp                  7.1502  2   0.028013 *
gender                   1.4806  1   0.223688
initiator                0.9485  1   0.330095
age_grp:gender           5.4573  2   0.065308 .
age_grp:initiator        1.1152  2   0.572570
gender:initiator         2.2692  1   0.131969
age_grp:gender:initiator 3.8771  2   0.143916
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.
```

glmer()

```
> car::Anova(mod_prop, type = 3)
Analysis of Deviance Table (Type III Wald chisquare tests)

Response: cue_presence
                         Chisq Df Pr(>Chisq)
(Intercept)              7.3127  1   0.006847 **
age_grp                  7.1703  2   0.027733 *
gender                   1.4914  1   0.221992
initiator                0.9546  1   0.328563
age_grp:gender           5.4769  2   0.064669 .
age_grp:initiator        1.1196  2   0.571324
gender:initiator         2.2849  1   0.130638
age_grp:gender:initiator 3.8938  2   0.142719
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Age group X gender

glmmTMB::glmmTMB()

```
> car::Anova(mod_prop_alt, type = 3)
Analysis of Deviance Table (Type III Wald chisquare tests)

Response: cue_presence
                         Chisq Df Pr(>Chisq)
(Intercept)             7.2597  1   0.007052 **
age_grp                 7.1502  2   0.028013 *
gender                  1.4806  1   0.223688
initiator               0.9485  1   0.330095
age_grp:gender          5.4573  2   0.065308 .
age_grp:initiator       1.1152  2   0.572570
gender:initiator        2.2692  1   0.131969
age_grp:gender:initiator 3.8771  2   0.143916
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

X

- Pairwise comparisons post-hoc analysis
- When effect is significant (main effect or interaction)
- Prioritise most complex effect (interactions) to avoid misinterpretations

Interactions

	Report stats	Visualise
Categorical x Categorical	emmeans::	ggeffects::
Continuous x Categorical	modelbased::	ggeffects::
Continuous x Continuous	modelbased:: + plot()	ggeffects:: / contour

“Band of influence”: estimate_means() + ggplot() + geom_contour_filled()

Why study interactions?

1. Handling multiple comparisons
2. Determine which group differences are significant

Why use emmeans?

Estimated Marginal Means (emmeans)

1. Calculate adjusted means for factor combinations
2. Estimate effects of controlled conditions using a model-based approach

```
> emmeans(the model, pairwise ~ variable A | variable B)
```

Categorical x Categorical: Report stats

```
> emmeans(mod_prop_alt, pairwise ~ age_grp | gender)
```

```
gender = F:
 contrast              estimate    SE  df z.ratio p.value
 age_grp0 - age_grp1      0.308 0.399 Inf   0.773  0.7197
 age_grp0 - age_grp2      1.214 0.433 Inf   2.805  0.0139
 age_grp1 - age_grp2      0.906 0.396 Inf   2.289  0.0573

gender = M:
 contrast              estimate    SE  df z.ratio p.value
 age_grp0 - age_grp1      0.143 0.351 Inf   0.407  0.9129
 age_grp0 - age_grp2      0.304 0.300 Inf   1.014  0.5678
 age_grp1 - age_grp2      0.162 0.310 Inf   0.521  0.8612
```

Categorical x Categorical: Report stats

```
> emmeans(mod_prop_alt, pairwise ~ gender | age_grp)
```

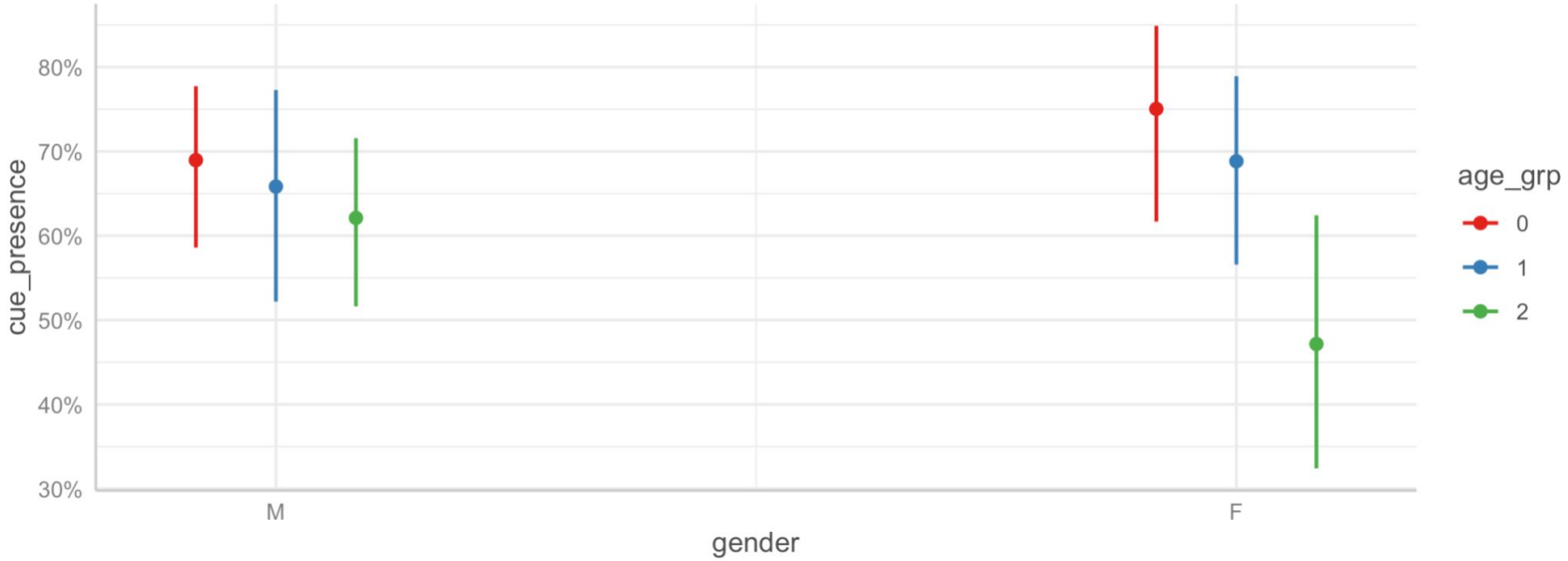
```
age_grp = 0:
 contrast estimate    SE  df z.ratio p.value
 F - M       0.302 0.384 Inf   0.787  0.4312

