How to conduct and interpret correlation analysis

Dr. Margriet Groen

Overview

- How to build a scatterplot
- How to run a correlation analysis and interpret the result

Step 1: Setup - load libraries

```
1 library(here)
2 library(tidyverse)
3 library(easystats)
```

1. The `here` package helps with finding your files
2. The `tidyverse` is a set of packages that handles almost everything you need to do with data
3. `easystats` is a set of packages that handles many statistical analyses; including the function that we need to run correlation analysis

Step 2: Reading in the data

```
1 data_mw_subscales <- read_csv(here("PSYC4141/files/week3", "data_bmw
```

1. `read_csv` reads in data; we need to tell R where the file is and what it is called

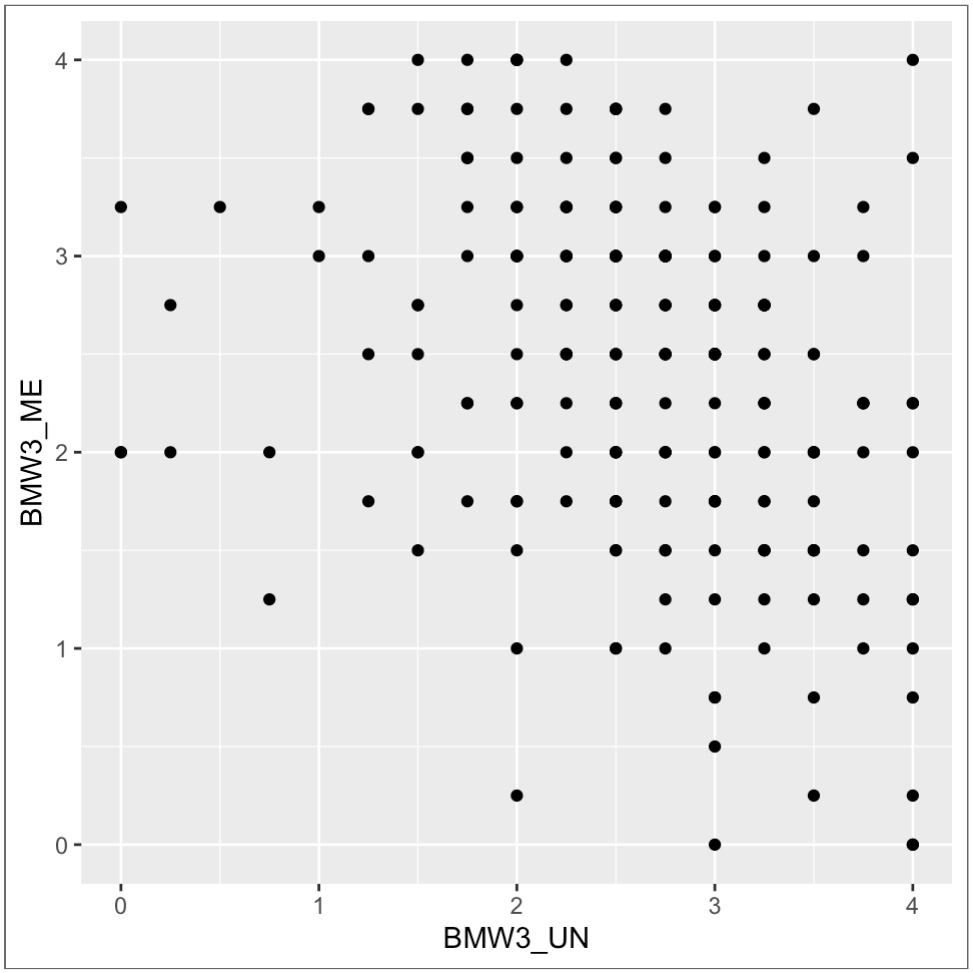
   `<-` is an assignment arrow, telling R that we want to put the data in an object (or dataframe) called `data_mw_subscales`

Step 3a: Build a scatterplot

Code | **Plot**

```
1 ggplot(data = data_mw_subscales, mapping = aes(x = BMW3_UN, y= BMW3_
2   geom_point()
```

