

# Research Methods:

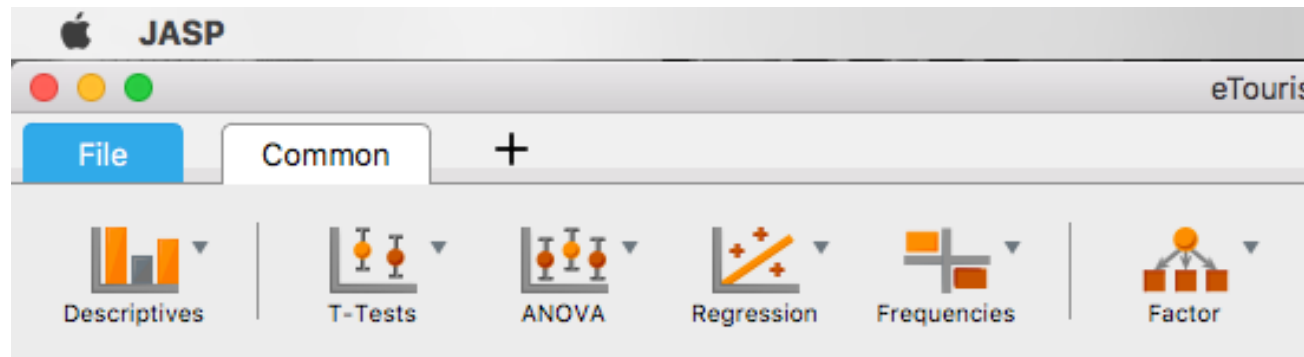
## Correlation

**Laurens Rook (Delft University of Technology)**

# Today

❖ **Correlation** (or correlational research methods and correlational statistics)

--> procedures, coefficients, and techniques



# Learning goals (Ch. 9)

- ❖ Describe the difference between strong, moderate, and weak correlation coefficients
- ❖ Capable of visual inspection and interpretation of scatter plots (valence, form, direction and strength of the correlation)
- ❖ Understand how correlations allows for prediction

# Correlation tables (example)

**Table 2**

Means, standard deviations and correlations (N = 493).

		M	SD	1	2	3	4	5	6	7
1	Gender	1.37 <sup>a</sup>	0.49	—						
2	Age	22.13	2.15	-0.04	—					
3	Price	0.48 <sup>a</sup>	0.50	0.01	-0.01	—				
4	Parking convenience	0.50 <sup>a</sup>	0.50	-0.08	-0.01	-0.03	—			
5	Car Type	0.53 <sup>a</sup>	0.50	-0.03	-0.06	0.06	0.04	—		
6	Psychological Ownership	5.72	0.97	-0.07	0.09*	-0.04	0.05	0.06	—	
7	Intention to select a shared car	0.17	1.82	0.01	-0.04	0.34**	0.19**	0.13**	0.00	—

\* $p < 0.05$ . \*\* $p < 0.01$ .

<sup>a</sup> These items take only two values.

# Correlation tables (example)

Table 6  
Correlations Between the BFI and the BFAS in Two Samples

Domain	BFI					BFAS					10 aspects									
	N	A	C	E	O	N	A	C	E	O	N <sub>V</sub>	N <sub>W</sub>	A <sub>C</sub>	A <sub>P</sub>	C <sub>I</sub>	C <sub>O</sub>	E <sub>B</sub>	E <sub>A</sub>	O <sub>I</sub>	O <sub>O</sub>
N (BFI)	—	<b>-.38</b>	<b>-.27</b>	<b>-.18</b>	<b>-.11</b>	<b>.75</b>	<b>-.13</b>	<b>-.14</b>	<b>-.26</b>	<b>-.04</b>	<b>.67</b>	<b>.67</b>	<b>-.06</b>	<b>-.17</b>	<b>-.30</b>	<b>.06</b>	<b>-.27</b>	<b>-.16</b>	<b>-.15</b>	<b>.10</b>
A (BFI)	<b>-.24</b>	—	<b>.25</b>	<b>.10</b>	<b>.03</b>	<b>-.38</b>	<b>.59</b>	<b>.13</b>	<b>.21</b>	<b>.00</b>	<b>-.44</b>	<b>-.24</b>	<b>.45</b>	<b>.55</b>	<b>.17</b>	<b>.06</b>	<b>.38</b>	<b>-.04</b>	<b>-.02</b>	<b>.02</b>
C (BFI)	<b>-.24</b>	<b>.38</b>	—	<b>.25</b>	<b>.15</b>	<b>-.29</b>	<b>.18</b>	<b>.71</b>	<b>.33</b>	<b>.13</b>	<b>-.19</b>	<b>-.32</b>	<b>.15</b>	<b>.16</b>	<b>.65</b>	<b>.54</b>	<b>.20</b>	<b>.34</b>	<b>.27</b>	<b>-.07</b>
E (BFI)	<b>-.33</b>	<b>.15</b>	<b>.18</b>	—	<b>.29</b>	<b>-.14</b>	<b>.07</b>	<b>.21</b>	<b>.76</b>	<b>.17</b>	<b>.05</b>	<b>-.30</b>	<b>.22</b>	<b>-.11</b>	<b>.24</b>	<b>.12</b>	<b>.60</b>	<b>.67</b>	<b>.22</b>	<b>.06</b>
O (BFI)	<b>-.13</b>	<b>.11</b>	<b>.11</b>	<b>.26</b>	—	<b>-.14</b>	<b>.10</b>	<b>-.01</b>	<b>.35</b>	<b>.77</b>	<b>-.06</b>	<b>-.20</b>	<b>.25</b>	<b>-.09</b>	<b>.10</b>	<b>-.11</b>	<b>.16</b>	<b>.42</b>	<b>.64</b>	<b>.64</b>
N (BFAS)	<b>.80</b>	<b>-.34</b>	<b>-.33</b>	<b>-.26</b>	<b>-.15</b>	—	<b>-.20</b>	<b>-.22</b>	<b>-.32</b>	<b>-.12</b>	<b>.89</b>	<b>.89</b>	<b>-.09</b>	<b>-.25</b>	<b>-.41</b>	<b>.04</b>	<b>-.28</b>	<b>-.27</b>	<b>-.26</b>	<b>.07</b>
A (BFAS)	<b>-.01</b>	<b>.68</b>	<b>.36</b>	<b>.06</b>	<b>.09</b>	<b>-.14</b>	—	<b>.18</b>	<b>.13</b>	<b>.12</b>	<b>-.25</b>	<b>-.10</b>	<b>.85</b>	<b>.84</b>	<b>.18</b>	<b>.12</b>	<b>.33</b>	<b>-.11</b>	<b>-.01</b>	<b>.23</b>
C (BFAS)	<b>-.15</b>	<b>.24</b>	<b>.77</b>	<b>.08</b>	<b>-.04</b>	<b>-.25</b>	<b>.22</b>	—	<b>.25</b>	<b>-.01</b>	<b>-.13</b>	<b>-.26</b>	<b>.11</b>	<b>.20</b>	<b>.83</b>	<b>.84</b>	<b>.14</b>	<b>.29</b>	<b>.12</b>	<b>-.14</b>
E (BFAS)	<b>-.36</b>	<b>.31</b>	<b>.33</b>	<b>.78</b>	<b>.34</b>	<b>-.33</b>	<b>.23</b>	<b>.24</b>	—	<b>.34</b>	<b>-.10</b>	<b>-.47</b>	<b>.32</b>	<b>-.10</b>	<b>.35</b>	<b>.07</b>	<b>.85</b>	<b>.84</b>	<b>.40</b>	<b>.15</b>
O (BFAS)	<b>-.21</b>	<b>.17</b>	<b>.31</b>	<b>.22</b>	<b>.67</b>	<b>-.20</b>	<b>.28</b>	<b>.19</b>	<b>.37</b>	—	<b>-.04</b>	<b>-.18</b>	<b>.27</b>	<b>-.07</b>	<b>.14</b>	<b>-.15</b>	<b>.19</b>	<b>.38</b>	<b>.85</b>	<b>.81</b>
Volatility (N <sub>V</sub> )	<b>.67</b>	<b>-.40</b>	<b>-.25</b>	<b>-.10</b>	<b>-.08</b>	<b>.90</b>	<b>-.24</b>	<b>-.17</b>	<b>-.16</b>	<b>-.15</b>	—	<b>.59</b>	<b>-.09</b>	<b>-.34</b>	<b>-.28</b>	<b>.06</b>	<b>-.12</b>	<b>-.06</b>	<b>-.14</b>	<b>.07</b>
Withdrawal (N <sub>W</sub> )	<b>.76</b>	<b>-.20</b>	<b>-.34</b>	<b>-.38</b>	<b>-.19</b>	<b>.88</b>	<b>.00</b>	<b>-.29</b>	<b>-.44</b>	<b>-.21</b>	<b>.59</b>	—	<b>-.08</b>	<b>-.10</b>	<b>-.46</b>	<b>.01</b>	<b>-.38</b>	<b>-.43</b>	<b>-.32</b>	<b>.05</b>
Compassion (A <sub>C</sub> )	<b>.02</b>	<b>.54</b>	<b>.32</b>	<b>.22</b>	<b>.19</b>	<b>-.03</b>	<b>.84</b>	<b>.18</b>	<b>.40</b>	<b>.40</b>	<b>-.07</b>	<b>.02</b>	—	<b>.43</b>	<b>.13</b>	<b>.05</b>	<b>.44</b>	<b>.10</b>	<b>.11</b>	<b>.35</b>
Politeness (A <sub>P</sub> )	<b>-.04</b>	<b>.62</b>	<b>.28</b>	<b>-.12</b>	<b>-.04</b>	<b>-.20</b>	<b>.86</b>	<b>.20</b>	<b>.00</b>	<b>.08</b>	<b>-.32</b>	<b>-.02</b>	<b>.45</b>	—	<b>.18</b>	<b>.17</b>	<b>.12</b>	<b>-.29</b>	<b>-.14</b>	<b>.03</b>
Industriousness (C <sub>I</sub> )	<b>-.32</b>	<b>.25</b>	<b>.72</b>	<b>.17</b>	<b>.04</b>	<b>-.42</b>	<b>.17</b>	<b>.84</b>	<b>.31</b>	<b>.23</b>	<b>-.28</b>	<b>-.49</b>	<b>.12</b>	<b>.16</b>	—	<b>.39</b>	<b>.21</b>	<b>.39</b>	<b>.31</b>	<b>-.09</b>
Orderliness (C <sub>O</sub> )	<b>.09</b>	<b>.14</b>	<b>.55</b>	<b>-.05</b>	<b>-.11</b>	<b>.02</b>	<b>.20</b>	<b>.81</b>	<b>.07</b>	<b>.07</b>	<b>.01</b>	<b>.03</b>	<b>.17</b>	<b>.17</b>	<b>.38</b>	—	<b>.03</b>	<b>.09</b>	<b>-.10</b>	<b>-.15</b>
Enthusiasm (E <sub>B</sub> )	<b>-.27</b>	<b>.42</b>	<b>.24</b>	<b>.69</b>	<b>.20</b>	<b>-.25</b>	<b>.36</b>	<b>.16</b>	<b>.88</b>	<b>.22</b>	<b>-.15</b>	<b>-.31</b>	<b>.46</b>	<b>.15</b>	<b>.20</b>	<b>.06</b>	—	<b>.43</b>	<b>.18</b>	<b>.13</b>
Assertiveness (E <sub>A</sub> )	<b>-.36</b>	<b>.11</b>	<b>.34</b>	<b>.68</b>	<b>.39</b>	<b>-.33</b>	<b>.04</b>	<b>.25</b>	<b>.87</b>	<b>.44</b>	<b>-.13</b>	<b>-.46</b>	<b>.22</b>	<b>-.15</b>	<b>.34</b>	<b>.06</b>	<b>.52</b>	—	<b>.49</b>	<b>.13</b>
Intellect (O <sub>I</sub> )	<b>-.37</b>	<b>.10</b>	<b>.39</b>	<b>.25</b>	<b>.46</b>	<b>-.37</b>	<b>.15</b>	<b>.31</b>	<b>.42</b>	<b>.82</b>	<b>-.25</b>	<b>-.41</b>	<b>.24</b>	<b>.01</b>	<b>.40</b>	<b>.10</b>	<b>.21</b>	<b>.52</b>	—	<b>.37</b>
Openness (O <sub>O</sub> )	<b>.03</b>	<b>.17</b>	<b>.09</b>	<b>.09</b>	<b>.62</b>	<b>.06</b>	<b>.33</b>	<b>-.02</b>	<b>.18</b>	<b>.80</b>	<b>.02</b>	<b>.08</b>	<b>.42</b>	<b>.14</b>	<b>-.04</b>	<b>.01</b>	<b>.14</b>	<b>.17</b>	<b>.33</b>	—

