

## Hoofdstuk 12. Correlatie – Oefenreeks – Oplossingen

### Oefening 12.1. Correlatie.

**Antwoord:** Nee, bij deze subset van de data is er geen significante samenhang tussen angst en eenzaamheid.

**Stap 1.** Kiezen van de juiste analysemethode.

**1. Hoeveel variabelen bevat mijn onderzoeksvraag?**

Twee variabelen, namelijk angst en eenzaamheid.

**2. Welk meetniveau hebben de variabelen in mijn onderzoeksvraag?**

'W1\_Angst1' tot en met 'W1\_Angst6': zes ordinale schaalitems die omgezet kunnen worden naar één metrische schaalvariabele, 'W1\_Angst\_Gem'

'W1\_Eenz1' tot en met 'W1\_Eenz4': vier ordinale schaalitems die omgezet kunnen worden naar één metrische schaalvariabele, 'W1\_Eenz\_Gem'

**3. Welke rol vervullen de variabelen?**

In deze onderzoeksvraag worden er geen specifieke rollen toegewezen aan de variabelen: je bent geïnteresseerd in het één op één verband tussen de twee variabelen. Er is hier dus geen sprake van een onderscheid tussen onafhankelijke en afhankelijke variabelen: alle variabelen kunnen zowel onafhankelijk als afhankelijk zijn.

Flowchart:

Eén afhankelijke variabele → Metrisch → Eén onafhankelijke variabele → Metrisch

→ Enkelvoudige lineaire regressie of Pearson correlatietoets.

Kijk naar de verwoording van de onderzoeksvraag: aangezien het hier gaat om het onderzoeken van samenhang tussen twee variabelen en niet om het voorspellen van een variabele aan de hand van een andere variabele, is **Pearson correlatietoets** de juiste analysetechniek.

**Stap 2.** Datacontrole.

Deze analyse moet op een subset van de data uitgevoerd worden, namelijk mensen die een diploma lagere graad middelbaar onderwijs of lager hebben (W1\_Diploma = 1). Maak eerst die selectie, voor je de inconsistentiecheck en de assumptiecheck uitvoert:

USE ALL.

COMPUTE filter\_\$=(W1\_Diploma = 1).

VARIABLE LABELS filter\_\$ 'W1\_Diploma = 1 (FILTER)'.  
VALUE LABELS filter\_\$ 0 'Not Selected' 1 'Selected'.  
FORMATS filter\_\$ (f1.0).

FILTER BY filter\_\$.

EXECUTE.

Inconsistentiecheck:

**W1\_Angst1 ANGST - Zich zenuwachtig, angstig of gespannen voelen**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Helemaal niet	4	23,5	23,5	23,5
2 Meerdere dagen, maar minder dan de helft	11	64,7	64,7	88,2
4 Bijna elke dag	2	11,8	11,8	100,0
Total	17	100,0	100,0	

**W1\_Angst2 ANGST - Niet in staat zijn om te stoppen met piekeren of om controle te krijgen over het piekeren**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Helemaal niet	8	47,1	47,1	47,1
2 Meerdere dagen, maar minder dan de helft	6	35,3	35,3	82,4
3 Meer dan de helft van de dagen	2	11,8	11,8	94,1
4 Bijna elke dag	1	5,9	5,9	100,0
Total	17	100,0	100,0	

**W1\_Angst3 ANGST - Zich te veel zorgen maken over verschillende dingen**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Helemaal niet	2	11,8	11,8	11,8
2 Meerdere dagen, maar minder dan de helft	11	64,7	64,7	76,5
3 Meer dan de helft van de dagen	4	23,5	23,5	100,0
Total	17	100,0	100,0	

**W1\_Angst4 ANGST - Moeite om je te ontspannen**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Helemaal niet	7	41,2	41,2	41,2
2 Meerdere dagen, maar minder dan de helft	6	35,3	35,3	76,5
3 Meer dan de helft van de dagen	4	23,5	23,5	100,0
Total	17	100,0	100,0	

**W1\_Angst5 ANGST - Zo rusteloos zijn dat het moeilijk is om stil te zitten**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Helemaal niet	13	76,5	76,5	76,5
2 Meerdere dagen, maar minder dan de helft	3	17,6	17,6	94,1
3 Meer dan de helft van de dagen	1	5,9	5,9	100,0
Total	17	100,0	100,0	

**W1\_Angst6 ANGST - Snel geïrriteerd of prikkelbaar zijn**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 Helemaal niet	5	29,4	29,4	29,4
2 Meerdere dagen, maar minder dan de helft	8	47,1	47,1	76,5
3 Meer dan de helft van de dagen	3	17,6	17,6	94,1
4 Bijna elke dag	1	5,9	5,9	100,0
Total	17	100,0	100,0	

**W1\_Eenz1 EENZAAMHEID - Ik mis gezelschap**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 (Bijna) nooit	2	11,8	11,8	11,8
3 Soms	4	23,5	23,5	35,3
4 Vaak	7	41,2	41,2	76,5
5 Zeer vaak	4	23,5	23,5	100,0
Total	17	100,0	100,0	

**W1\_Eenz2 EENZAAMHEID - Ik voel me geïsoleerd van anderen**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 (Bijna) nooit	1	5,9	5,9	5,9
2 Zelden	7	41,2	41,2	47,1
3 Soms	3	17,6	17,6	64,7
4 Vaak	4	23,5	23,5	88,2
5 Zeer vaak	2	11,8	11,8	100,0
Total	17	100,0	100,0	

**W1\_Eenz3 EENZAAMHEID - Ik voel me alleen**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 (Bijna) nooit	3	17,6	17,6	17,6
2 Zelden	2	11,8	11,8	29,4
3 Soms	6	35,3	35,3	64,7
4 Vaak	4	23,5	23,5	88,2
5 Zeer vaak	2	11,8	11,8	100,0
Total	17	100,0	100,0	

**W1\_Eenz4 EENZAAMHEID - Er is niemand bij wie ik terecht kan**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 (Bijna) nooit	4	23,5	23,5	23,5
2 Zelden	5	29,4	29,4	52,9
3 Soms	4	23,5	23,5	76,5
4 Vaak	3	17,6	17,6	94,1
5 Zeer vaak	1	5,9	5,9	100,0
Total	17	100,0	100,0	

Geen inconsistenties. Je kunt het best nog nagaan of de schalen intern consistent zijn aan de hand van een Cronbach's alpha toets, maar dit valt buiten de scope van deze oefening. Aangezien de schaalitems over angstgevoelens en eenzaamheid gecontroleerd zijn, kun je die omzetten in twee metrische schaalvariabelen door de gemiddelde score te berekenen van elk concept (zie hoofdstuk 5):

**COMPUTE** W1\_Angst\_Gem=**MEAN**(W1\_Angst1,W1\_Angst2,W1\_Angst3,W1\_Angst4,W1\_Angst5,W1\_Angst6).  
**VARIABLE LABELS** W1\_Angst\_Gem 'Gemiddelde score op zes items over angstgevoelens'.  
**EXECUTE**.

**COMPUTE** W1\_Eenz\_Gem=**MEAN**(W1\_Eenz1,W1\_Eenz2,W1\_Eenz3,W1\_Eenz4).  
**VARIABLE LABELS** W1\_Eenz\_Gem 'Gemiddelde score op vier items over eenzaamheid'.  
**EXECUTE**.

Assumptiecheck:

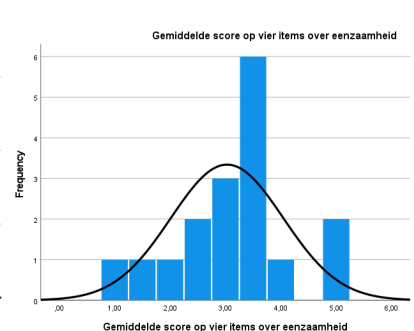
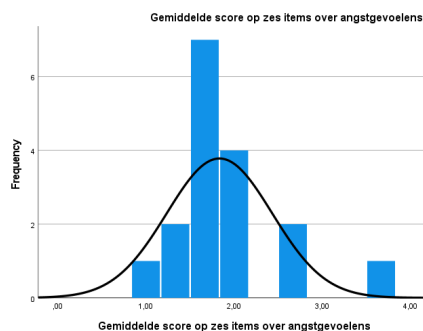
1. Beide variabelen zijn van metrisch meetniveau. **IN ORDE**
2. Gepaarde observaties: beide variabelen werden bij dezelfde groep onderzoekseenheden gemeten. **IN ORDE**
3. Normaliteit: beide variabelen zijn normaal verdeeld (of CLS:  $n > 100$ ). **CLS NIET VAN TOEPASSING** ( $n = 17$ )
4. Lineariteit: beide variabelen hebben een lineair verband met elkaar.
5. Homoscedasticiteit: homogene (of anders gezegd: gelijke) afwijkingen van alle datapunten tegenover de best passende rechte.