1. We use `ggplot()` to tell R what data to plot and which variables to use
2. We use `geom_point()` to tell R that we want a scatterplot

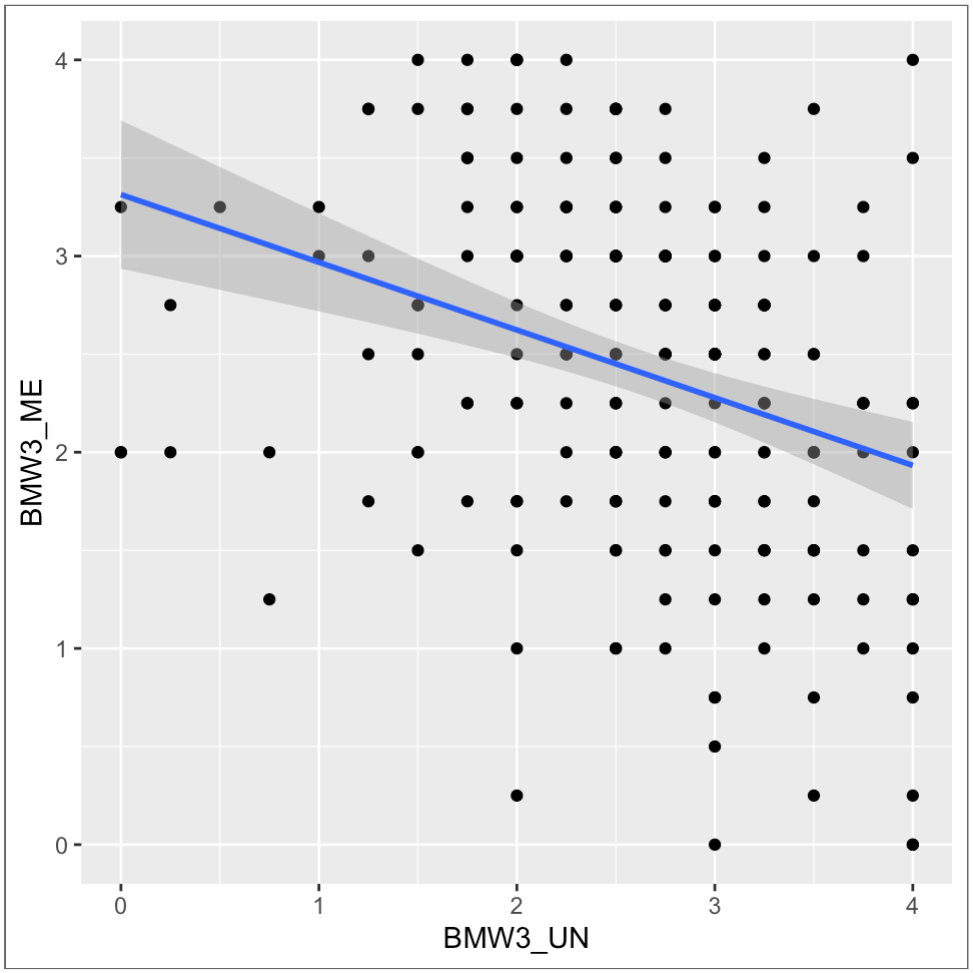

BMW3_ME
BMW3_UN
0
1
2
3
4

Step 3b: Add a line of best fit

Code | Plot

```
1 ggplot(data = data_mw_subscales, mapping = aes(x = BMW3_UN, y= BMW3_
2   geom_point() +
3   geom_smooth(method = "lm", se = TRUE)
```

3. We use `geom_smooth()`, with the `method = "lm"` to add a line of best fit; `se` refers to the standard error around the line