Note. The university sample is below the diagonal; the Eugene-Springfield community sample is above. Validity coefficients across instruments are in bold italics. Correlations among the Big Five within instrument are in bold; BFI = Big Five Inventory; BFAS = Big Five Aspect Scales; N = Neuroticism; A = Agreeableness; C = Conscientiousness; E = Extraversion; O = Openness/Intellect; subscript letters represent the first letter of the aspect.

# Correlation

1.

# The goals of science

Jackson (Ch. 1, mod. 1) describes 3 basic goals of scientific research:

- ❖ **Description:** via careful observation of behavior
- ❖ **Prediction:** via identification of factors impacting behavior
- ❖ **Explanation:** via causes and mechanisms that explain the when and why of behavior

# Research methods in science

- ❖ Descriptive methods: observational, case study, survey
- ❖ Predictive methods: correlational, quasi-experimental
- ❖ Explanatory method: the experiment



# The correlational method

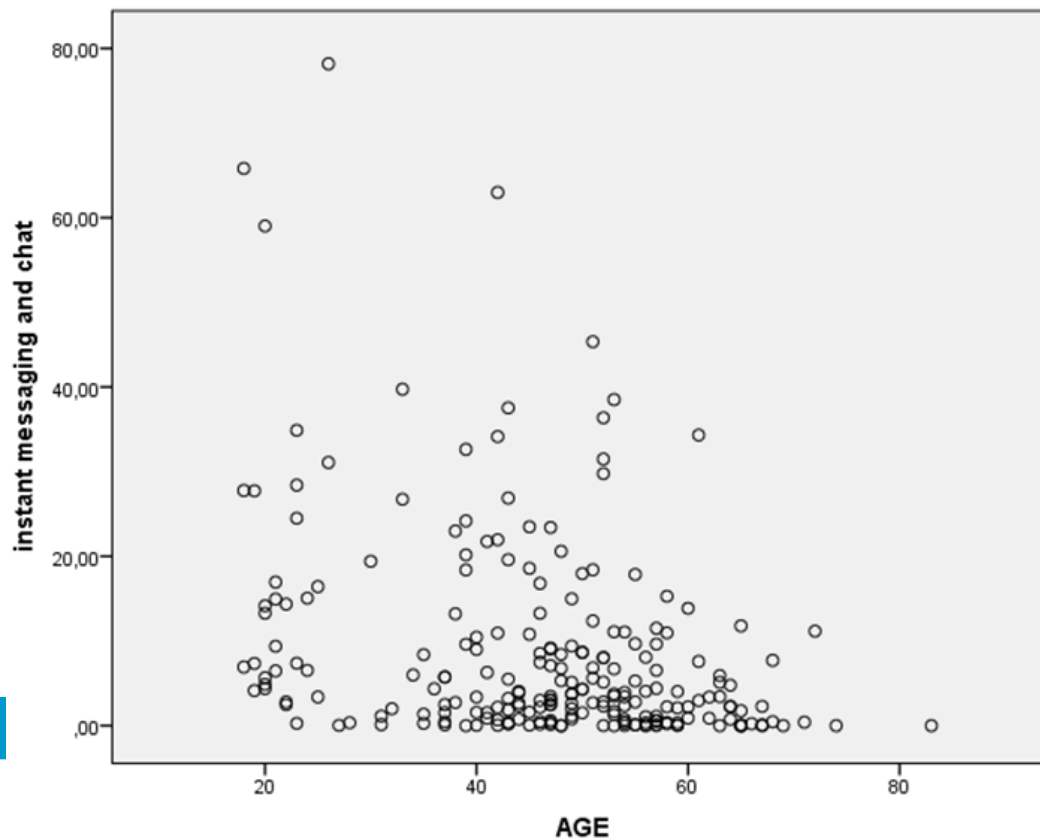
- ❖ The correlational method is a non-experimental method
- ❖ It describes the relationship between two naturally occurring (measured) variables -> with the aim to explore if and how they are 'co-related'
- ❖ Correlation between 2 variables gives us tools to develop fairly accurate predictions