**Assumptie 3.** Normaliteit checken we via een histogram en z-scores van scheefheid en kurtosis van elke variabele:

```
FREQUENCIES VARIABLES=W1_Angst_Gem W1_Eenz_Gem
  /STATISTICS=STDDEV MEAN SKEWNESS SESKEW KURTOSIS SEKURT
  /HISTOGRAM NORMAL
  /ORDER=ANALYSIS.
```

**Output.** De twee histogrammen lijken een normale vorm te benaderen. Als we kijken naar de z-scores, benadert eenzaamheid de normale vorm (alle z-scores binnen +/- 1.96), maar angst niet: de z-scores van scheefheid en kurtosis vallen voor deze variabele buiten de grenzen van +/- 1.96. Beide variabelen moeten normaal verdeeld zijn, dus **is de derde assumptie niet vervuld**. In het handboek zagen we dat de flowchart in dat geval een non-parametrisch alternatief aanbiedt, namelijk de **Spearman rangcorrelatietoets**. Eén van de assumpties van deze toets stelt echter dat er minstens 30 onderzoekseenheden moeten zijn, hier is  $n$  slechts gelijk aan 17. Is deze niet vervuld, dan is er een tweede non-parametrisch alternatief voor de Pearson correlatietoets, namelijk **Kendall's tau correlatietoets**.

Statistics			
		W1_Angst_Gem	W1_Eenz_Gem
		Gemiddelde score op zes items over angstgevoelens	Gemiddelde score op vier items over eenzaamheid
N	Valid	17	17
	Missing	0	0
Mean		1,8333	3,0294
Std. Deviation		,59803	1,01505
Skewness		1,463	-,300
Std. Error of Skewness		,550	,550
Kurtosis		2,871	,313
Std. Error of Kurtosis		1,063	1,063



$$\text{Angst: } z_{\text{scheefheid}} = \frac{1.46}{0.55} = 2.66$$

$$\text{Eenzaamheid: } z_{\text{scheefheid}} = \frac{-0.30}{0.55} = -0.55$$

$$z_{\text{kurtosis}} = \frac{2.87}{1.06} = 2.70$$

$$z_{\text{kurtosis}} = \frac{0.31}{1.06} = 0.29$$

Assumptiecheck Kendall's tau correlatietoets:

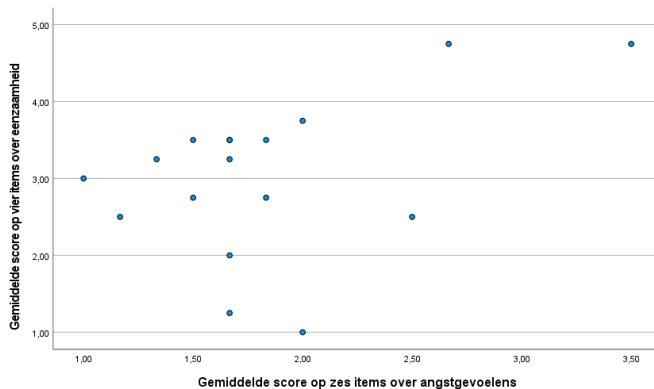
1. Beide variabelen zijn van ordinaal of metrisch meetniveau. **IN ORDE**
2. Gepaarde observaties: beide variabelen werden bij dezelfde groep onderzoekseenheden gemeten. **IN ORDE**
3. Lineariteit: beide variabelen hebben een lineair verband met elkaar. **VIA SCATTERPLOT NAGAAN**
4. Er worden minder dan 30 onderzoekseenheden onderzocht ( $n < 30$ ). **IN ORDE**

**Assumptie 3.** Lineariteit moet je controleren aan de hand van een scatterplot, wat je kunt doen aan de hand van het onderstaande SPSS-commando:

GRAPH

```
/SCATTERPLOT(BIVAR)=W1_Angst_Gem WITH W1_Eenz_Gem  
/MISSING=LISTWISE.
```

**Output.** In deze scatterplot merken we geen non-lineair verband op (zoals bv. een parabool), dus mogen we ervan uitgaan dat het verband lineair is. De derde assumptie is bij deze dus vervuld.



**Stap 3.** Hypothesen formuleren.

$H_0$ : *Er is geen samenhang tussen angstgevoelens en eenzaamheid bij mensen die een diploma lagere graad middelbaar onderwijs of lager hebben: de correlatiecoëfficiënt tussen die twee variabelen is gelijk aan nul.*

$H_a$ : *Er is wel een samenhang tussen angstgevoelens en eenzaamheid bij mensen die een diploma lagere graad middelbaar onderwijs of lager hebben: de correlatiecoëfficiënt tussen die twee variabelen is groter of kleiner dan nul.*

**Stap 4.** Significantieniveau bepalen.

Het gaat om een steekproef met minder dan 400 onderzoekseenheden ( $n = 17$ ), dus hanteer  $\alpha = .05$  of 5 %.

**Stap 5.** Toetsingsgrootte en p-waarde berekenen.

Doorloop de reeks SPSS-acties in het boek om het volgende commando te verkrijgen:

NONPAR CORR

```
/VARIABLES=W1_Angst_Gem W1_Eenz_Gem  
/PRINT=KENDALL TWOTAIL NOSIG FULL  
/MISSING=PAIRWISE.
```

## Output.

Correlations			W1_Angst_Gem Gemiddelde score op zes items over angstgevoelens	W1_Eenz_Gem Gemiddelde score op vier items over eenzaamheid
Kendall's tau_b	W1_Angst_Gem Gemiddelde score op zes items over angstgevoelens	Correlation Coefficient	1,000	,241
		Sig. (2-tailed)	.	,204
		N	17	17
	W1_Eenz_Gem Gemiddelde score op vier items over eenzaamheid	Correlation Coefficient	,241	1,000
		Sig. (2-tailed)	,204	.
		N	17	17

**Stap 6.** Conclusie en rapportering. Vergeet niet de vrijheidsgraden te vermelden bij de correlatiecoëfficiënt ( $n - 2$ ).

“Omwille van de kleine steekproefgrootte ( $n = 17$ ) werd een Kendall's tau correlatietoets gehanteerd om het bivariate verband tussen angstgevoelens en eenzaamheid te onderzoeken bij mensen die een diploma lagere graad middelbaar onderwijs of lager hebben. De toets toont aan dat het verband niet significant is, met  $\tau_b(15) = .24, p = .204$ ”.

### Oefening 12.2. Correlatie.

**Antwoord:** Ja, er is een significante en hoge samenhang tussen angst en depressie wanneer we controleren voor financiële stress, maar de samenhang is even hoog wanneer we niet voor financiële stress controleren.

**Stap 1.** Kiezen van de juiste analysemethode.

#### 1. Hoeveel variabelen bevat mijn onderzoeksvraag?

Drie variabelen, namelijk angst, depressie en financiële stress.

#### 2. Welk meetniveau hebben de variabelen in mijn onderzoeksvraag?

'W2\_Angst1' tot en met 'W2\_Angst4': vier ordinale schaalitems die omgezet kunnen worden naar één metrische schaalvariabele, 'W2\_Angst\_Gem'

'W2\_Depressie1' tot en met 'W2\_Depressie3': drie ordinale schaalitems die omgezet kunnen worden naar één metrische schaalvariabele, 'W2\_Depressie\_Gem'

'W2\_Finstress1' tot en met 'W2\_Finstress3': drie ordinale schaalitems die omgezet kunnen worden naar één metrische schaalvariabele, 'W2\_Finstress\_Gem'

#### 3. Welke rol vervullen de variabelen?

In deze onderzoeksvraag worden er geen specifieke rollen toegewezen aan de variabelen: je bent geïnteresseerd in het één op één verband tussen de twee variabelen. Er is hier dus geen sprake van een onderscheid tussen onafhankelijke en afhankelijke variabelen: alle variabelen kunnen zowel onafhankelijk als afhankelijk zijn. Financiële stress daarentegen vervult wel een specifieke rol in de onderzoeksvraag, namelijk die van covariaat.

Flowchart:

Eén afhankelijke variabele → Metrisch → Twee of meer onafhankelijke variabelen → Metrisch

→ Meervoudige lineaire regressie of partiële correlatietoets.

Kijk naar de verwoording van de onderzoeksvraag: aangezien het hier gaat om het onderzoeken van samenhang tussen twee variabelen (gecontroleerd voor een derde variabele) en niet om het voorspellen van een variabele aan de hand van een andere variabele, is **partiële correlatietoets** de juiste analysetechniek.