age_grp = 1:
 contrast estimate    SE  df z.ratio p.value
 F - M       0.137 0.357 Inf   0.383  0.7019

age_grp = 2:
 contrast estimate    SE  df z.ratio p.value
 F - M      -0.608 0.357 Inf  -1.702  0.0887
```

Categorical x Categorical: Visualise

```
> ggeffects::ggemmeans(model, terms = c("variableA", "variableB"))
> ggemmeans(model, terms = c("variableA", "variableB"))
```

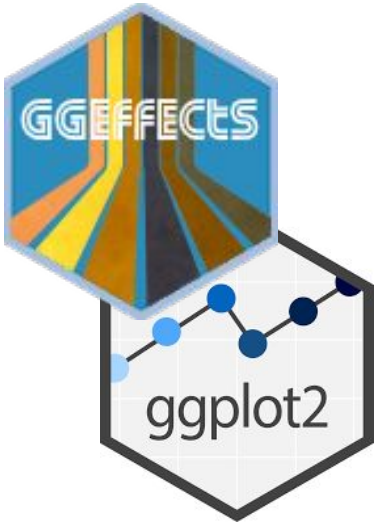
Categorical x Categorical: Visualise

```
> ggeffects::ggemmeans(mod_prop_alt, terms = c("gender",
  "age_grp")) %>% plot()
```



ggeffects (& ggplot2, etc)

gg = grammar of graphics



Define overall visuals

Set limits for graph

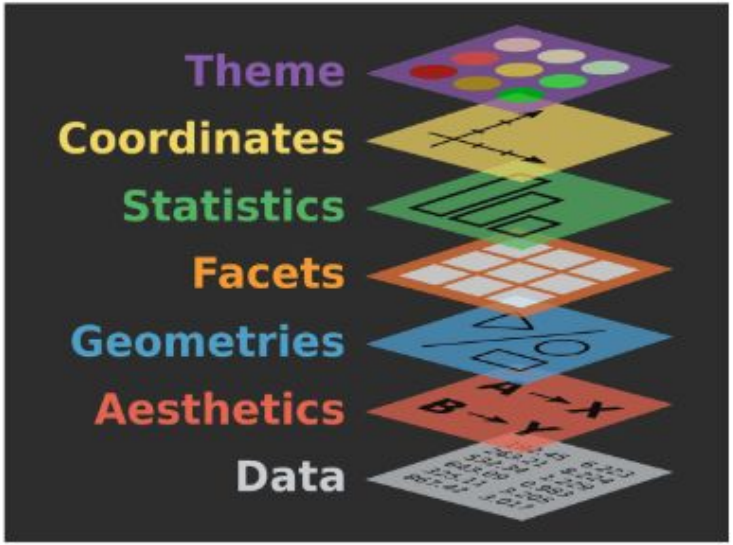
Summarise data

Plot data into facets

3. Declare what graphical geometries to use

2. Map variables to aesthetics

1. Provide data



from [The Grammar of Graphics]

Model 2: Refit model with continuous variable

```
> glmer(cue_presence ~ age * gender * initiator + (1 + cbid | id),
    family = binomial, data = dfmod_prop) -> mod_prop_age
```

```
> glmer(cue_presence ~ age * gender * initiator + (1 + cbid | id),  f
amily = binomial, data = dfmod_prop) -> mod_prop_age
Warning message:
In checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv,
:
  Model failed to converge with max|grad| = 0.0183897 (tol = 0.002, c
omponent 1)
```

Refit model with continuous variable

```
> glmer(cue_presence ~ scale(age) * gender * initiator + (1 | id) + (0 + cbid | id),  family = binomial, data = dfmod_prop) -> mod_prop_age_alt
```