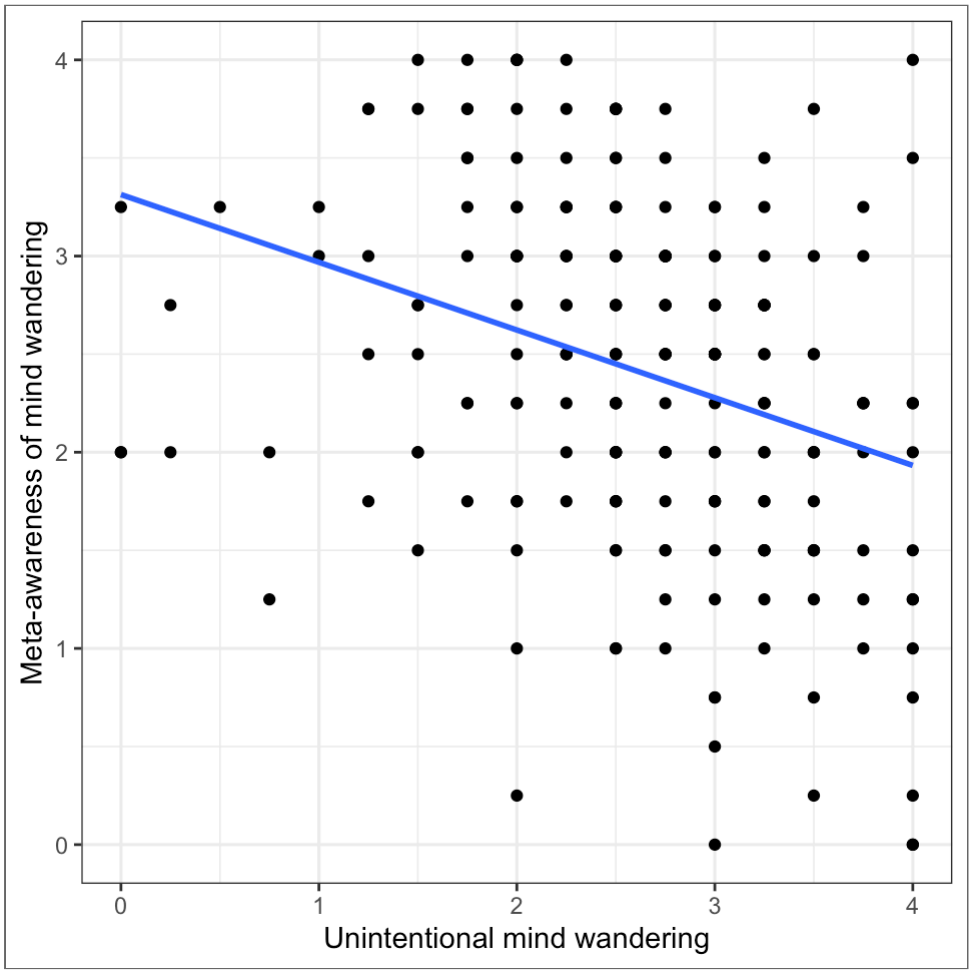

BMW3_ME
BMW3_UN
0
1
2
3
4

Step 3c: Tidy up the scatterplot

Code | Plot

```
1 ggplot(data = data_mw_subscales, mapping = aes(x = BMW3_UN, y= BMW3_
2   geom_point() +
3   geom_smooth(method = "lm", se = FALSE) +
4   theme_bw() +
5   labs(x = "Unintentional mind wandering", y = "Meta-awareness of mi
```

4. We use `theme_bw()` to change the background
5. We use `labs()` to add informative titles to the axes


Meta-awareness of mind wandering
Unintentional mind wandering
0
1
2
3
4

Step 4: Correlation analysis

Code | **Output**

```r
correlation(data = data_mw_subscales,
            select = c("BMW3_UN", "BMW3_ME"),
            method = "pearson",
            p_adjust = "none")
```

1. We use the `correlation()` function and tell R what data to use
2. We specify the variables we want to correlate using `select()`
3. We specify which method to use; `"pearson"` is the default
4. We'll talk about that in a couple of weeks

```
# Correlation Matrix (pearson-method)

Parameter1 | Parameter2 |     r |         95% CI | t(213) |          p
-------------------------------------------------------------------
BMW3_UN    |    BMW3_ME | -0.32 | [-0.44, -0.20] |  -4.94 | < .001***
```

```
p-value adjustment method: none
Observations: 215
```