## Stap 2. Datacontrole.

Inconsistentiecheck:

**W2\_Angst1 ANGST - Zich zenuwachtig, angstig of gespannen voelen**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Helemaal niet	106	20,7	20,7	20,7
	2 Meerdere dagen, maar minder dan de helft	195	38,2	38,2	58,9
	3 Meer dan de helft van de dagen	101	19,8	19,8	78,7
	4 Bijna elke dag	109	21,3	21,3	100,0
	Total	511	100,0	100,0	

**W2\_Angst2 ANGST - Niet in staat zijn om te stoppen met piekeren of om controle te krijgen over het piekeren**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Helemaal niet	164	32,1	32,1	32,1
	2 Meerdere dagen, maar minder dan de helft	171	33,5	33,5	65,6
	3 Meer dan de helft van de dagen	97	19,0	19,0	84,5
	4 Bijna elke dag	79	15,5	15,5	100,0
	Total	511	100,0	100,0	

**W2\_Angst3 ANGST - Zich te veel zorgen maken over verschillende dingen**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Helemaal niet	111	21,7	21,7	21,7
	2 Meerdere dagen, maar minder dan de helft	179	35,0	35,0	56,8
	3 Meer dan de helft van de dagen	117	22,9	22,9	79,6
	4 Bijna elke dag	104	20,4	20,4	100,0
	Total	511	100,0	100,0	

**W2\_Angst4 ANGST - Moeite om je te ontspannen**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Helemaal niet	126	24,7	24,7	24,7
	2 Meerdere dagen, maar minder dan de helft	171	33,5	33,5	58,1
	3 Meer dan de helft van de dagen	99	19,4	19,4	77,5
	4 Bijna elke dag	115	22,5	22,5	100,0
	Total	511	100,0	100,0	

**W2\_Depressie1 DEPRESSIE - ...was ik gelukkig**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Zelden of nooit	25	4,9	4,9	4,9
	2 Soms of weinig	159	31,1	31,1	36,0
	3 Regelmatig	220	43,1	43,1	79,1
	4 Meestal of altijd	107	20,9	20,9	100,0
	Total	511	100,0	100,0	

**W2\_Depressie2 DEPRESSIE - ...had ik plezier in het leven**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Zelden of nooit	31	6,1	6,1	6,1
	2 Soms of weinig	170	33,3	33,3	39,3
	3 Regelmatig	216	42,3	42,3	81,6
	4 Meestal of altijd	94	18,4	18,4	100,0
	Total	511	100,0	100,0	

**W2\_Depressie3 DEPRESSIE - ...voelde ik me droevig**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Zelden of nooit	92	18,0	18,0	18,0
	2 Soms of weinig	226	44,2	44,2	62,2
	3 Regelmatig	161	31,5	31,5	93,7
	4 Meestal of altijd	32	6,3	6,3	100,0
	Total	511	100,0	100,0	

**W2\_FINSTRESS1 FINANCIËLE STRESS - Met mijn/ons huidige inkomen is het moeilijk om veel meer te veroorloven dan de basisbehoeften**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Niet akkoord	194	38,0	51,2	51,2
	2 Eerder niet akkoord	89	17,4	23,5	74,7
	3 Noch niet akkoord/noch akkoord	35	6,8	9,2	83,9
	4 Eerder akkoord	40	7,8	10,6	94,5
	5 Akkoord	21	4,1	5,5	100,0
	Total	379	74,2	100,0	
Missing	System	132	25,8		
Total		511	100,0		

**W2\_FINSTRESS2 FINANCIËLE STRESS - Ik heb het gevoel dat mijn/ons huidige inkomen toelaat een levensstandaard te behouden die ik wens**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Niet akkoord	17	3,3	4,5	4,5
	2 Eerder niet akkoord	34	6,7	9,0	13,5
	3 Noch niet akkoord/noch akkoord	36	7,0	9,5	23,0
	4 Eerder akkoord	132	25,8	34,8	57,8
	5 Akkoord	160	31,3	42,2	100,0
	Total	379	74,2	100,0	
Missing	System	132	25,8		
Total		511	100,0		

**W2\_FINSTRESS3 FINANCIËLE STRESS - Met mijn/ons huidige inkomen is het moeilijk om rond te komen (dit wil zeggen 'de eindjes aan elkaar te knopen')**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Niet akkoord	234	45,8	61,7	61,7
	2 Eerder niet akkoord	90	17,6	23,7	85,5
	3 Noch niet akkoord/noch akkoord	27	5,3	7,1	92,6
	4 Eerder akkoord	17	3,3	4,5	97,1
	5 Akkoord	11	2,2	2,9	100,0
	Total	379	74,2	100,0	
Missing	System	132	25,8		
Total		511	100,0		

Enkele zaken vallen hier op. De twee eerste items over depressie moeten omgekeerd worden, zodat een hogere score meer depressiviteit betekent. Ook het tweede item over financiële stress moet om die reden omgekeerd worden. Daarnaast hebben we 132 ontbrekende waarden bij de items rond financiële stress, maar dit komt omdat studenten deze vraag niet kregen. De correlatieanalyse zal dus slechts op 379 onderzoekseenheden uitgevoerd worden. Keer de items om waar nodig en maak dan de drie schaalvariabelen aan die we nodig hebben. Vooraf kun je het best nog nagaan of de schalen intern consistent zijn aan de hand van een Cronbach's alpha toets, maar dit valt buiten de scope van deze oefening.

```
COMPUTE W2_Angst_Gem=MEAN(W2_Angst1,W2_Angst2,W2_Angst3,W2_Angst4).  
VARIABLE LABELS W2_Angst_Gem 'Gemiddelde score op vier items over angstgevoelens'.  
EXECUTE.
```

```
RECODE W2_Depressie1 (1=4) (2=3) (3=2) (4=1) (MISSING=SYSMIS) INTO W2_Depressie1_Recode.  
VARIABLE LABELS W2_Depressie1_Recode 'Omkering van W2_Depressie1'.  
EXECUTE.
```

```
RECODE W2_Depressie2 (1=4) (2=3) (3=2) (4=1) (MISSING=SYSMIS) INTO W2_Depressie2_Recode.  
VARIABLE LABELS W2_Depressie2_Recode 'Omkering van W2_Depressie2'.  
EXECUTE.
```

```
COMPUTE W2_Depressie_Gem=MEAN(W2_Depressie1_Recode,W2_Depressie2_Recode,W2_Depressie3).  
VARIABLE LABELS W2_Depressie_Gem 'Gemiddelde score op drie items rond depressie'.  
EXECUTE.
```

```
RECODE W2_FINSTRESS2 (1=5) (2=4) (3=3) (4=2) (5=1) (MISSING=SYSMIS) INTO W2_FINSTRESS2_Recode.  
VARIABLE LABELS W2_FINSTRESS2_Recode 'Omkering van W2_FINSTRESS2'.  
EXECUTE.
```

```
COMPUTE W2_FINSTRESS_Gem=MEAN(W2_FINSTRESS1, W2_FINSTRESS2_Recode, W2_FINSTRESS3).  
VARIABLE LABELS W2_FINSTRESS_Gem 'Gemiddelde score op drie items over financiële stress'.  
EXECUTE.
```

Assumptiecheck:

1. Beide variabelen en de covariaat zijn van metrisch meetniveau. **IN ORDE**
2. Gepaarde observaties: alle variabelen werden bij dezelfde groep onderzoekseenheden gemeten. **IN ORDE**
3. Normaliteit: alle variabelen zijn normaal verdeeld (of CLS:  $n > 100$ ). **IN ORDE** ( $n = 379$ )
4. Lineariteit: alle variabelen hebben een lineair verband met elkaar. **NAGAAN VIA SCATTERPLOTS**
5. Homoscedasticiteit: homogene (of anders gezegd: gelijke) afwijkingen van alle datapunten tegenover de best passende rechte. **NAGAAN VIA SCATTERPLOTS**

**Assumptie 4 en 5.** Lineariteit en homoscedasticiteit moet je controleren aan de hand van een scatterplot tussen elk paar variabelen, wat je kunt doen aan de hand van de onderstaande SPSS-commando's:

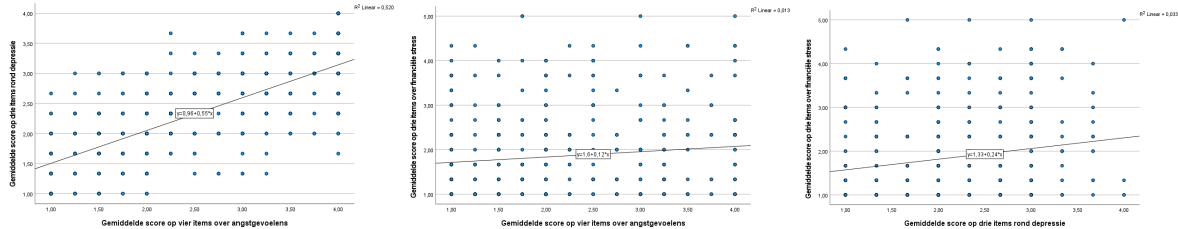
```
GRAPH  
/SCATTERPLOT(BIVAR)=W2_Angst_Gem WITH W2_Depressie_Gem  
/MISSING=LISTWISE.
```

```
GRAPH  
/SCATTERPLOT(BIVAR)=W2_Angst_Gem WITH W2_Finstress_Gem  
/MISSING=LISTWISE.
```

## GRAPH

```
/SCATTERPLOT(BIVAR)=W2_Depressie_Gem WITH W2_Finstress_Gem  
/MISSING=LISTWISE.
```

**Output.** In deze scatterplots merken we geen non-lineaire verbanden op (zoals bv. een parabool), dus mogen we ervan uitgaan dat het verband telkens lineair is. De derde assumptie is bij deze dus vervuld. De datapunten wijken ook niet ongelijk af van de best passende rechte als we die van links naar rechts afgaan (bv. geen kegelvorm, zie handboek p. 450-451), dus kunnen we van homoscedasticiteit uitgaan en zijn alle assumpties vervuld.