# Understanding correlation coefficients

- ❖ **Magnitude** = an indication of the strength of the relationship between variable
- ❖ **Correlation coefficient** = a measure of the degree of relationship between two sets of score. For correlation ( $r$ ) it varies **between -1.00 and +1.00**
  - when  $r$  is between  $\pm .70$  and  $\pm 1.00$ : strong
  - when  $r$  is between  $\pm .30$  and  $\pm .69$ : moderate
  - when  $r$  is between  $\pm .00$  and  $\pm .29$ : none (.00) to weak

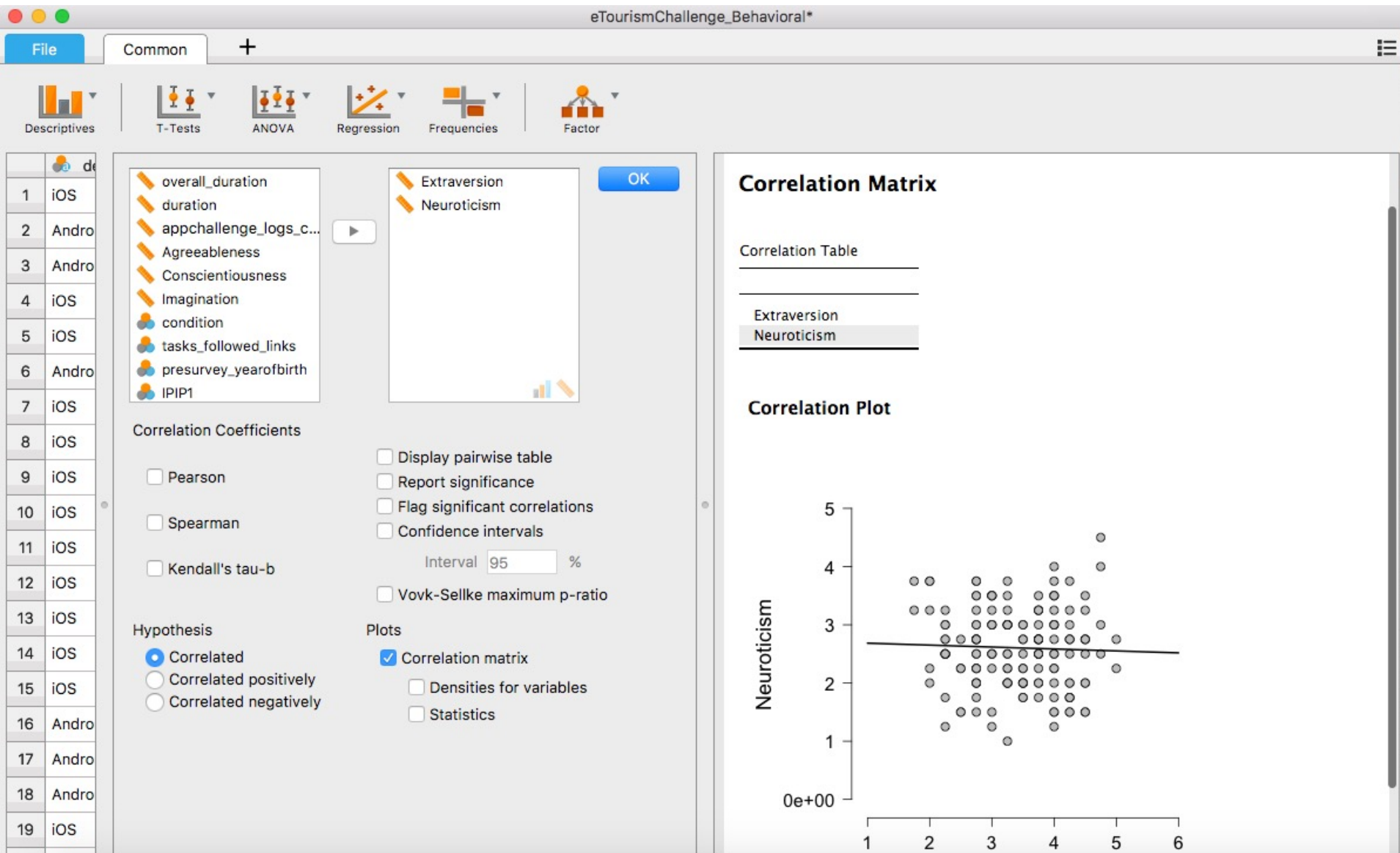
# Scatterplots

- ❖ A figure that graphically represents the relationships between two variables



Note that the x-axis and y-axis could easily be reversed on the axes. This would not make a difference in interpretation.

# In JASP



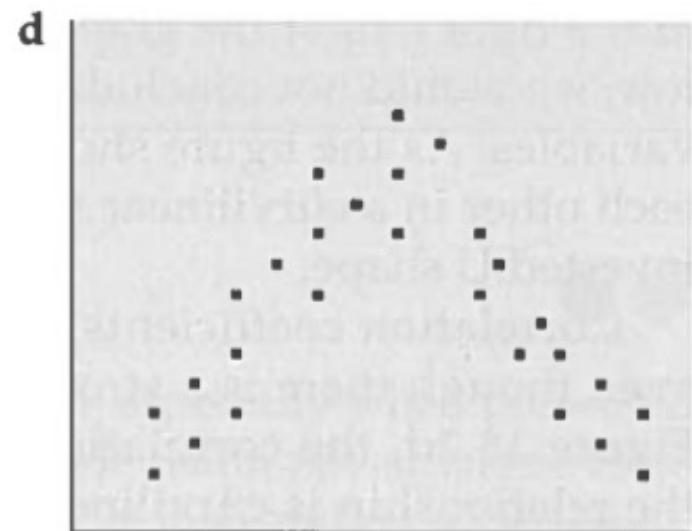
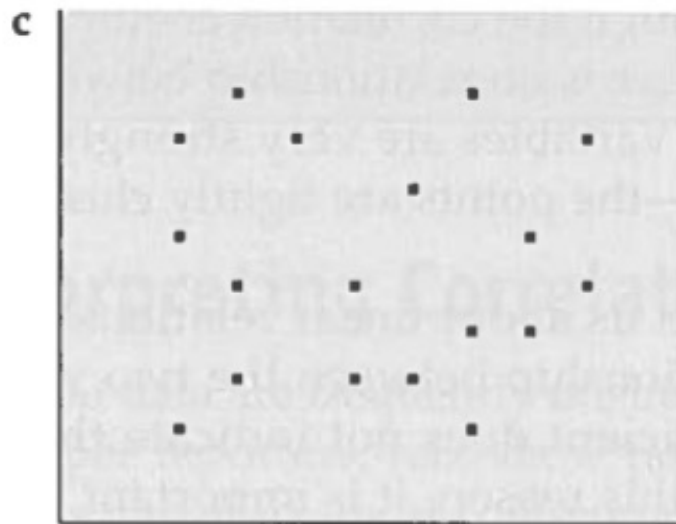
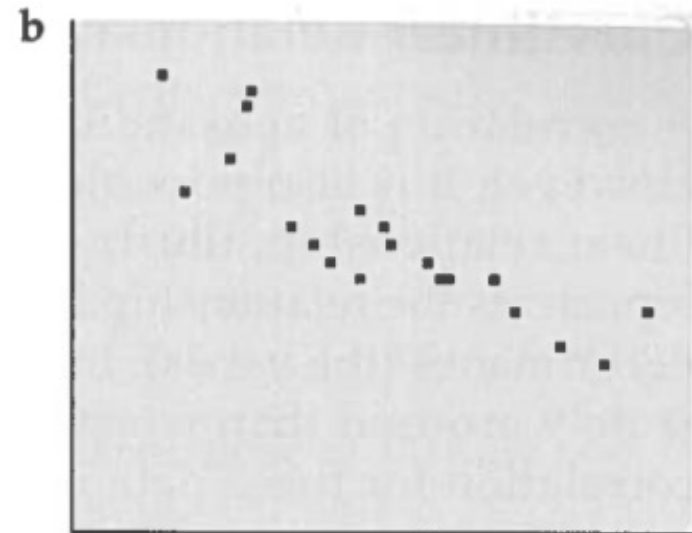
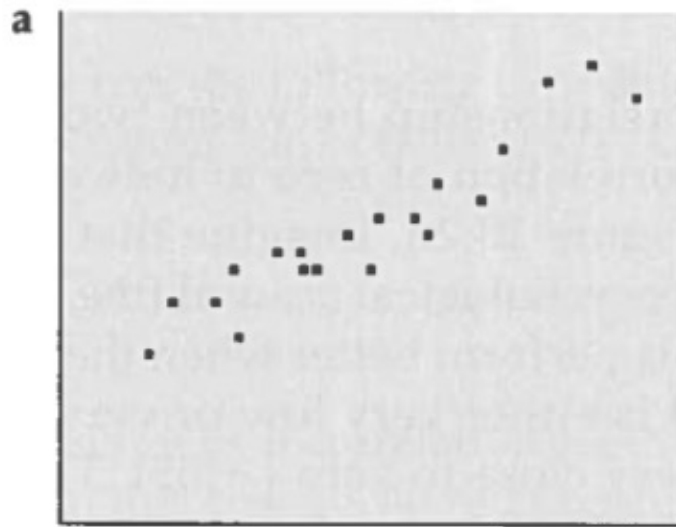
# Possible types of correlational relationship (I)