```
> glmer(cue_presence ~ scale(age) * gender * initiator + (1 | id) +
(0 + cbid | id),  family = binomial, data = dfmod_prop) -> mod_prop_a
ge
>
```

yay!

Compare the converged and non-converged models

scale(age) … + (1 | id) + (0 + cbid | id)

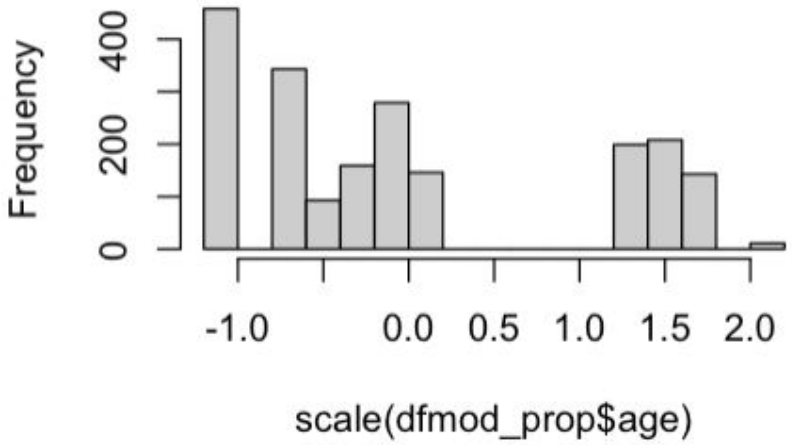
```
> car::Anova(mod_prop_age_alt, type = "III")
Analysis of Deviance Table (Type III Wald chisquare te

Response: cue_presence
                            Chisq Df Pr(>Chisq)
(Intercept)                21.4313  1  3.667e-06 ***
scale(age)                  6.8524  1   0.008852 **
gender                      0.4803  1   0.488294
initiator                   2.9968  1   0.083430 .
scale(age):gender           7.4308  1   0.006412 **
scale(age):initiator        1.7001  1   0.192268
gender:initiator            3.0403  1   0.081222 .
scale(age):gender:initiator 5.6233  1   0.017723 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0
```

age … + (1 + cbid | id)

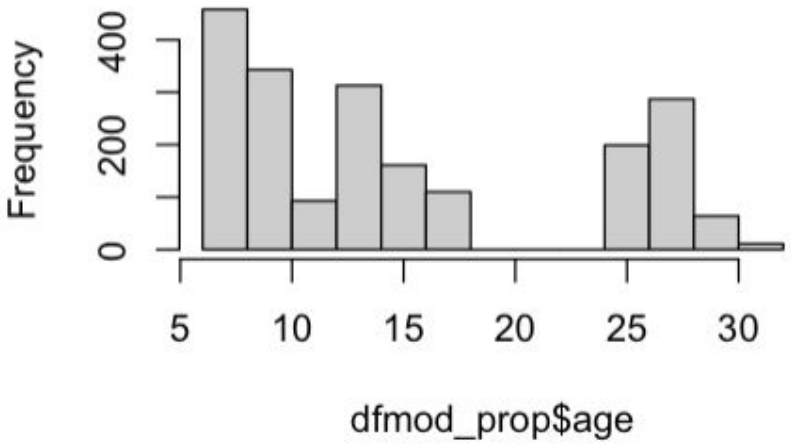
```
> car::Anova(mod_prop_age, type = "III")
Analysis of Deviance Table (Type III Wald chisquare tests)

Response: cue_presence
                          Chisq Df Pr(>Chisq)
(Intercept)             13.7123  1  0.0002131 ***
age                      8.1103  1  0.0044015 **
gender                   5.5190  1  0.0188105 *
initiator                2.7980  1  0.0943800 .
age:gender               6.3967  1  0.0114336 *
age:initiator            1.5558  1  0.2122772
gender:initiator         6.3908  1  0.0114716 *
age:gender:initiator     5.4959  1  0.0190606 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Compare terms of the converged and non-converged

scale(age) ... + (1 | id) +
(0 + cbid | id)



age ... + (1 + cbid | id)



Should you scale?

Not always.

Pros:

1) Can improve convergence
2) Standardizes ranges and magnitude so more manageable
3) Makes variables more comparable to each other

Cons:

1) Can reduce interpretability
2) Requires extra preprocessing steps and computation
3) Can amplify noise (e.g., outliers)

Compare terms of the converged and non-converged

scale(age) ... + **(1 | id) + (0 + cbid | id)**

- assume an **independent** random intercept and slope

➔ (1 | id): **random intercept** for each level of "id"

➔ (0 + cbid | id): **random slope** for "cbid" for each level of "id", **without** a **random intercept**