### Stap 3. Hypothesen formuleren.

$H_0$ : Er is geen samenhang tussen angstgevoelens en depressie, gecontroleerd voor financiële stress: de correlatiecoëfficiënt tussen die twee variabelen is gelijk aan nul.

$H_a$ : Er is wel een samenhang tussen angstgevoelens en depressie, gecontroleerd voor financiële stress: de correlatiecoëfficiënt tussen die twee variabelen is groter of kleiner dan nul.

### Stap 4. Significantieniveau bepalen.

Het gaat om een steekproef met minder dan 400 onderzoekseenheden ( $n = 379$ ), dus hanteer  $\alpha = .05$  of 5 %.

### Stap 5. Toetsingsgrootheid en p-waarde berekenen.

Doorloop de reeks SPSS-acties in het boek om het volgende commando te verkrijgen:

## PARTIAL CORR

```
/VARIABLES=W2_Angst_Gem W2_Depressie_Gem BY W2_FINSTRESS_Gem  
/SIGNIFICANCE=TWOTAIL  
/STATISTICS=DESCRIPTIVES CORR  
/MISSING=LISTWISE.
```

## Output.



### Correlations

Control Variables			W2_Angst_Gem Gemiddelde score op vier items over angstgevoelens	W2_Depressie_Gem Gemiddelde score op drie items rond depressie	W2_FINSTRESS_Gem Gemiddelde score op drie items over financiële stress
-none- <sup>a</sup>	W2_Angst_Gem Gemiddelde score op vier items over angstgevoelens	Correlation	1,000	,725	,115
		Significance (2-tailed)	.	<,001	,025
		df	0	377	377
	W2_Depressie_Gem Gemiddelde score op drie items rond depressie	Correlation	,725	1,000	,183
		Significance (2-tailed)	<,001	.	<,001
		df	377	0	377
W2_FINSTRESS_Gem Gemiddelde score op drie items over financiële stress	Correlation	,115	,183	1,000	
	Significance (2-tailed)	,025	<,001	.	
	df	377	377	0	
W2_FINSTRESS_Gem Gemiddelde score op drie items over financiële stress	W2_Angst_Gem Gemiddelde score op vier items over angstgevoelens	Correlation	1,000	,721	
		Significance (2-tailed)	.	<,001	
		df	0	376	
	W2_Depressie_Gem Gemiddelde score op drie items rond depressie	Correlation	,721	1,000	
		Significance (2-tailed)	<,001	.	
		df	376	0	

a. Cells contain zero-order (Pearson) correlations.

### Descriptive Statistics

	Mean	Std. Deviation	N
W2_Angst_Gem Gemiddelde score op vier items over angstgevoelens	2,1979	,94426	379
W2_Depressie_Gem Gemiddelde score op drie items rond depressie	2,1750	,73273	379
W2_FINSTRESS_Gem Gemiddelde score op drie items over financiële stress	1,8584	,97750	379

**Stap 6.** Conclusie en rapportering. Vergeet niet de vrijheidsgraden te vermelden bij de correlatiecoëfficiënt, wat hier ( $n - 3$ ) is vanwege de covariaat.

“Aan de hand van een partiële Pearson correlatietoets werd het verband onderzocht tussen angst ( $M = 2.20$ ,  $SD = 0.94$ ) en depressieve gevoelens ( $M = 2.18$ ,  $SD = 0.73$ ), gecontroleerd voor financiële stress ( $M = 1.86$ ,  $SD = 0.98$ ). De correlatie tussen angst en depressieve gevoelens is, zonder controle voor financiële stress, significant ( $r(377) = .73$ ,  $p < .001$ ). Deze correlatie is ongeveer even hoog als wanneer we controleren voor financiële stress ( $r(376) = .72$ ,  $p < .001$ ). In beide gevallen is er sprake van een hoge correlatie tussen angst en depressieve gevoelens.”

### Oefening 12.3. Correlatie.

**Antwoord:** Ja, een Spearman rangcorrelatietoets toont aan dat er een significant maar zwak verband is tussen werk- of studiestress en concentratie.

**Stap 1.** Kiezen van de juiste analysemethode.

#### 1. Hoeveel variabelen bevat mijn onderzoeksvraag?

Twee variabelen, namelijk werk- of studiestress en concentratie.

#### 2. Welk meetniveau hebben de variabelen in mijn onderzoeksvraag?

‘W1\_W\_STRESS2’: ordinale variabele met vijf antwoordcategorieën

‘W1\_W\_CONC2’: ordinale variabele met vijf antwoordcategorieën

#### 3. Welke rol vervullen de variabelen?

In deze onderzoeksvraag worden er geen specifieke rollen toegewezen aan de variabelen: je bent geïnteresseerd in het één op één verband tussen de twee variabelen. Er is hier dus geen sprake van een onderscheid tussen onafhankelijke en afhankelijke variabelen: alle variabelen kunnen zowel onafhankelijk als afhankelijk zijn.

Flowchart:

Eén afhankelijke variabele → Categorical → Eén onafhankelijke variabele → Categorical → Ordinaal → Meer dan twee → Verschillende

→ Kruskal-Wallis toets of Spearman rangcorrelatietoets.

Kijk naar de verwoording van de onderzoeksvraag: aangezien het hier gaat om het onderzoeken van samenhang tussen twee variabelen en niet om het vergelijken van twee of meer gemiddelden van verschillende groepen, is **Spearman rangcorrelatietoets** de juiste analysetechniek.

## Stap 2. Datacontrole.

Inconsistentiecheck:

W1_W_STRESS2 WERK/STUDIESTRESS: hoe vaak... - ...had je het gevoel dat je gestresseerd was door je werk/studie?					W1_W_CONC2 WERK/STUDIECONCENTRATIE: in welke mate... - ...kan jij ongestoord werken/studie?					
		Frequency	Percent	Valid Percent	Cumulative Percent		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 (Bijna) nooit	156	7,7	7,7	7,7	Valid	1 Veel minder dan gewoonlijk	324	16,0	16,0
	2 Zelden	259	12,8	12,8	20,5		2 Minder dan gewoonlijk	695	34,4	34,4
	3 Soms	635	31,4	31,4	52,0		3 Net zoveel als gewoonlijk	606	30,0	30,0
	4 Vaak	594	29,4	29,4	81,4		4 Meer dan gewoonlijk	322	15,9	15,9
	5 Heel vaak	376	18,6	18,6	100,0		5 Veel meer dan gewoonlijk	72	3,6	3,6
	Total	2.020	100,0	100,0			Total	2.019	100,0	100,0
						Missing	999	1	,0	
						Total		2.020	100,0	

Geen inconsistenties, behalve bij de ordinale variabele rond concentratie: daar is er één ontbrekende waarde, maar dat lijkt een willekeurig ontbrekende waarde en kunnen we niets aan doen. Die onderzoekseenheid zal niet meegenomen worden in de analyse.