- ❖ **Positive:** A relationship in which the two variables move together (an in/decrease in one is related to an in/decrease in the other)
- ❖ **Negative:** An inverse relationship in which an in/decrease in one is related to a de/increase in the other, and v.v.

# Possible types of correlational relationship (II)

- ❖ **No relationship**: The relationship between two variables is (close to) zero, and data points are scattered in random fashion
- ❖ **Curvilinear**: Variables increase together up to a point and then, as one continues to increase, the other decreases. A curvilinear relationship between two variables also is summarized in a correlation coefficient (close to) zero (but there is a relationship!)

# Example



# Misinterpreting correlations (Ia)

- ❖ The most common mistake = interpret the observed relationship (the correlation) as a causal relationship
- ❖ **Causality** = the assumption that a correlation indicates a causal relationship between the two variables
- ❖ The correlation coefficient indicates that two variables 'move together', not that a change in Var. 1 will cause a change in Var. 2



# Misinterpreting correlations (Ib)

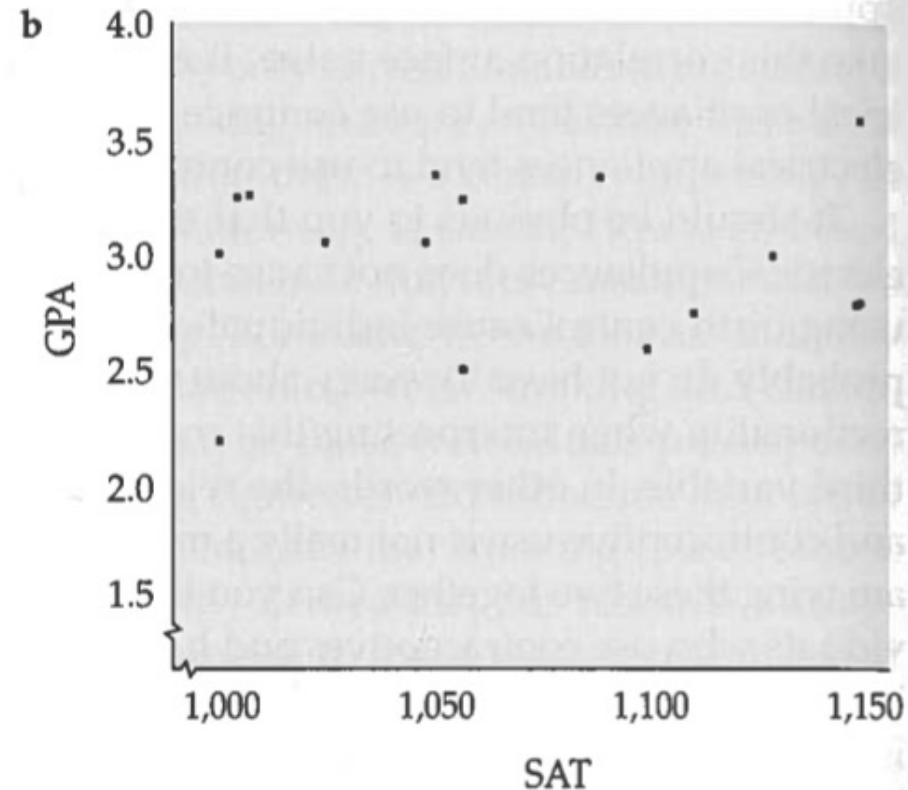
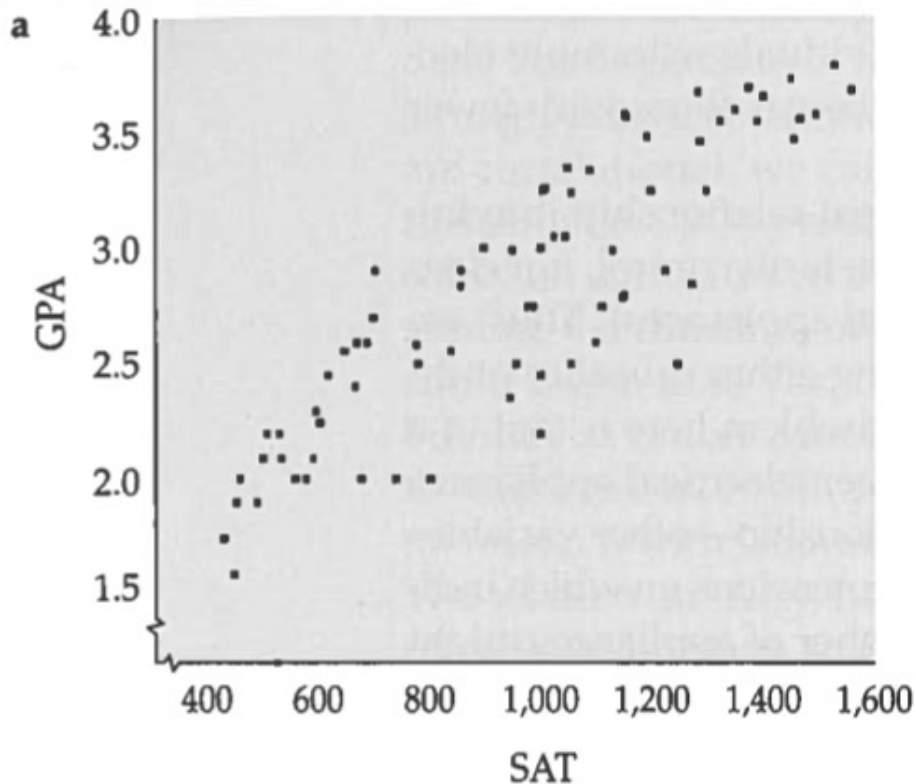
- ❖ **Directionality** = the inference made with respect to the direction of a relationship between two variables
- ❖ Again, the correlation coefficient indicates that two variables 'move together', not that a  $\pm$  change in Var. 1 will cause a  $\pm$  change in Var. 2
- ❖ You cannot draw causality-based conclusions from correlational data!