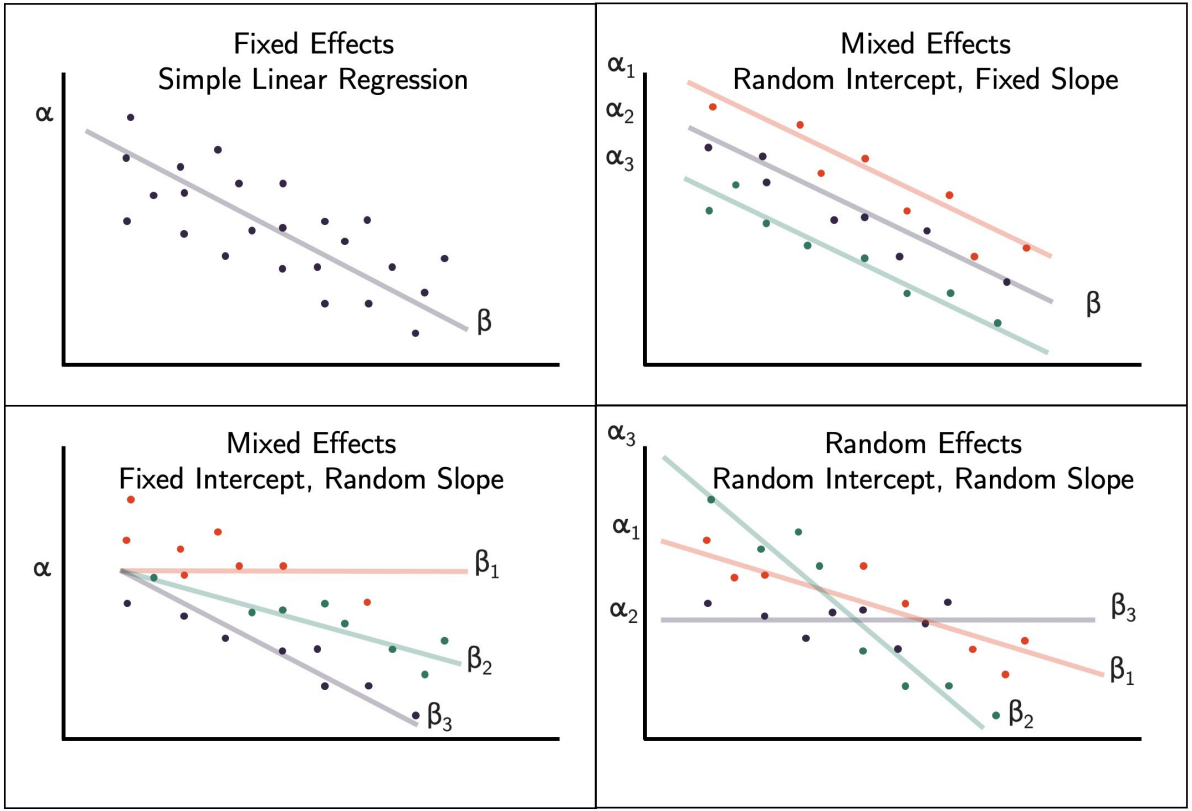
age ... **+ (1 + cbid | id)**

- assume a **correlated** random intercept and slope

➔ (1 + cbid | id): **random intercept and slope** for "cbid", grouped by "id"

➔ allows intercept and effect of "cbid" to vary and correlate across levels of "id"

random intercept and slope

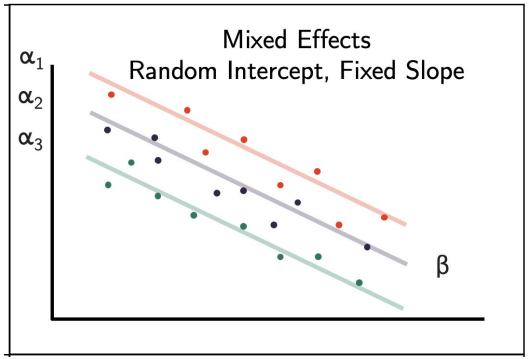


https://bookdown.org/steve_midway/DAR/random-effects.html

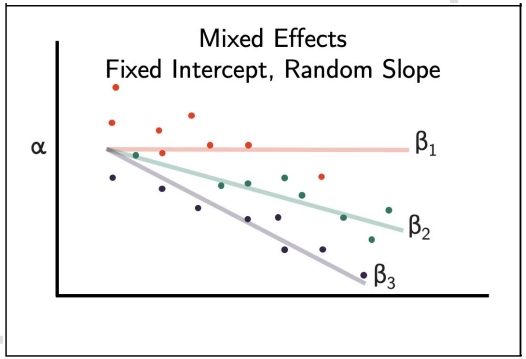
Compare terms of the converged and non-converged

scale(age) ... +

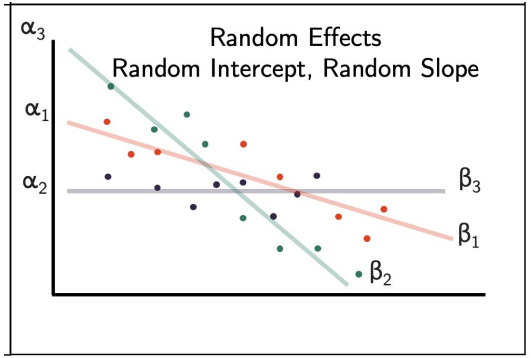
(1 | id) +



(0 + cbid | id)



age ... + **(1 + cbid | id)**



Age (continuous) X gender (categorical)

scale(age) … + (1 | id) +

(0 + cbid | id)

```
> car::Anova(mod_prop_age_alt, type = "III")
Analysis of Deviance Table (Type III Wald chisquare tests)

Response: cue_presence
                            Chisq Df Pr(>Chisq)
(Intercept)                21.4313  1  3.667e-06 ***
scale(age)                  6.8524  1   0.008852 **
gender                      0.4803  1   0.488294
initiator                   2.9968  1   0.083430 .
scale(age):gender           7.4308  1   0.006412 **
scale(age):initiator        1.7001  1   0.192268
gender:initiator            3.0403  1   0.081222 .
scale(age):gender:initiator 5.6233  1   0.017723 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

- Significant interaction between age (continuous) and gender (categorical)

Continuous x Categorical: Report stats

```
> modelbased::estimate_slopes(model, trend = "continuous variable", at = "categorical variable")
```

Continuous x Categorical: Report stats