Assumptiecheck:

1. Beide variabelen zijn van ordinaal of metrisch meetniveau. **IN ORDE**
2. Gepaarde observaties: beide variabelen werden bij dezelfde groep onderzoekseenheden gemeten. **IN ORDE**
3. Lineariteit: beide variabelen hebben een lineair verband met elkaar. **NAGAAN VIA SCATTERPLOT**
4. Er werden 30 of meer onderzoekseenheden onderzocht ( $n \geq 30$ ). **IN ORDE**

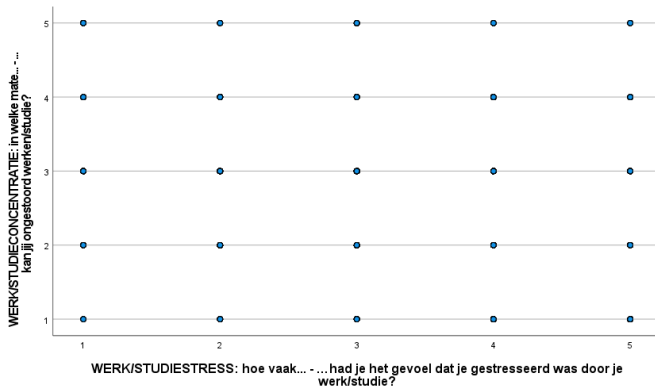
**Assumptie 3.** Lineariteit checken we via een scatterplot van de onderzochte variabelen:

GRAPH

/SCATTERPLOT(BIVAR)=W1\_W\_STRESS2 WITH W1\_W\_CONC2

/MISSING=LISTWISE.

**Output.** In deze scatterplot merken we geen non-lineair verband op (zoals bv. een parabool), dus mogen we ervan uitgaan dat het verband lineair is. De derde assumptie is bij deze dus vervuld.



### Stap 3. Hypothesen formuleren.

$H_0$ : Er is geen samenhang tussen werk- of studiestress en concentratie: de rangcorrelatiecoëfficiënt tussen die twee variabelen is gelijk aan nul.

$H_a$ : Er is wel een samenhang tussen werk- of studiestress en concentratie: de rangcorrelatiecoëfficiënt tussen die twee variabelen is groter of kleiner dan nul.

### Stap 4. Significantieniveau bepalen.

Het gaat om een steekproef met meer dan 400 onderzoekseenheden ( $n = 2019$ ), dus hanteer  $\alpha = .01$  of 1 %.

### Stap 5. Toetsingsgrootheid en p-waarde berekenen.

Doorloop de reeks SPSS-acties in het boek om het volgende commando te verkrijgen:

#### NONPAR CORR

```
/VARIABLES=W1_W_STRESS2 W1_W_CONC2
/PRINT=SPEARMAN TWOTAIL NOSIG FULL
/MISSING=PAIRWISE.
```

### Output.

Correlations				
	W1_W_STRESS2 WERKSTUDIESTRESS: hoe vaak... - ...had je het gevoel dat je gestresseerd was door je werk/studie?		W1_W_CONC2 WERKSTUDIECONCENTRATIE: in welke mate... - ...kan jij ongestoord werken/studie?	
Spearman's rho	W1_W_STRESS2	Correlation Coefficient	1,000	-,271**
	WERKSTUDIESTRESS: hoe vaak... - ...had je het gevoel dat je gestresseerd was door je werk/studie?	Sig. (2-tailed)	.	<,001
		N	2.020	2.019
	W1_W_CONC2	Correlation Coefficient	-,271**	1,000
	WERKSTUDIECONCENTRATIE: in welke mate... - ...kan jij ongestoord werken/studie?	Sig. (2-tailed)	<,001	.
		N	2.019	2.019

\*\* Correlation is significant at the 0.01 level (2-tailed).

### Stap 6. Conclusie en rapportering. Vergeet niet de vrijheidsgraden te vermelden bij de correlatiecoëfficiënt ( $n - 2$ ).

“Een Spearman rangcorrelatietoets werd gebruikt om het bivariate verband tussen werk- of studiestress en concentratie te onderzoeken. De toets toont aan dat dit verband significant is, met  $r_s(2017) = -.27, p < .001$ . Het gaat echter om een lage correlatie.”

#### Oefening 12.4. Correlatie.

**Antwoord:** Ja, een Pearson correlatietoets toont aan dat er een significant maar zeer zwak verband is tussen coronastress en werk- of studieconcentratie.

**Stap 1.** Kiezen van de juiste analysemethode.

**1. Hoeveel variabelen bevat mijn onderzoeksvraag?**

Twee variabelen, namelijk coronastress en werk- of studieconcentratie.

**2. Welk meetniveau hebben de variabelen in mijn onderzoeksvraag?**

‘W1\_Corstress1’ tot en met ‘W1\_Corstress3’: drie ordinale schaalitems die omgezet kunnen worden naar één metrische schaalvariabele, ‘W1\_Corstress\_Gem’

‘W1\_W\_CONC1’ tot en met ‘W1\_W\_CONC3’: drie ordinale schaalitems die omgezet kunnen worden naar één metrische schaalvariabele, ‘W1\_W\_CONC\_Gem’

**3. Welke rol vervullen de variabelen?**

In deze onderzoeksvraag worden er geen specifieke rollen toegewezen aan de variabelen: je bent geïnteresseerd in het één op één verband tussen de twee variabelen. Er is hier dus geen sprake van een onderscheid tussen onafhankelijke en afhankelijke variabelen: alle variabelen kunnen zowel onafhankelijk als afhankelijk zijn.

Flowchart:

Eén afhankelijke variabele → Metrisch → Eén onafhankelijke variabele → Metrisch

→ Enkelvoudige lineaire regressie of Pearson correlatietoets.

Kijk naar de verwoording van de onderzoeksvraag: aangezien het hier gaat om het onderzoeken van samenhang tussen twee variabelen en niet om het voorspellen van een variabele aan de hand van een andere variabele, is **Pearson correlatietoets** de juiste analysetechniek.

**Stap 2.** Datacontrole.

Inconsistentiecheck:

W1_Corstress1 CORONASTRESS - Ik maak mij zorgen dat ik met het coronavirus besmet zal raken						W1_Corstress2 CORONASTRESS - Ik maak mij zorgen dat mijn naasten met het coronavirus besmet zullen raken					
		Frequency	Percent	Valid Percent	Cumulative Percent			Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Niet akkoord	247	12,2	12,2	12,2	Valid	1 Niet akkoord	51	2,5	2,5	2,5
	2 Eerder niet akkoord	549	27,2	27,2	39,4		2 Eerder niet akkoord	164	8,1	8,1	10,6
	3 Noch niet akkoord, noch akkoord	380	18,8	18,8	58,2		3 Noch niet akkoord, noch akkoord	215	10,6	10,6	21,3
	4 Eerder akkoord	617	30,5	30,5	88,8		4 Eerder akkoord	869	43,0	43,0	64,3
	5 Akkoord	227	11,2	11,2	100,0		5 Akkoord	721	35,7	35,7	100,0
	Total	2.020	100,0	100,0		Total	2.020	100,0	100,0		

**W1\_Corstress3 CORONASTRESS - Ik vind de huidige coronasituatie zeer stresserend**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Niet akkoord	108	5,3	5,3	5,3
	2 Eerder niet akkoord	336	16,6	16,6	22,0
	3 Noch niet akkoord, noch akkoord	427	21,1	21,1	43,1
	4 Eerder akkoord	757	37,5	37,5	80,6
	5 Akkoord	392	19,4	19,4	100,0
	Total	2.020	100,0	100,0	

**W1\_W\_CONC1 WERK/STUDIECONCENTRATIE: in welke mate... - ...kan jij je concentreren op je werk/studie?**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Veel minder dan gewoonlijk	370	18,3	18,3	18,3
	2 Minder dan gewoonlijk	906	44,9	44,9	63,2
	3 Net zoveel als gewoonlijk	535	26,5	26,5	89,7
	4 Meer dan gewoonlijk	177	8,8	8,8	98,4
	5 Veel meer dan gewoonlijk	32	1,6	1,6	100,0
	Total	2.020	100,0	100,0	

**W1\_W\_CONC2 WERK/STUDIECONCENTRATIE: in welke mate... - ...kan jij ongestoord werken/studie?**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Veel minder dan gewoonlijk	324	16,0	16,0	16,0
	2 Minder dan gewoonlijk	695	34,4	34,4	50,5
	3 Net zoveel als gewoonlijk	606	30,0	30,0	80,5
	4 Meer dan gewoonlijk	322	15,9	15,9	96,4
	5 Veel meer dan gewoonlijk	72	3,6	3,6	100,0
	Total	2.019	100,0	100,0	
Missing	999	1	,0		
	Total	2.020	100,0		

**W1\_W\_CONC3 WERK/STUDIECONCENTRATIE: in welke mate... - ...ben jij afgeleid terwijl je werkt/studeert?**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Veel minder dan gewoonlijk	61	3,0	3,0	3,0
	2 Minder dan gewoonlijk	233	11,5	11,5	14,6
	3 Net zoveel als gewoonlijk	547	27,1	27,1	41,7
	4 Meer dan gewoonlijk	913	45,2	45,2	86,9
	5 Veel meer dan gewoonlijk	265	13,1	13,1	100,0
	Total	2.019	100,0	100,0	
Missing	999	1	,0		
	Total	2.020	100,0		

Geen inconsistenties, behalve dat item drie over werk- of studieconcentratie omgekeerd moet worden, zodat een hogere score meer concentratie betekent. Zoals we in oefening 12.3 al opmerkten, hebben twee items rond concentratie een willekeurig ontbrekende waarde waar we weinig aan kunnen doen. Na het omkeren van 'W1\_W\_CONC3' kun je de twee schaalvariabelen laten berekenen. Vooraf kun je het best nog nagaan of de schalen intern consistent zijn aan de hand van een Cronbach's alpha toets, maar dit valt buiten de scope van deze oefening.