# Misinterpreting correlations (II)

- ❖ The **third-variable problem** = the problem of a correlation between two variables being dependent on another (third) variable
- ❖ You can statistically control for the third-variable problem using a procedure called **partial correlation**:
  - ❑ Measuring three variables, then remove the effect of the third variable from the correlation of the remaining two variables

# Misinterpreting correlations (III)

- ❖ **Restrictive range** = A variable does not vary enough, and it is therefore impossible to observe a relationship



# Misinterpreting correlations (IV)

- ❖ **Curvilinear relationships** = A partial relationship exists between two variables, but this relationship is cancelled out in the other half of the curve
- ❖ The correlation coefficient would be .00 or close to .00
- ❖ You would erroneously conclude that no correlation existed!

# Correlation and prediction

- ❖ **Correlation ≠ Causality.** Still, correlation indicates that when one variable is present at a certain level, the second variable *may / tend to*\* also be present at a certain level
  - \* NOTE the phrasing / vocabulary here!
- ❖ Thus, you may observe (linear) trends, but must acknowledge that not all cases will fit in – there will be outliers / exceptions

# Recall: The goals of science

Jackson (Ch. 1, mod. 1) describes 3 basic goals of scientific research:

- ❖ **Description:** via careful observation of behavior
- ❖ **Prediction:** via identification of factors impacting behavior
- ❖ **Explanation:** via causes and mechanisms that explain the when and why of behavior

# Correlation coefficients

2a.

JASP

84%

 U.S.
Wed 16:24

eTourismChallenge\_Behavioral\*

File

Common

+

Descriptives

T-Tests

ANOVA

Regression

Frequencies

Factor

	overall_duration	duration	appchallenge_logs_c...	Agreeableness	Conscientiousness	Imagination	condition	tasks_followed_links	presurvey_yearofbirth	IPIP1
1	iOS									
2	Andro									
3	Andro									
4	iOS									
5	iOS									
6	Andro									
7	iOS									
8	iOS									
9	iOS									
10	iOS									
11	iOS									
12	iOS									
13	iOS									
14	iOS									
15	iOS									
16	Andro									
17	Andro									
18	Andro									
19	iOS									

Extraversion

Neuroticism

OK

Correlation Coefficients

☒ Pearson
☐ Spearman
☒ Kendall's tau-b

☐ Display pairwise table
☐ Report significance
☐ Flag significant correlations
☐ Confidence intervals

Interval 95 %
☐ Vovk-Sellke maximum p-ratio

Hypothesis

☒ Correlated
☐ Correlated positively
☐ Correlated negatively

Plots

☐ Correlation matrix
☐ Densities for variables
☐ Statistics

Results

Correlation Matrix

Correlation Table

		Extraversion	Neuroticism
Extraversion	Pearson's r	—	
	Spearman's rho	—	
	Kendall's tau B	—	
Neuroticism	Pearson's r	−0.037	—
	Spearman's rho	−0.048	—
	Kendall's tau B	−0.037	—



# Learning goals

- ❖ Describe when it is appropriate to use the Pearson correlation coefficient, and the Spearman correlation coefficient or Kendall's tau-b (the latter one is "slides only")
- ❖ Interpret these correlation coefficients in JASP output form

# The correlational method

- ❖ The correlational method describes the relationship between two naturally occurring (measured) variables  
-> with the aim to explore if and how they are 'co-related'
- ❖ Correlation between 2 variables gives us tools to develop fairly accurate predictions

# Scales of measurement (Ch. 1)




- ❖ **Nominal scale** = a scale in which objects or individuals are broken into categories that have no numerical properties
- ❖ **Ordinal scale** = a scale in which objects or individuals are categorized, and those categories for a rank order
- ❖ **Interval scale** = a scale in which the units of measurement (intervals) between the numbers on a scale are all equal in size (and with an absolute zero for **ratio scales** --> the absence of the variable being measured)

# Scales of measurement

## FEATURES OF SCALES OF MEASUREMENT

### SCALE OF MEASUREMENT

	Nominal	Ordinal	Interval	Ratio
Examples	Ethnicity Religion Sex	Class rank Letter grade	Temperature (Fahrenheit and Celsius) Many psychological tests	Weight Height Time
Properties	Identity	Identity Magnitude	Identity Magnitude Equal unit size	Identity Magnitude Equal unit size Absolute zero
Mathematical Operations	None	Rank order	Add Subtract Multiply Divide	Add Subtract Multiply Divide
Typical Statistics Used	Mode Chi-square	Mode Median Wilcoxon tests	Mode Median Mean $t$ test ANOVA	Mode Median Mean $t$ test ANOVA

 Extraversion  
 tasks\_followed\_links  
 condition



# Pearson's $r$

- ❖ Pearson's product-moment correlation coefficient (Pearson's  $r$ ) = the most commonly used correlation coefficient
- ❖ Suits **variables at the interval and ratio scale**
- ❖ Interpretation:  $r$  between **-1.00 and + 1.00** (none, weak, moderate, strong)

# Calculating Pearson's $r$ (step 1)

- ❖ Raw scores are converted into  $z$  scores (the number of standard deviation units a raw score is above / below the mean):

$$Z = \frac{x - \mu}{\sigma}$$

$X$  = each individual score

$\mu$  = the population mean

$\sigma$  = the population standard deviation

## Calculating Pearson's $r$ (step 2)

Calculate the cross-products (multiplication) of all the individual  $z$  scores for both variables. Enter the cross-product of the  $z$  scores into the formula for Pearson's  $r$ :

$$r = \frac{\sum Z_X Z_Y}{N}$$

$Z_X$  = the  $z$  score for variable  $x$  for each individual

$Z_Y$  = the  $z$  score for variable  $y$  for each individual

$N$  = the number of individuals in the sample

# In JASP

The screenshot shows the JASP software interface with the 'Common' tab selected. The left sidebar lists 19 rows of data, with the first row labeled 'iOS' and the last row labeled 'iOS'. The main panel displays various statistical tests: Descriptives, T-Tests, ANOVA, Regression, Frequencies, and Factor. The 'Correlation Coefficients' section is active, showing options for Pearson, Spearman, and Kendall's tau-b. The 'Hypothesis' section is set to 'Correlated'. The 'Plots' section includes 'Correlation matrix', 'Densities for variables', and 'Statistics'. The 'Results' panel on the right displays a 'Correlation Matrix' for 'Extraversion' and 'Neuroticism'.

**Correlation Matrix**

Pearson Correlations

	Extraversion	Neuroticism
Extraversion	—	—
Neuroticism	-0.037	—



# Alternative correlation coefficients

- ❖ **Spearman's rank-order correlation coefficient** = The correlation coefficient ( $\rho$ ) used when one or more of the variables is measured on an ordinal (ranking) scale -> **between -1.00 and + 1.00**
- ❖ **Kendall's tau-b** = A non-parametric correlation coefficient similar to Spearman's, used for small data sets with large numbers of tied ranks (many scores have the same rank) -> **between -1.00 and + 1.00**

# In JASP

File

Common

+

Descriptives

T-Tests

ANOVA

Regression

Frequencies

Factor

	de
1	iOS
2	Andro
3	Andro
4	iOS
5	iOS
6	Andro
7	iOS
8	iOS
9	iOS
10	iOS
11	iOS
12	iOS
13	iOS
14	iOS
15	iOS
16	Andro
17	Andro
18	Andro
19	iOS

overall\_duration

duration

appchallenge\_logs\_c...

Neuroticism

Agreeableness

Conscientiousness

Imagination

condition

presurvey\_yearofbirth

IPIP1

▶

tasks\_followed\_links

Extraversion

OK

Correlation Coefficients

☐ Pearson

☒ Spearman

☒ Kendall's tau-b

Interval 95 %

☐ Vovk-Sellke maximum p-ratio

Hypothesis

☒ Correlated

☐ Correlated positively

☐ Correlated negatively

Plots

☐ Correlation matrix

☐ Densities for variables

☐ Statistics

## Results

### Correlation Matrix

Correlation Table

		tasks_followed_links	Extraversion
tasks_followed_links	Spearman's rho	—	
	Kendall's tau B	—	
Extraversion	Spearman's rho	−0.009	—
	Kendall's tau B	−0.004	—

\* p < .05, \*\* p < .01, \*\*\* p < .001

# Other alternatives (book)

- ❖ **Point-biserial correlation** = the correlation coefficient used when you have one variable on a nominal scale and another variable at the interval / ratio scale
- ❖ **Phi coefficient** = the correlation coefficient used when both variables are nominal

# Ergo

- ❖ You must be careful to check the scale of measurement of the two variables you wish to compare
- ❖ ...As the type of correlation coefficient you select depends on this!