```
> modelbased::estimate_slopes(mod_prop_age_alt, trend = "age",
  at = "gender")
```

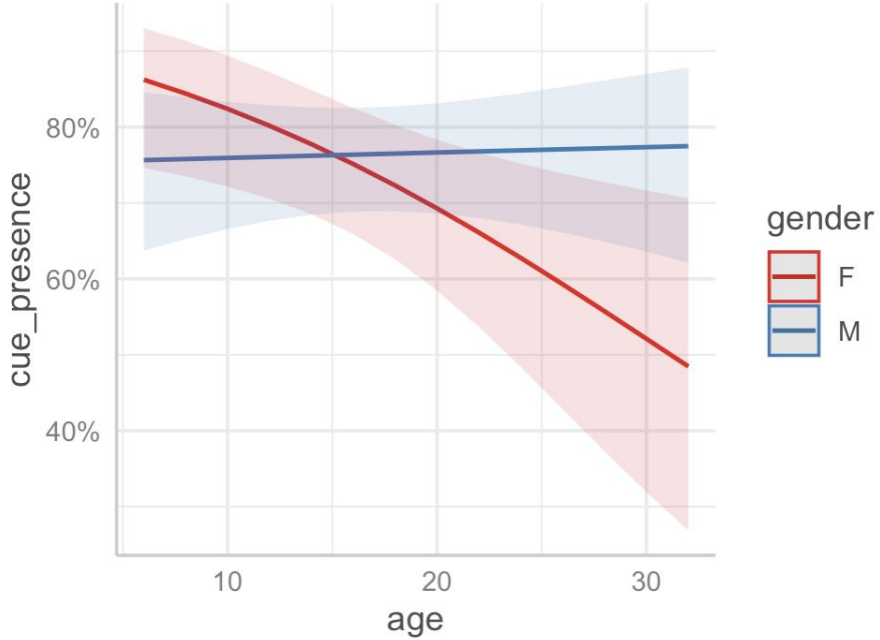
```
> modelbased::estimate_slopes(mod_prop_age_alt, trend = "age", at
= "gender")
NOTE: Results may be misleading due to involvement in interactions
Estimated Marginal Effects

gender | Coefficient |   SE |         95% CI |     z |  df |     p
-------------------------------------------------------------------
M      |    3.95e-03 | 0.02 | [-0.04,  0.04] |  0.19 | Inf | 0.850
F      |       -0.07 | 0.03 | [-0.13, -0.02] | -2.58 | Inf | 0.010
Marginal effects estimated for age>
```

Continuous x Categorical: Visualise

```
> ggeffects::ggemmeans(model, terms = c("continuousVar",
  "categoricalVar"))
```

Continuous x Categorical: Visualise

```
> ggeffects::ggemmeans(mod_prop_age_alt, terms = c("age",
  "gender")) %>% plot()
```



Saving graph as a png

```
> ggsave()
> ggsave("plot.png", width = 7.5, height = 7)
```

- Save last graph, unless specified:

```
> ggeffects::ggemmeans(mod_prop_age_alt, terms = c("age",
  "gender")) %>% plot() -> myplot
> ggsave("plot.png", plot = myplot, width = 7.5, height = 7)
```

Continuous x Categorical: Report stats

```
> modelbased::estimate_slopes(mod_prop_age_alt, trend = "age",
  at = "gender")
```

```
> modelbased::estimate_slopes(mod_prop_age_alt, trend = "age", at
= "gender")
NOTE: Results may be misleading due to involvement in interactions
Estimated Marginal Effects

gender | Coefficient |   SE |         95% CI |     z |  df |     p
-------------------------------------------------------------------
M      |    3.95e-03 | 0.02 | [-0.04,  0.04] |  0.19 | Inf | 0.850
F      |       -0.07 | 0.03 | [-0.13, -0.02] | -2.58 | Inf | 0.010
Marginal effects estimated for age>
```

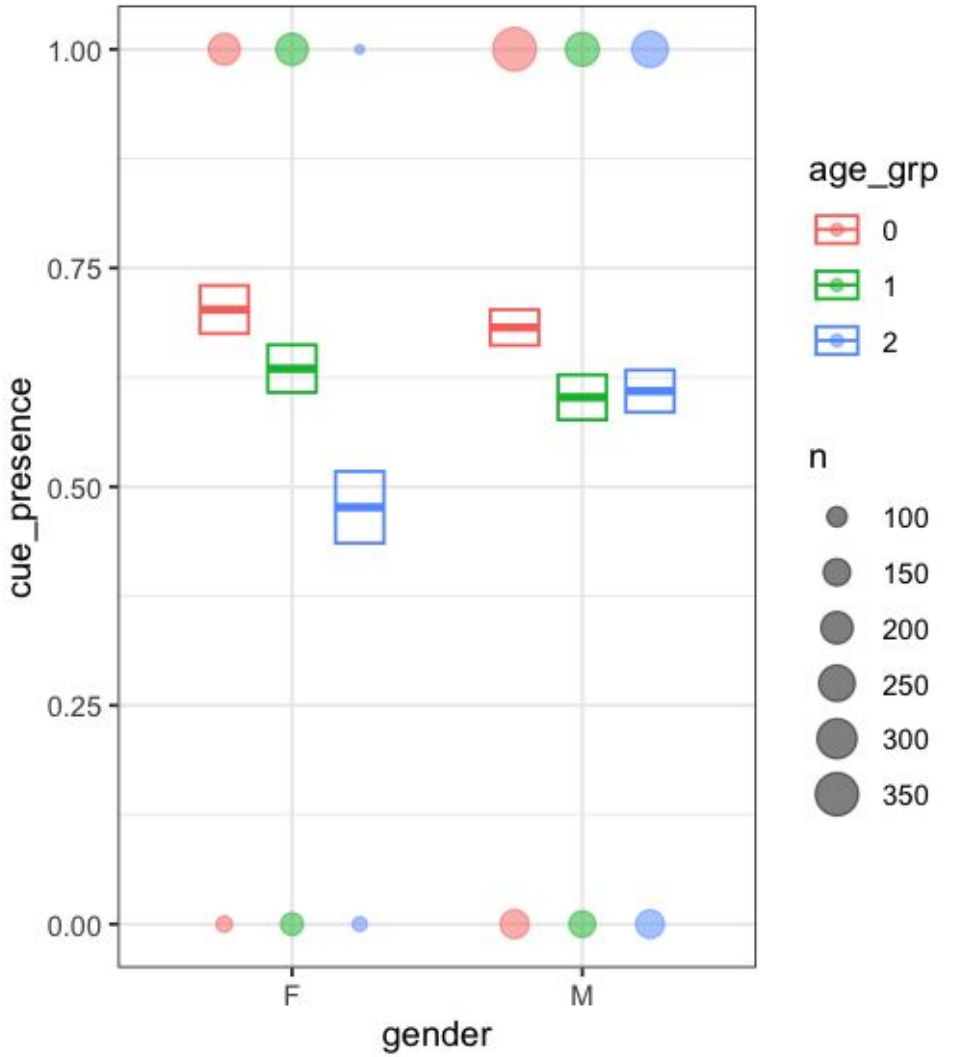
Visualisation with ggplot