`COMPUTE W1_Corstress_Gem=MEAN(W1_Corstress1,W1_Corstress2,W1_Corstress3).`

`VARIABLE LABELS W1_Corstress_Gem 'Gemiddelde score op drie items over coronastress'.`

`EXECUTE.`

`RECODE W1_W_CONC3 (1=5) (2=4) (3=3) (4=2) (5=1) (MISSING=SYSMIS) INTO W1_W_CONC3_Recode.`

`VARIABLE LABELS W1_W_CONC3_Recode 'Omkering van W1_W_CONC3'.`

`EXECUTE.`

`COMPUTE W1_W_CONC_Gem=MEAN(W1_W_CONC1,W1_W_CONC2,W1_W_CONC3_Recode).`

`VARIABLE LABELS W1_W_CONC_Gem 'Gemiddelde score op drie items over werk- of studieconcentratie'.`

`EXECUTE.`

Assumptiecheck:

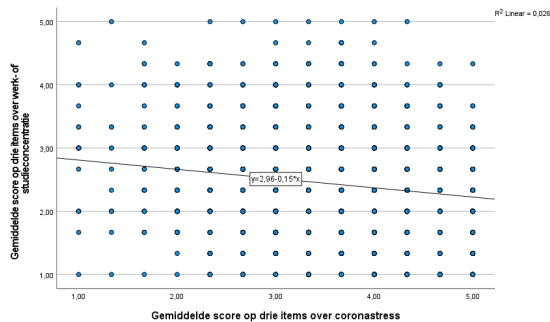
1. Beide variabelen zijn van metrisch meetniveau. **IN ORDE**
2. Gepaarde observaties: beide variabelen werden bij dezelfde groep onderzoekseenheden gemeten. **IN ORDE**
3. Normaliteit: beide variabelen zijn normaal verdeeld (of CLS:  $n > 100$ ). **IN ORDE** ( $n = 2019$ )
4. Lineariteit: beide variabelen hebben een lineair verband met elkaar. **NAGAAN VIA SCATTERPLOT**
5. Homoscedasticiteit: homogene (of anders gezegd: gelijke) afwijkingen van alle datapunten tegenover de best passende rechte. **NAGAAN VIA SCATTERPLOT**

**Assumptie 4 en 5.** Lineariteit en homoscedasticiteit moet je controleren aan de hand van een scatterplot tussen elk paar variabelen, wat je kunt doen aan de hand van de onderstaande SPSS-commando's:

`GRAPH`

```
/SCATTERPLOT(BIVAR)=W1_Corstress_Gem WITH W1_W_CONC_Gem
/MISSING=LISTWISE.
```

**Output.** In deze scatterplot merken we geen non-lineair verband op (zoals bv. een parabool), dus mogen we ervan uitgaan dat het verband lineair is. De derde assumptie is bij deze dus vervuld. De datapunten wijken ook niet ongelijk af van de best passende rechte als we die van links naar rechts afgaan (bv. geen kegelvorm, zie handboek p. 450-451), dus kunnen we van homoscedasticiteit uitgaan en zijn alle assumpties vervuld.



### Stap 3. Hypothesen formuleren.

$H_0$ : Er is geen samenhang tussen coronastress en werk- of studieconcentratie: de correlatiecoëfficiënt tussen die twee variabelen is gelijk aan nul.

$H_a$ : Er is wel een samenhang tussen coronastress en werk- of studieconcentratie: de correlatiecoëfficiënt tussen die twee variabelen is groter of kleiner dan nul.

### Stap 4. Significantieniveau bepalen.

Het gaat om een steekproef met meer dan 400 onderzoekseenheden ( $n = 2020$ ), dus hanteer  $\alpha = .01$  of 1 %.

### Stap 5. Toetsingsgroottheid en p-waarde berekenen.

Doorloop de reeks SPSS-acties in het boek om het volgende commando te verkrijgen:

#### CORRELATIONS

```
/VARIABLES=W1_Corstress_Gem W1_W_CONC_Gem
/PRINT=TWOTAIL NOSIG FULL
/STATISTICS DESCRIPTIVES
/MISSING=PAIRWISE.
```

#### Output.

Descriptive Statistics				Correlations			
	Mean	Std. Deviation	N		W1_Corstress_Gem Gemiddelde score op drie items over coronastress	W1_W_CONC_Gem Gemiddelde score op drie items over werk- of studieconcentratie	
W1_Corstress_Gem Gemiddelde score op drie items over coronastress	3,5053	,90226	2.020	W1_Corstress_Gem Gemiddelde score op drie items over coronastress	Pearson Correlation	1	-,163**
					Sig. (2-tailed)		<,001
					N	2.020	2.020
W1_W_CONC_Gem Gemiddelde score op drie items over werk- of studieconcentratie	2,4432	,81252	2.020	W1_W_CONC_Gem Gemiddelde score op drie items over werk- of studieconcentratie	Pearson Correlation	-,163**	1
					Sig. (2-tailed)	<,001	
					N	2.020	2.020

\*\* . Correlation is significant at the 0.01 level (2-tailed).

**Stap 6.** Conclusie en rapportering. Vergeet niet de vrijheidsgraden te vermelden bij de correlatiecoëfficiënt ( $n - 2$ ).

“Een Pearson correlatietoets werd gebruikt om het bivariate verband tussen coronastress en werk- of studieconcentratie te onderzoeken. Deze toets toont aan dat dit verband significant is, met  $r(2018) = -.16$ ,  $p < .001$ . Het gaat echter om een zeer lage correlatie.”

## Oefening 12.5. Correlatie.

**Antwoord:** Nee, er is geen significante samenhang tussen depressie en financiële stress wanneer we controleren voor toekomstige financiële stress.

**Stap 1.** Kiezen van de juiste analysemethode.

### 1. Hoeveel variabelen bevat mijn onderzoeksvraag?

Drie variabelen, namelijk depressie, financiële stress en toekomstige financiële stress.

### 2. Welk meetniveau hebben de variabelen in mijn onderzoeksvraag?

‘W2\_Depressie1’ tot en met ‘W2\_Depressie3’: drie ordinale schaalitems die omgezet kunnen worden naar één metrische schaalvariabele, ‘W2\_Depressie\_Gem’

‘W2\_FINSTRESS1’ tot en met ‘W2\_FINSTRESS3’: drie ordinale schaalitems die omgezet kunnen worden naar één metrische schaalvariabele, ‘W2\_FINSTRESS\_Gem’

‘W2\_FINSTRESS4’ tot en met ‘W2\_FINSTRESS6’: drie ordinale schaalitems die omgezet kunnen worden naar één metrische schaalvariabele, ‘W2\_TOEK\_FINSTRESS\_Gem’

### 3. Welke rol vervullen de variabelen?

In deze onderzoeksvraag worden er geen specifieke rollen toegewezen aan de variabelen: je bent geïnteresseerd in het één op één verband tussen de twee variabelen. Er is hier dus geen sprake van een onderscheid tussen onafhankelijke en afhankelijke variabelen: alle variabelen kunnen zowel onafhankelijk als afhankelijk zijn. Toekomstige financiële stress daarentegen vervult wel een specifieke rol in de onderzoeksvraag, namelijk die van covariaat.

Flowchart:

Eén afhankelijke variabele → Metrisch → Twee of meer onafhankelijke variabelen → Metrisch

→ Meervoudige lineaire regressie of partiële correlatietoets.

Kijk naar de verwoording van de onderzoeksvraag: aangezien het hier gaat om het onderzoeken van samenhang tussen twee variabelen (gecontroleerd voor een derde variabele) en niet om het voorspellen van een variabele aan de hand van een andere variabele, is **partiële correlatietoets** de juiste analysetechniek.