# Correlation and significance testing

2b.

# Correlation and significance testing

- ❖ Apart from looking at the valence and magnitude of a correlation coefficient, we can also check whether it is statistically significant or not
- ❖  $H_0$  = the true population correlation coefficient is zero (so, not related)
- ❖  $H_a$  = the observed correlation is not equal to zero (hence they are related)

(NOTE this is for a two-tailed test)

# Correlation and significance testing (2)

- ❖ Your book provides a table where you can manually inspect the critical values for one-tailed / two-tailed tests, and see whether you have statistical significance (also see slides on Probability & Hypothesis Testing)
- ❖ JASP can do this automatically for you --> when you select that option

# Example

- ❖ Let's explore the following:

This app came with a survey, in which participants provided scores on two variables: Behavioral Inhibition (BIS) Anxiety (BIS\_Anxiety) and FFFS\_Fear).

- ❖ How do these two sub-dimensions correlate?





# Question 1

❖ BIS\_Anxiety and FFFFS-Fear were measured on a 4-point Likert scale; (*not at all*) to (*very much so*)

☐ Which correlation coefficients are we going to select?

☐ And why?

# Question 1

❖ BIS\_Anxiety and FFFFS-Fear were measured on a 4-point Likert scale; (*not at all*) to (*very much so*)

☐ Which correlation coefficients are we going to select?

☐ And why?

Answer: Both are interval scales. In JASP this is indicated with the “ruler”-icon. Hence, we go for the Pearson coefficient.

# Question 2

❖ Do BIS\_Anxiety and FFFFS-Fear correlate with each other?

☐ If so, how, and to what extent?

☐ And why?



Descriptives



T-Tests



ANOVA



Regression



Frequencies



Factor

Neuroticism  
Agreeableness  
Conscientiousness  
Imagination  
condition  
tasks\_followed\_links  
year\_of\_birth  
IPIP1  
IPIP2  
IPIP3



BIS\_Anxiety  
FFFS\_Fear

OK

## Correlation Coefficients

☒ Pearson

☐ Spearman

☐ Kendall's tau-b

☐ Display pairwise table

☒ Report significance

☒ Flag significant correlations

☐ Confidence intervals

Interval 95 %

☐ Vovk-Sellke maximum p-ratio

## Hypothesis

☒ Correlated

☐ Correlated positively

☐ Correlated negatively

## Plots

☐ Correlation matrix

☐ Densities for variables

☐ Statistics

## Results

## Correlation Matrix

## Pearson Correlations

		BIS_Anxiety	FFFS_Fear
BIS_Anxiety	Pearson's r	—	
	p-value	—	
FFFS_Fear	Pearson's r	0.440***	—
	p-value	< .001	—

\* p &lt; .05, \*\* p &lt; .01, \*\*\* p &lt; .001

# Interpretation

❖ Pearson's  $r = 0.44$ ,  $p < .001$

□ --> we reject  $H_0$  that the variables are not related, and we accept  $H_a$  (that they are related), instead!

# Question 3

❖ Let's explore the following:

□ We have our sub-dimensions (BIS\_Anxiety and FFFS\_Fear)

□ We also have a Big Five personality test result: Extraversion, Agreeableness, Conscientiousness, Neuroticism, Imagination

❖ **Theory claims associations between BIS\_Anxiety, FFFS\_Fear, and Neuroticism**

❖ Does a correlation analysis show this?



Agreeableness  
 Conscientiousness  
 Imagination  
 condition  
 tasks\_followed\_links  
 year\_of\_birth  
 IPIP1  
 IPIP2  
 IPIP3  
 IPIP4

BIS\_Anxiety  
 FFFS\_Fear  
 Neuroticism

OK

## Correlation Coefficients

☒ Pearson☐ Spearman☐ Kendall's tau-b☐ Display pairwise table☒ Report significance☒ Flag significant correlations☐ Confidence intervals

Interval 95 %

☐ Vovk-Sellke maximum p-ratio

## Hypothesis

☒ Correlated☐ Correlated positively☐ Correlated negatively

## Plots

☐ Correlation matrix☐ Densities for variables☐ Statistics

## Results

## Correlation Matrix

## Pearson Correlations

		BIS_Anxiety	FFFS_Fear	Neuroticism
BIS_Anxiety	Pearson's r	—		
	p-value	—		
FFFS_Fear	Pearson's r	0.440***	—	
	p-value	< .001	—	
Neuroticism	Pearson's r	0.411***	0.468***	—
	p-value	< .001	< .001	—

\* p &lt; .05, \*\* p &lt; .01, \*\*\* p &lt; .001

# Question 3: Answer

❖ Answer:

- ❖ The Pearson's  $r = 0.41$  for BIS Anxiety-Neuroticism; and  $r = 0.47$  for FFFS\_Fear-Neuroticism. This indicates correlations with moderate strength in both cases.
- ❖ Also, in both cases, these correlations are highly significant, with p-values  $< 0.001$ .
  - --> we reject  $H_0$  that the variables are not related, and we accept  $H_a$  (that they are related), instead!



# Another example (different variables)

❖ Do the variables MaxG and CS correlate?

The screenshot shows a statistical software interface with a toolbar at the top containing icons for Descriptives, T-Tests, ANOVA, Regression, Frequencies, Factor, Network, Machine Learning, and SEM. The main window is titled "Chatbot Data\*" and is divided into two panes. The left pane is titled "Correlation Matrix" and contains a list of variables on the left: CPerson, BotPersonality, Gender, NPS, MaxG\_1, MaxG\_2, MaxG\_3, MaxG\_4, MaxG\_5, and MaxG\_6. A "Variables" box on the right contains "CS" and "MaxG". Below the variable lists are "Correlation Coefficient" options (Pearson, Spearman, Kendall's tau-b) and "Additional Options" (Display pairwise table, Report significance, Flag significant correlations, Confidence intervals). The "Report significance" option is checked. The "Interval" is set to 95%. The right pane is titled "Results" and contains a "Correlation Matrix" section with a table of Pearson Correlations.

Chatbot Data\*

Descriptives T-Tests ANOVA Regression Frequencies Factor Network Machine Learning SEM

▼ Correlation Matrix

CPerson BotPersonality Gender NPS MaxG\_1 MaxG\_2 MaxG\_3 MaxG\_4 MaxG\_5 MaxG\_6

Variables

CS MaxG

Correlation Coefficient

☒ Pearson ☐ Spearman ☐ Kendall's tau-b

Additional Options

☐ Display pairwise table ☒ Report significance ☐ Flag significant correlations ☐ Confidence intervals

Interval 95 %

Results

Correlation Matrix

Pearson Correlations

		CS	MaxG
CS	Pearson's r	—	
	p-value	—	
MaxG	Pearson's r	-0.045	—
	p-value	0.625	—

# Interpretation

❖ Pearson's  $r = -0.045$ ,  $p = .625$

□ --> we reject  $H_a$  (that they are related), and we accept  $H_0$  that the variables are not related, instead!