```
> ggplot(dfmod_prop,
        aes(y = cue_presence, col = age_grp, x = cues)) +
        stat_summary(geom = "crossbar",
              fun.data = "mean_se", position = position_dodge(width = .7),
        width = .5) +
        coord_flip() +
        theme_bw() +
        geom_count( position = position_dodge(width = .7), alpha = .5)
```

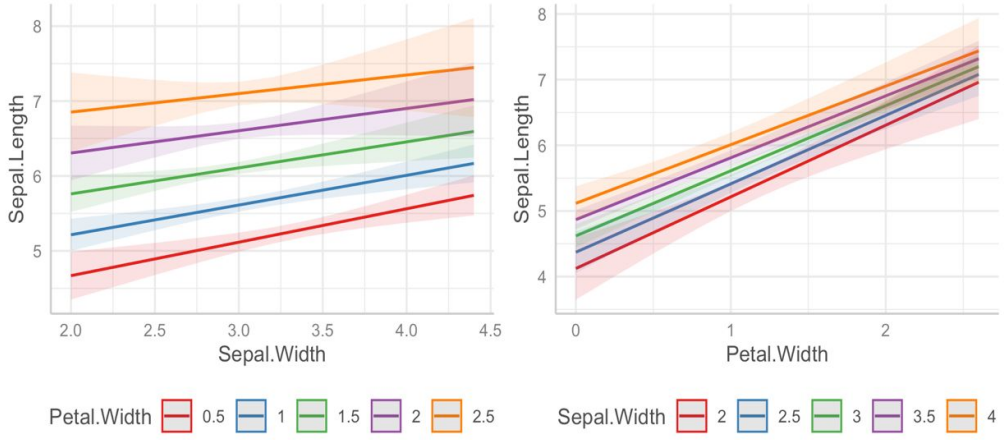
Visualisatio

tion: Does the use of
ges as children age?