**Stap 2.** Datacontrole.

Inconsistentiecheck:

W2_FINSTRESS4 TOEK FINANCIËLE STRESS - Ik denk dat ik in de komende maanden mijn levensstandaard zal moeten verminderen					W2_FINSTRESS5 TOEK FINANCIËLE STRESS - Ik heb schrik dat ik in de nabije toekomst niet al mijn rekeningen zal kunnen betalen						
		Frequency	Percent	Valid Percent	Cumulative Percent			Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Niet akkoord	182	35,6	48,0	48,0	Valid	1 Niet akkoord	222	43,4	58,6	58,6
	2 Eerder niet akkoord	83	16,2	21,9	69,9		2 Eerder niet akkoord	95	18,6	25,1	83,6
	3 Noch niet akkoord/noch akkoord	51	10,0	13,5	83,4		3 Noch niet akkoord/noch akkoord	29	5,7	7,7	91,3
	4 Eerder akkoord	53	10,4	14,0	97,4		4 Eerder akkoord	22	4,3	5,8	97,1
	5 Akkoord	10	2,0	2,6	100,0		5 Akkoord	11	2,2	2,9	100,0
	Total	379	74,2	100,0		Total	379	74,2	100,0		
Missing	System	132	25,8			Missing	System	132	25,8		
	Total	511	100,0				Total	511	100,0		



**W2\_FINSTRESS6 TOEK FINANCIËLE STRESS - Ik verwacht dat ik of mijn gezin effectieve moeilijkheden (zoals gebrekkige huisvesting, tekort aan medische zorg of voedsel) zal ondervinden in de komende maanden**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Niet akkoord	295	57,7	77,8	77,8
	2 Eerder niet akkoord	53	10,4	14,0	91,8
	3 Noch niet akkoord/noch akkoord	24	4,7	6,3	98,2
	4 Eerder akkoord	5	1,0	1,3	99,5
	5 Akkoord	2	,4	,5	100,0
	Total	379	74,2	100,0	
Missing	System	132	25,8		
	Total	511	100,0		

Depressie en financiële stress controleerden we al in oefening 12.2. Enkele zaken vallen hier op. De twee eerste items over depressie moeten omgekeerd worden, zodat een hogere score meer depressiviteit betekent. Ook het tweede item over financiële stress moet om die reden omgekeerd worden. Daarnaast hebben we 132 ontbrekende waarden bij de items rond (toekomstige) financiële stress, maar dit komt omdat studenten deze vraag niet kregen. De correlatieanalyse zal dus slechts op 379 onderzoekseenheden uitgevoerd worden. Keer de items om indien nodig en maak dan de drie schaalvariabelen aan die we nodig hebben. Vooraf kun je het best nog nagaan of de schalen intern consistent zijn aan de hand van een Cronbach's alpha toets, maar dit valt buiten de scope van deze oefening.

`RECODE W2_Depressie1 (1=4) (2=3) (3=2) (4=1) (MISSING=SYSMIS) INTO W2_Depressie1_Recode.`

`VARIABLE LABELS W2_Depressie1_Recode 'Omkering van W2_Depressie1'.`

`EXECUTE.`

`RECODE W2_Depressie2 (1=4) (2=3) (3=2) (4=1) (MISSING=SYSMIS) INTO W2_Depressie2_Recode.`

`VARIABLE LABELS W2_Depressie2_Recode 'Omkering van W2_Depressie2'.`

`EXECUTE.`

`COMPUTE W2_Depressie_Gem=MEAN(W2_Depressie1_Recode,W2_Depressie2_Recode,W2_Depressie3).`

`VARIABLE LABELS W2_Depressie_Gem 'Gemiddelde score op drie items rond depressie'.`

`EXECUTE.`

`RECODE W2_FINSTRESS2 (1=5) (2=4) (3=3) (4=2) (5=1) (MISSING=SYSMIS) INTO W2_FINSTRESS2_Recode.`

`VARIABLE LABELS W2_FINSTRESS2_Recode 'Omkering van W2_FINSTRESS2'.`

`EXECUTE.`

`COMPUTE W2_FINSTRESS_Gem=MEAN(W2_FINSTRESS1, W2_FINSTRESS2_Recode, W2_FINSTRESS3).`

`VARIABLE LABELS W2_FINSTRESS_Gem 'Gemiddelde score op drie items over financiële stress'.`

`EXECUTE.`

`COMPUTE W2_TOEK_FINSTRESS_Gem=MEAN(W2_FINSTRESS4, W2_FINSTRESS5, W2_FINSTRESS6).`

`VARIABLE LABELS W2_TOEK_FINSTRESS_Gem 'Gemiddelde score op drie items over toekomstige financiële stress'.`

`EXECUTE.`

Assumptiecheck:

1. Beide variabelen en de covariaat zijn van metrisch meetniveau. **IN ORDE**
2. Gepaarde observaties: alle variabelen werden bij dezelfde groep onderzoekseenheden gemeten. **IN ORDE**
3. Normaliteit: alle variabelen zijn normaal verdeeld (of CLS:  $n > 100$ ). **IN ORDE** ( $n = 379$ )
4. Lineariteit: alle variabelen hebben een lineair verband met elkaar. **NAGAAN VIA SCATTERPLOTS**

5. Homoscedasticiteit: homogene (of anders gezegd: gelijke) afwijkingen van alle datapunten tegenover de best passende rechte. **NAGAAN VIA SCATTERPLOTS**

**Assumptie 4 en 5.** Lineariteit en homoscedasticiteit moet je controleren aan de hand van een scatterplot tussen elk paar variabelen, wat je kunt doen aan de hand van de onderstaande SPSS-commando's:

GRAPH

```
/SCATTERPLOT(BIVAR)=W2_Depressie_Gem WITH W2_FINSTRESS_Gem
/MISSING=LISTWISE.
```

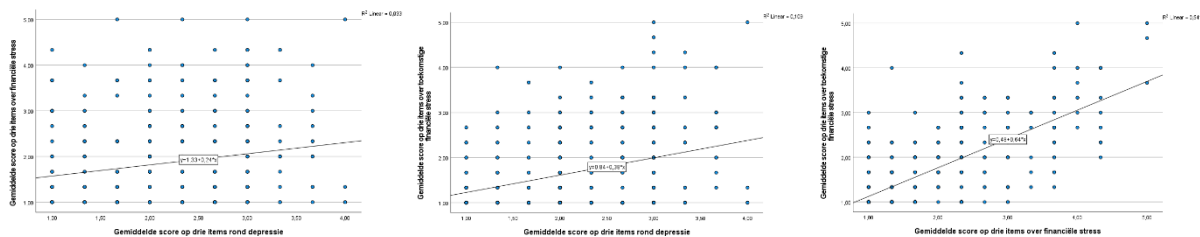
GRAPH

```
/SCATTERPLOT(BIVAR)=W2_Depressie_Gem WITH W2_TOEK_FINSTRESS_Gem
/MISSING=LISTWISE.
```

GRAPH

```
/SCATTERPLOT(BIVAR)=W2_FINSTRESS_Gem WITH W2_TOEK_FINSTRESS_Gem
/MISSING=LISTWISE.
```

**Output.** In deze scatterplots merken we geen non-lineaire verbanden op (zoals bv. een parabool), dus mogen we ervan uitgaan dat het verband telkens lineair is. De derde assumptie is bij deze dus vervuld. De datapunten wijken ook niet ongelijk af van de best passende rechte als we die van links naar rechts afgaan (bv. geen kegelvorm, zie handboek p. 450-451), dus kunnen we van homoscedasticiteit uitgaan en zijn alle assumpties vervuld.



### Stap 3. Hypothesen formuleren.

$H_0$ : Er is geen samenhang tussen depressie en financiële stress, gecontroleerd voor toekomstige financiële stress: de correlatiecoëfficiënt tussen die twee variabelen is gelijk aan nul.

$H_a$ : Er is wel een samenhang tussen depressie en financiële stress, gecontroleerd voor toekomstige financiële stress: de correlatiecoëfficiënt tussen die twee variabelen is groter of kleiner dan nul.

### Stap 4. Significantieniveau bepalen.

Het gaat om een steekproef met minder dan 400 onderzoekseenheden ( $n = 379$ ), dus hanteer  $\alpha = .05$  of 5%.