# Question 4

- ❖ Let's explore the following:
  - We have our sub-dimensions (BIS\_Anxiety and FFFS\_Fear)
  - We also have the Big Five sub-dimensions: Extraversion, Agreeableness, Conscientiousness, Neuroticism, Imagination
- ❖ **How do the two BIS dimensions (BIS\_Anxiety, FFFS\_Fear) and all five dimensions of the Big Five relate to each other?**
- ❖ Let's build a bigger correlation table and interpret it



overall\_duration  
duration  
appchallenge\_logs\_c...  
condition  
tasks\_followed\_links  
year\_of\_birth  
IPIP1  
IPIP2  
IPIP3  
IPIP4

BIS\_Anxiety  
FFFS\_Fear  
Extraversion  
Neuroticism  
Agreeableness  
Conscientiousness  
Imagination

OK

## Correlation Coefficients

☒ Pearson☐ Spearman☐ Kendall's tau-b☐ Display pairwise table☒ Report significance☒ Flag significant correlations☐ Confidence intervals

Interval 95 %

☐ Vovk-Sellke maximum p-ratio

## Hypothesis

☒ Correlated☐ Correlated positively☐ Correlated negatively

## Plots

☐ Correlation matrix☐ Densities for variables☐ Statistics

Imagination	Pearson's r	-0.088	-0.024	-0.067	0.187*	0.160*
	p-value	0.274	0.764	0.405	0.019	0.045

\* p &lt; .05, \*\* p &lt; .01, \*\*\* p &lt; .001

## Correlation Matrix

## Pearson Correlations

		BIS_Anxiety	FFFS_Fear	Extraversion	Neuroticism	Agreeableness
BIS_Anxiety	Pearson's r	—				
	p-value	—				
FFFS_Fear	Pearson's r	0.440***	—			
	p-value	< .001	—			
Extraversion	Pearson's r	-0.026	-0.029	—		
	p-value	0.743	0.716	—		
Neuroticism	Pearson's r	0.411***	0.468***	-0.037	—	
	p-value	< .001	< .001	0.644	—	
Agreeableness	Pearson's r	0.160*	0.113	0.267***	-0.018	—
	p-value	0.046	0.162	< .001	0.820	—
Conscientiousness	Pearson's r	0.042	-0.007	-0.037	-0.017	0.009
	p-value	0.603	0.928	0.644	0.833	0.915
Imagination	Pearson's r	-0.088	-0.024	0.187*	-0.067	0.160*
	p-value	0.274	0.764	0.019	0.405	0.045

\* p &lt; .05, \*\* p &lt; .01, \*\*\* p &lt; .001

# Correlation Matrix

Pearson Correlations

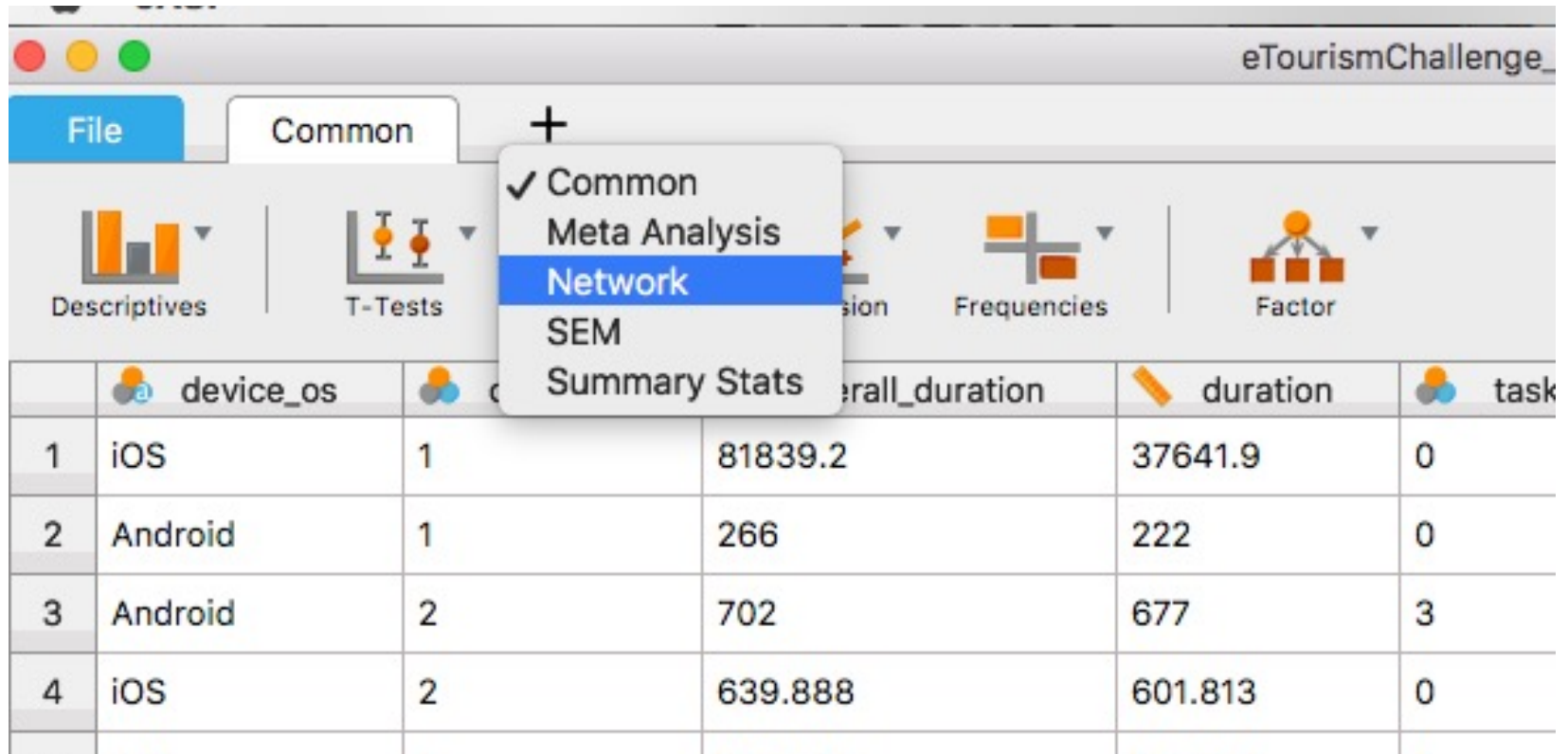
		BIS_Anxiety	FFFS_Fear	Extraversion	Neuroticism	Agreeableness	Conscientiousness	Imagination
BIS_Anxiety	Pearson's r	—						
	p-value	—						
FFFS_Fear	Pearson's r	0.440***	—					
	p-value	< .001	—					
Extraversion	Pearson's r	−0.026	−0.029	—				
	p-value	0.743	0.716	—				
Neuroticism	Pearson's r	0.411***	0.468***	−0.037	—			
	p-value	< .001	< .001	0.644	—			
Agreeableness	Pearson's r	0.160*	0.113	0.267***	−0.018	—		
	p-value	0.046	0.162	< .001	0.820	—		
Conscientiousness	Pearson's r	0.042	−0.007	−0.037	−0.017	0.009	—	
	p-value	0.603	0.928	0.644	0.833	0.915	—	
Imagination	Pearson's r	−0.088	−0.024	0.187*	−0.067	0.160*	0.107	—
	p-value	0.274	0.764	0.019	0.405	0.045	0.184	—

\* p < .05, \*\* p < .01, \*\*\* p < .001

# Interpretation

- ❖ The two main patterns are:
- ❖ We see highly significant – positive – correlations of moderate strength between Neuroticism, BIS-Anxiety and FFFS-Fear -> they move together such that low scores link with low scores, and high scores link with high scores.
- ❖ Agreeableness and Extraversion (two social components of the Big Five) also move together, in weak, positive, but significant manner