```
> ggplot(dfm
      aes(y =
      stat_su
          fur                                    dge(width = .7),
      width =
      coord_
      theme_
      geom_                                      ), alpha = .5)
```

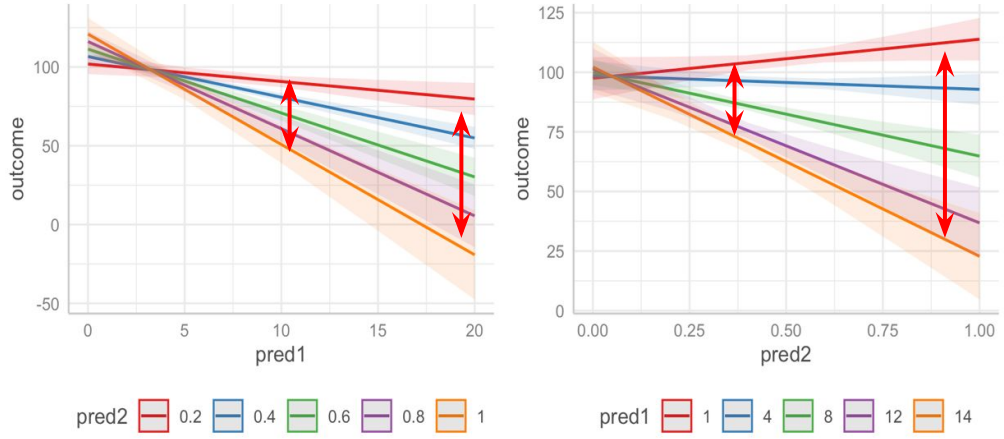


Continuous x Continuous



Model 1: predicting **Sepal.Length**

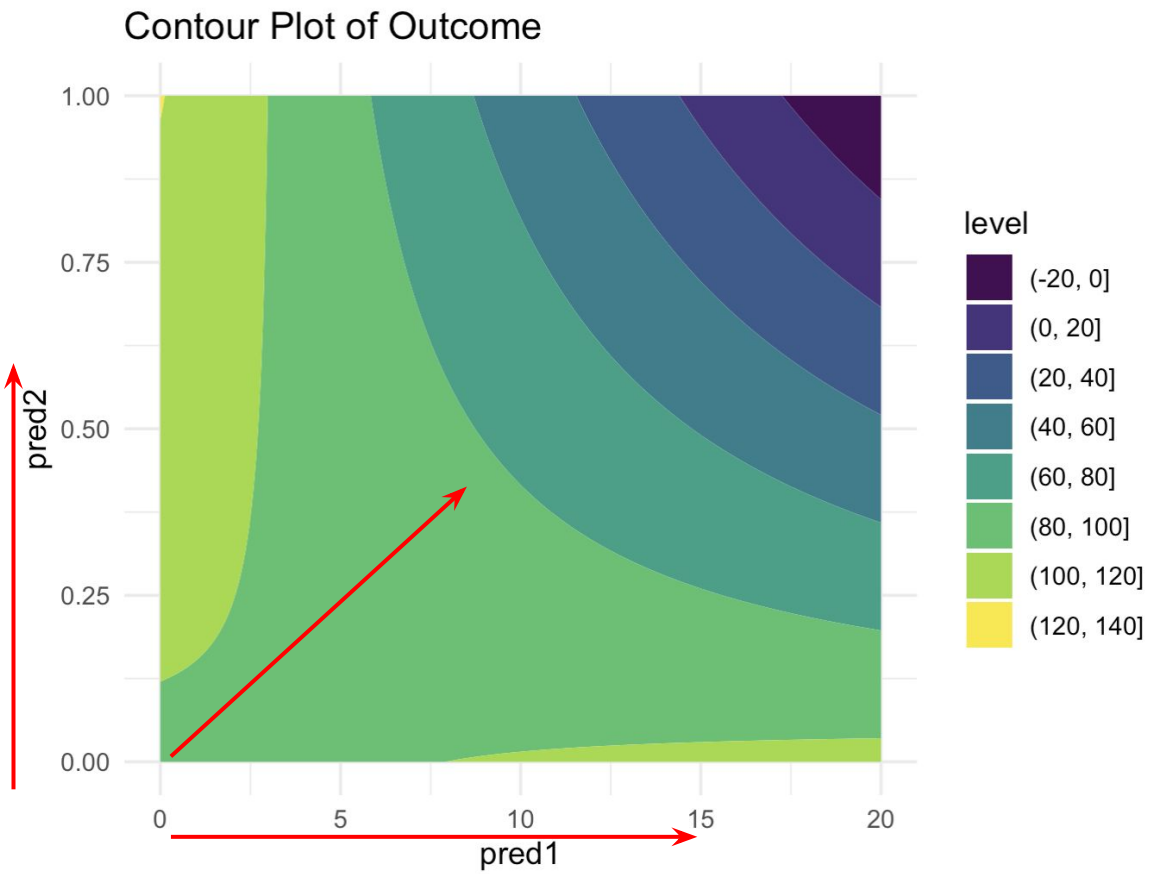
	Estimate	Pr(>|t|)	
(Intercept)	3.13046	2.1e-07	***
Sepal.Width	0.49642	0.00418	**
Petal.Width	1.29057	0.00729	**
Sepal.Width:Petal.Width	-0.09946	0.50038	



Model 2: predicting **outcome**

	Estimate	Pr(>|t|)	
(Intercept)	97.1423	< 2e-16	***
pred1	0.3617	0.5743	
pred2	23.7294	0.0155	*
pred1:pred2	-7.3663	9.6e-07	***

Continuous x Continuous



- 3-way change in values
- Effect of one predictor on the outcome depends on the level of the other predictor

Continuous ~ Continuous x Continuous

is actually a 3d space!

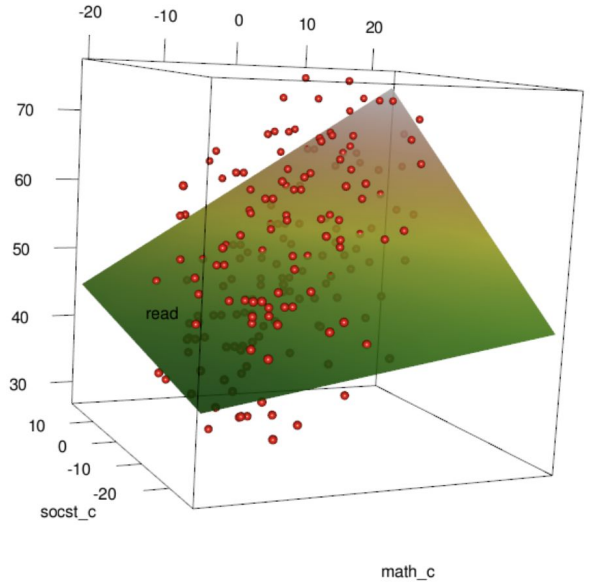