### Stap 5. Toetsingsgrootheid en p-waarde berekenen.

Doorloop de reeks SPSS-acties in het boek om het volgende commando te verkrijgen:

PARTIAL CORR

```
/VARIABLES=W2_Depressie_Gem W2_FINSTRESS_Gem BY W2_TOEK_FINSTRESS_Gem
/SIGNIFICANCE=TWOTAIL
/STATISTICS=DESCRIPTIVES CORR
/MISSING=LISTWISE.
```

## Output.

			Correlations		
Control Variables			W2_Depressie_Gem Gemiddelde score op drie items rond depressie	W2_FINSTRESS_Gem Gemiddelde score op drie items over financiële stress	W2_TOEK_FINSTRESS_Gem Gemiddelde score op drie items over toekomstige financiële stress
-none- <sup>a</sup>	W2_Depressie_Gem Gemiddelde score op drie items rond depressie	Correlation	1,000	,183	,330
		Significance (2-tailed)	.	<,001	<,001
		df	0	377	377
	W2_FINSTRESS_Gem Gemiddelde score op drie items over financiële stress	Correlation	,183	1,000	,736
		Significance (2-tailed)	<,001	.	<,001
		df	377	0	377
	W2_TOEK_FINSTRESS_Gem Gemiddelde score op drie items over toekomstige financiële stress	Correlation	,330	,736	1,000
		Significance (2-tailed)	<,001	<,001	.
		df	377	377	0
W2_TOEK_FINSTRESS_Gem Gemiddelde score op drie items over toekomstige financiële stress	W2_Depressie_Gem Gemiddelde score op drie items rond depressie	Correlation	1,000	-,094	
		Significance (2-tailed)	.	,069	
	W2_FINSTRESS_Gem Gemiddelde score op drie items over financiële stress	Correlation	-,094	1,000	
		Significance (2-tailed)	,069	.	
			df	376	0

Descriptive Statistics			
	Mean	Std. Deviation	N
W2_Depressie_Gem Gemiddelde score op drie items rond depressie	2,1750	,73273	379
W2_FINSTRESS_Gem Gemiddelde score op drie items over financiële stress	1,8584	,97750	379
W2_TOEK_FINSTRESS_Gem Gemiddelde score op drie items over toekomstige financiële stress	1,6781	,85390	379

a. Cells contain zero-order (Pearson) correlations.

**Stap 6.** Conclusie en rapportering. Vergeet niet de vrijheidsgraden te vermelden bij de correlatiecoëfficiënt, wat hier ( $n - 3$ ) is vanwege de covariaat.

“Aan de hand van een partiële Pearson correlatietoets werd het verband onderzocht tussen depressie ( $M = 2.18$ ,  $SD = 0.73$ ) en financiële stress ( $M = 1.86$ ,  $SD = 0.98$ ), gecontroleerd voor toekomstige financiële stress ( $M = 1.68$ ,  $SD = 0.85$ ). De correlatie tussen depressie en financiële stress is, zonder controle voor toekomstige financiële stress, significant ( $r(377) = .18$ ,  $p < .001$ ). Wanneer we echter controleren voor toekomstige financiële stress, is het verband niet langer significant ( $r(376) = -.09$ ,  $p = .069$ ). Het zwakke verband tussen depressie en financiële stress verdwijnt dus helemaal wanneer we rekening houden met toekomstige financiële stress, wat komt doordat depressie samenhangt met toekomstige financiële stress, en financiële stress een sterk verband heeft met toekomstige financiële stress.”

### Oefening 12.6. Correlatie.

**Antwoord:** Nee, een Spearman rangcorrelatietoets toont aan dat er geen significante samenhang is tussen zich alleen voelen en bezorgd zijn om met het coronavirus besmet te raken.

**Stap 1.** Kiezen van de juiste analysemethode.

**1. Hoeveel variabelen bevat mijn onderzoeksvraag?**

Twee variabelen, namelijk eenzaamheid en coronastress.

**2. Welk meetniveau hebben de variabelen in mijn onderzoeksvraag?**

‘W2\_Eenz3’: ordinale variabele met vijf antwoordcategorieën

‘W2\_CORSTRESS1’: ordinale variabele met vijf antwoordcategorieën

**3. Welke rol vervullen de variabelen?**

In deze onderzoeksvraag worden er geen specifieke rollen toegewezen aan de variabelen: je bent geïnteresseerd in het één op één verband tussen de twee variabelen. Er is hier dus geen sprake van een onderscheid tussen onafhankelijke en afhankelijke variabelen: alle variabelen kunnen zowel onafhankelijk als afhankelijk zijn.

Flowchart:

Eén afhankelijke variabele → Categorisch → Eén onafhankelijke variabele → Categorisch → Ordinaal → Meer dan twee → Verschillende

→ Kruskal-Wallis toets of Spearman rangcorrelatietoets.

Kijk naar de verwoording van de onderzoeksvraag: aangezien het hier gaat om het onderzoeken van samenhang tussen twee variabelen en niet om het vergelijken van twee of meer gemiddelden van verschillende groepen, is **Spearman rangcorrelatietoets** de juiste analysetechniek.

## Stap 2. Datacontrole.

Inconsistentiecheck:

W2_Eenz3 EENZAAMHEID - Ik voel me alleen					W2_Corstress1 CORONASTRESS - Ik maak mij zorgen dat ik met het coronavirus besmet zal raken					
		Frequency	Percent	Valid Percent	Cumulative Percent		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 (Bijna) nooit	104	20,4	20,4	20,4	Valid	1 Niet akkoord	90	17,6	17,6
	2 Zelden	123	24,1	24,1	44,4		2 Eerder niet akkoord	135	26,4	44,0
	3 Soms	152	29,7	29,7	74,2		3 Noch niet akkoord, noch akkoord	102	20,0	64,0
	4 Vaak	82	16,0	16,0	90,2		4 Eerder akkoord	149	29,2	93,2
	5 Zeer vaak	50	9,8	9,8	100,0		5 Akkoord	35	6,8	100,0
Total		511	100,0	100,0		Total		511	100,0	100,0

Geen inconsistenties.

Assumptiecheck:

1. Beide variabelen zijn van ordinaal of metrisch meetniveau. **IN ORDE**
2. Gepaarde observaties: beide variabelen werden bij dezelfde groep onderzoekseenheden gemeten. **IN ORDE**
3. Lineariteit: beide variabelen hebben een lineair verband met elkaar. **NAGAAN VIA SCATTERPLOT**
4. Er werden 30 of meer onderzoekseenheden onderzocht ( $n \geq 30$ ). **IN ORDE**

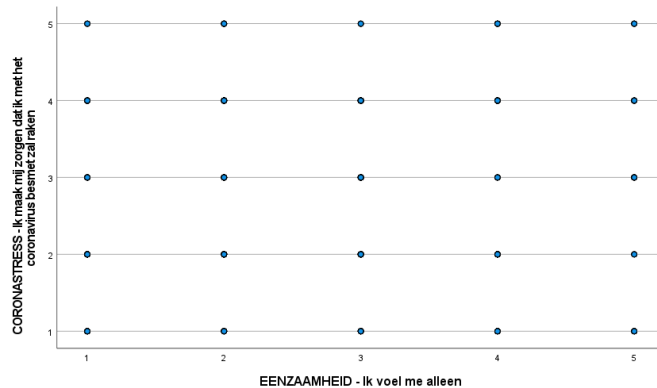
**Assumptie 3.** Lineariteit checken we via een scatterplot van de onderzochte variabelen:

GRAPH

/SCATTERPLOT(BIVAR)=W2\_Eenz3 WITH W2\_Corstress1

/MISSING=LISTWISE.

**Output.** In deze scatterplot merken we geen non-lineair verband op (zoals bv. een parabool), dus mogen we ervan uitgaan dat het verband lineair is. De derde assumptie is bij deze dus vervuld.



**Stap 3.** Hypothesen formuleren.

$H_0$ : Er is geen samenhang tussen eenzaamheid en coronastress: de rangcorrelatiecoëfficiënt tussen die twee variabelen is gelijk aan nul.

$H_a$ : Er is wel een samenhang tussen eenzaamheid en coronastress: de rangcorrelatiecoëfficiënt tussen die twee variabelen is groter of kleiner dan nul.

**Stap 4.** Significantieniveau bepalen.

Het gaat om een steekproef met meer dan 400 onderzoekseenheden ( $n = 511$ ), dus hanteer  $\alpha = .01$  of 1 %.

**Stap 5.** Toetsingsgrootte en p-waarde berekenen.

Doorloop de reeks SPSS-acties in het boek om het volgende commando te verkrijgen:

NONPAR CORR

```
/VARIABLES= W2_Eenz3 W2_Corstress1
/PRINT=SPEARMAN TWOTAIL NOSIG FULL
/MISSING=PAIRWISE.
```

**Output.**

**Correlations**

		W2_Eenz3 EENZAAMHEID - Ik voel me alleen		W2_Corstress1 CORONASTRESS - Ik maak mij zorgen dat ik met het coronavirus besmet zal raken
Spearman's rho	W2_Eenz3 EENZAAMHEID - Ik voel me alleen	Correlation Coefficient	1,000	,036
		Sig. (2-tailed)	.	,414
		N	511	511
	W2_Corstress1 CORONASTRESS - Ik maak mij zorgen dat ik met het coronavirus besmet zal raken	Correlation Coefficient	,036	1,000
		Sig. (2-tailed)	,414	.
		N	511	511

**Stap 6.** Conclusie en rapportering. Vergeet niet de vrijheidsgraden te vermelden bij de correlatiecoëfficiënt ( $n - 2$ ).

“Een Spearman rangcorrelatietoets werd gebruikt om het bivariate verband tussen eenzaamheid en coronastress te onderzoeken. De toets toont aan dat dit verband niet significant is, met  $r_s(509) = .04$ ,  $p = .41$ .”