# A cool feature in JASP (not in the exam)



The screenshot shows the JASP software interface. At the top, there's a menu bar with 'File' and 'Common'. Below the menu bar, there's a toolbar with icons for 'Descriptives', 'T-Tests', 'Network', 'Frequencies', and 'Factor'. The 'Network' icon is highlighted, and a dropdown menu is open, showing the following options: 'Common', 'Meta Analysis', 'Network' (highlighted), 'SEM', and 'Summary Stats'. Below the toolbar, there's a data table with the following columns: 'device\_os', 'c', 'overall\_duration', 'duration', and 'task'. The table contains four rows of data:

	device_os	c	overall_duration	duration	task
1	iOS	1	81839.2	37641.9	0
2	Android	1	266	222	0
3	Android	2	702	677	3
4	iOS	2	639.888	601.813	0

eTourismChallenge\_Behavioral\_Extended\*

FileCommonNetwork+

Network

IPIP4

IPIP5

IPIP6

IPIP7

IPIP8

IPIP9

IPIP10

IPIP11

IPIP12

IPIP13

IPIP14

IPIP15

IPIP16

IPIP17

Dependent Variables

BIS\_Anxiety

FFFS\_Fear

Extraversion

Neuroticism

Agreeableness

Conscientiousness

Imagination

Split By

OK

Estimator

EBICglasso

Plots

☒ Network plot

☐ Centrality plot

☐ Clustering plot

Analysis Options - EBICglasso

Correlation method

Centrality Measures

Network Plot

Network

BIS\_Anxiety

FFFS\_Fear

Neuroticism


Imagination

Extraversion

Agreeableness

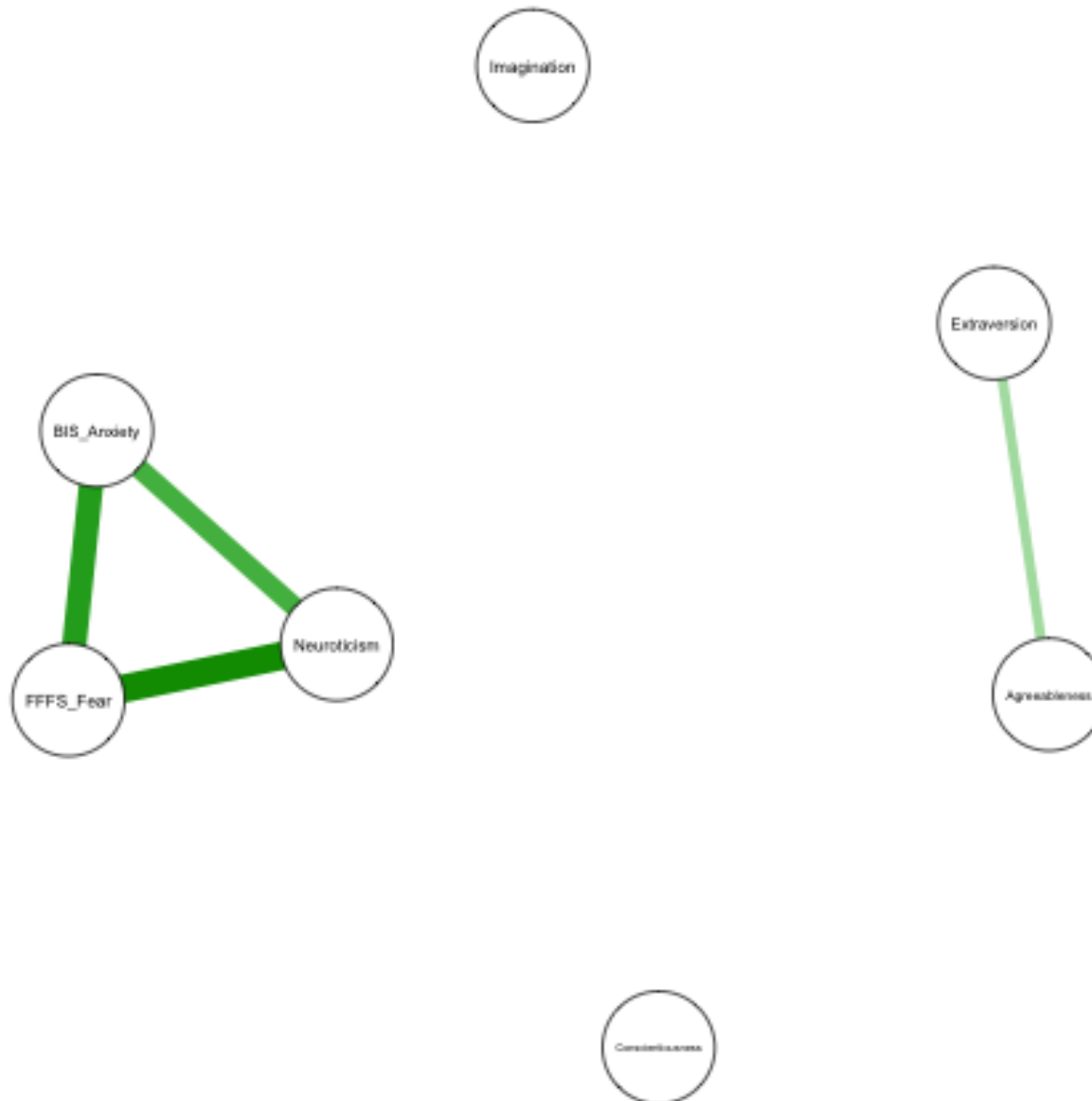
Conscientiousness

56

 TU Delft



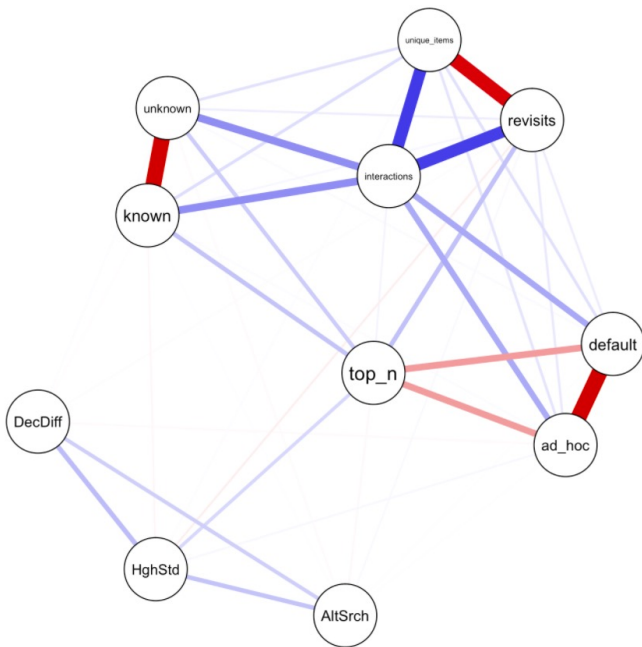
NOTE that this network graph visually confirms / illustrates what we stated for the correlation matrix: two main patterns.



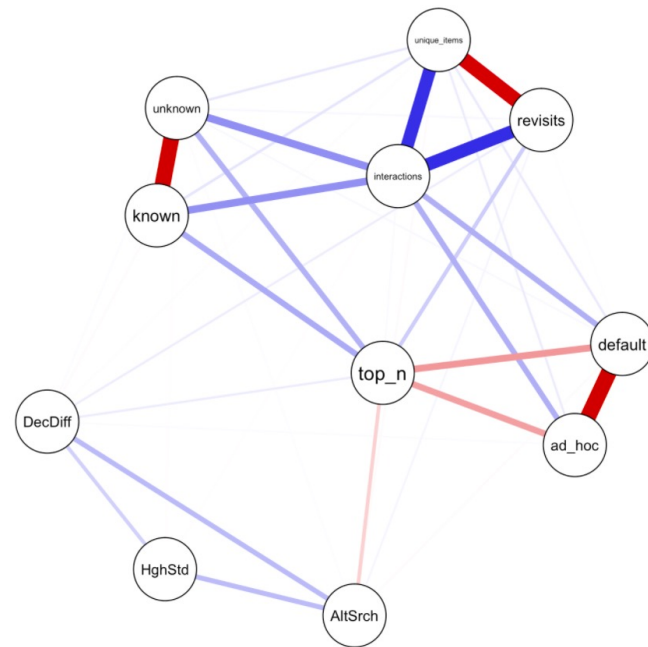
# Bigger network plots

Network Plot ▼

0



1



# In conclusion

3.

# Learning goals (I)

- ❖ Describe the difference between strong, moderate, and weak correlation coefficients
- ❖ Capable of visual inspection and interpretation of scatter plots (valence, form, direction and strength of the correlation)
- ❖ Understand how correlations allows for prediction

# Learning goals (II)

- ❖ Describe when it is appropriate to use the Pearson correlation coefficient, and when the Spearman correlation coefficient or Kendall's tau-b should be used (the latter one is "slides only")
- ❖ Interpret these correlation coefficients in JASP output form