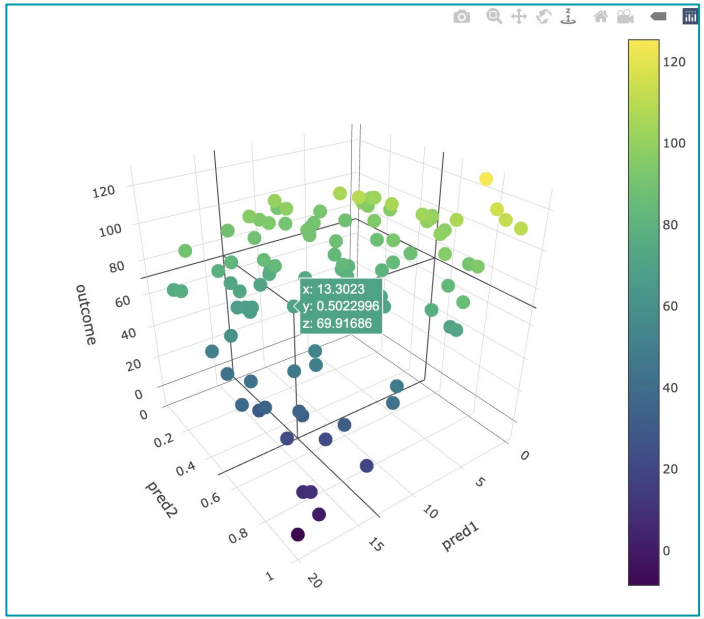
Figure 3: Interaction response surface.

https://web.pdx.edu/~joel8/resources/ConceptualPresentationResources/ContinuousByContinousInteractions_walkthrough_v2.pdf

Continuous ~ Continuous x Continuous

try it yourself

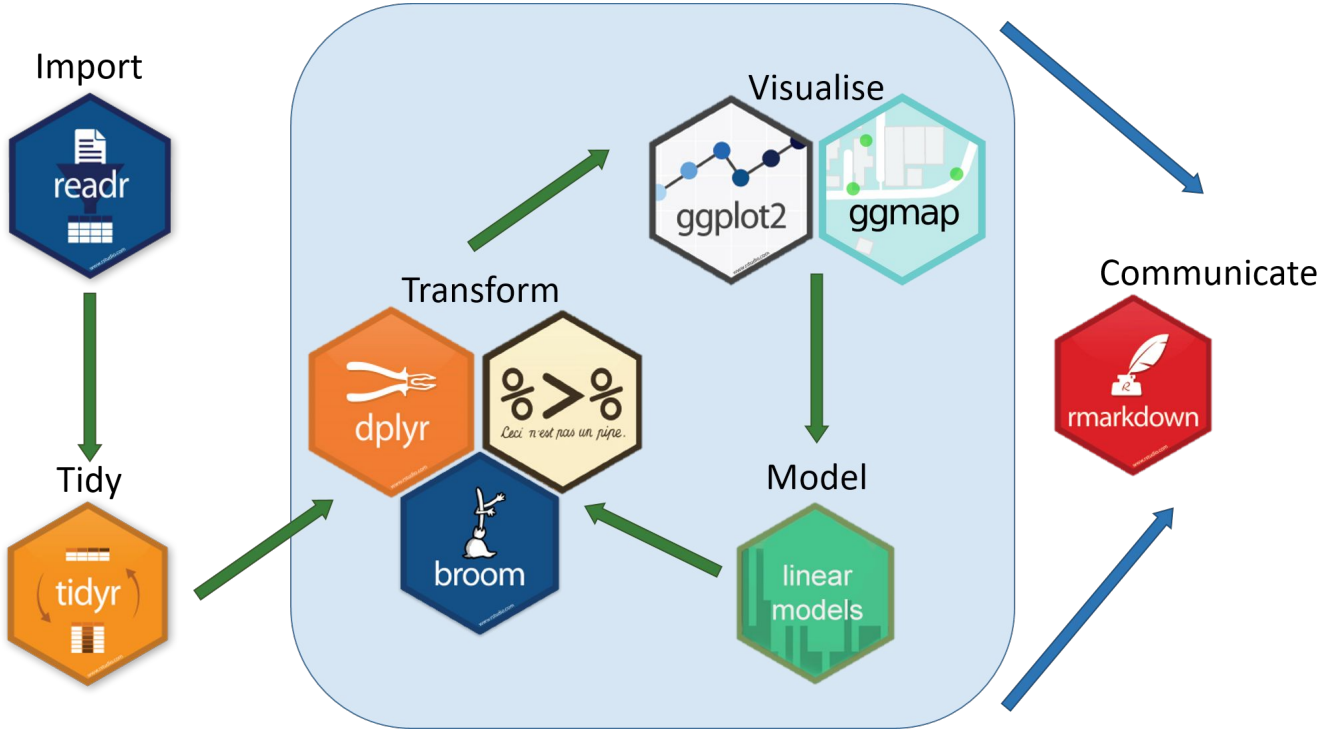
https://junhochai.github.io/figures/interactive_plot.html



Bonus questions

1. How the cue presence differ depending on who initiates the conversation and child's age and gender? (3-way interactions)
2. How does the cue distributions change over age and what are its implications on children's development?

Data Analysis Pipeline



Thank you!

junhoc94@gmail.com

about me

much more edgy and æsthetic - according to my dad

蔡骏豪
(cài jùn háo)
Chai Jun Ho
.
other variations:
Chai Chun Hao
Chai Chung Hou
Chai Chong How
.
.

Hello

Apa khabar?

你好!

नमस्ते

வணக்கம்

我是

客家人

kè jiā rén

客：guest

家：family

人：people