

# Werkblad Kwantitatieve data analyseren



## Introductie

Je hebt net de video bekeken over het analyseren van kwantitatieve data. Daarin is uitgelegd hoe je je dataset opschooft, de correlatie berekent en kritisch kijkt naar confounders en biases. Na het maken van dit werkblad ben je goed voorbereid op het uitvoeren van je data-analyse.

### Opdracht 1: Dataset opschonen

**A** Schrijf je onderzoeksvraag op.

.....

.....

.....

**B** Zet, indien je dit nog niet gedaan hebt, je data in Excel of in een ander programma dat passend is bij de gekozen methode. Dit helpt je om data gemakkelijker te analyseren.

**C** Bekijk of jouw verzamelde data, je dataset, past bij de onderzoeksvraag die je vooraf wilde beantwoorden. Indien bepaalde data hier niet aan voldoen, verwijder je deze uit de dataset.



## Opdracht 2: Confounder of bias

Geef bij de onderstaande onderzoeksvragen aan of er sprake is van een **confounder** of **bias**.

ONDERZOEKSVRAAG	SITUATIE	CONFOUNDER/ BIAS
Hoeveel inwoners uit Amsterdam maken gebruik van het openbaar vervoer als primair vervoermiddel naar het werk?	Het onderzoek wordt alleen uitgevoerd onder mensen die al bekend staan als frequente openbaar vervoer gebruikers.	.....
Wat zijn de persoonlijke motivaties en verwachtingen van studenten bij de keuze voor een bepaalde studierichting in het hoger onderwijs?	De sociaal emotionele achtergrond van de proefpersonen is niet meegenomen in het onderzoek.	.....
Is er een verband tussen lichaamsbeweging en de mate van stress bij volwassenen in de leeftijdsgroep van 25-40 jaar?	De geselecteerde participanten bestaan voornamelijk uit mensen die al veel bewegen.	.....
Wat zijn de ervaringen van oudere volwassenen met betrekking tot het gebruik van technologie voor communicatie met hun familie en vrienden?	Er werd verondersteld dat alle oudere volwassenen niet bekwaam zouden zijn in het gebruiken van technologische apparaten, maar in de werkelijkheid varieerde dit heel erg.	.....

Ga in je eigen onderzoek ook kritisch na of er sprake is van confounders of biases. Zo ja, noteer dit en beschrijf dit in de discussie van je onderzoek.

.....

.....

.....

### Opdracht 3: Kies het juiste type analyse

**A** Lees de verschillende types analyses hieronder met bijbehorende voorbeelden door.

#### T-toets

Gebruik je wanneer je verschil wilt toetsen. Is er verschil tussen 2 groepen?

De T-toets komt in 3 vormen voor:

**One-sample T-toets:** Toetst of het gemiddelde van een steekproef verschilt van een waarde die je van tevoren zou verwachten.

→ **Voorbeeld onderzoeksvragen:**

1. In hoeverre wijkt de lichaamstemperatuur van een groep patiënten af van de "normale" gemiddelde lichaamstemperatuur van 37 graden Celsius?
2. In hoeverre verschilt het gemiddelde examenresultaat van studenten aan de Radboud Universiteit die een nieuwe leer methode hebben gevolgd, van het algemene gemiddelde examenresultaat van de gehele studentenpopulatie?

**Paired samples T-toets:** Toetst of de gemiddeldes van twee condities die zijn gemeten aan dezelfde proefobjecten (mensen, planten, dieren en cellen), van elkaar verschillen.

→ **Voorbeeld onderzoeksvragen:**

1. Verschilt de reactietijd, nadat mensen water of energydrink drinken? (Dezelfde groep mensen doet de taak na het drinken van water en na het drinken van energydrink).
2. Wat is het effect van een nieuwe leer methode op de prestaties van individuele studenten? (Dezelfde groep studenten wordt getest voor en na de implementatie van de nieuwe methode).

**Independent samples T-toets:** Toetst of de gemiddeldes van twee condities die zijn gemeten aan verschillende proefobjecten (mensen, planten, dieren en cellen), van elkaar verschillen.

→ **Voorbeeld onderzoeksvragen:**

1. Verschilt de reactietijd, nadat mensen water of energydrink drinken? (twee groepen mensen, een groep drinkt water, een groep drinkt energydrink).
2. Wat is het verschil in prestaties tussen twee verschillende groepen studenten die ofwel de gebruikelijke leer methode of een nieuwe leer methode hebben gevolgd, gemeten door hun scores aan het einde van het semester?

#### Correlatie

Gebruik je wanneer je samenhang wilt toetsen. Is er samenhang tussen 2 variabelen?

→ **Voorbeeld onderzoeksvraag:**

1. In hoeverre hangt het percentage spiermassa samen met de mate waarin de sportschool wordt bezocht door mannen tussen de 18 en 30 jaar?  
Wanneer er samenhang is gevonden wil dit **niet** zeggen dat het een **causaal verband** is. Wanneer je meer spiermassa hebt hoeft het niet zo te zijn dat je per se vaker naar de sportschool gaat.

**Wanneer er samenhang is gevonden wil dit niet zeggen dat het een causaal verband is. Wanneer je meer spiermassa hebt hoeft het niet zo te zijn dat je per se vaker naar de sportschool gaat.**



**B** Welk type analyse is passend bij het type onderzoeksvraag? Vul in.

**1. Onderzoeksvraag:** In welke mate verschilt het gemiddelde aantal uren slaap per nacht tussen mensen die regelmatig sporten en mensen die niet regelmatig sporten?

**Omcirkel:** *One-sample T-toets / Paired samples T-toets / Independent samples T-toets / correlatie*

**2. Onderzoeksvraag:** In hoeverre is er een verschil in de gemiddelde hartslag tussen mensen die dagelijks cafeïne-houdende dranken consumeren en mensen die dat niet doen?

**Omcirkel:** *One-sample T-toets / Paired samples T-toets / Independent samples T-toets / correlatie*

**3. Onderzoeksvraag:** Wat is het effect van muziekgenre op de concentratie van studenten tijdens het studeren?

**Omcirkel:** *One-sample T-toets / Paired samples T-toets / Independent samples T-toets / correlatie*

**4. Onderzoeksvraag:** In welke mate verschilt de gemiddelde verkoop van een product gepromoot in twee verschillende reclamecampagnes?

**Omcirkel:** *One-sample T-toets / Paired samples T-toets / Independent samples T-toets / correlatie*

**5. Onderzoeksvraag:** Hoe verhouden de hoeveelheid tijd die mensen besteden aan sociale media en hun zelfgerapporteerde geluksniveau zich tot elkaar?

**Omcirkel:** *One-sample T-toets / Paired samples T-toets / Independent samples T-toets / correlatie*

**6. Onderzoeksvraag:** Wat is het verschil in gemiddelde scores op een wiskundetoets tussen leerlingen die extra bijles hebben gehad en leerlingen die dat niet hebben gehad?

**Omcirkel:** *One-sample T-toets / Paired samples T-toets / Independent samples T-toets / correlatie*

#### Opdracht 4: Kies het juiste type analyse

Kies op basis van bovenstaande informatie welke analyse passend is bij jouw onderzoeksvraag. Vervolgens kun je met je profielwerkstukbegeleider gaan kijken hoe je dit in de praktijk vorm gaat geven. Dit is per onderzoeksvraag net wat anders. Veel informatie is terug te vinden op het internet.

Kies je voor correlatie en wil je daar nog mee oefenen? Ga dan naar opdracht 5.

Kies je voor een T-toets? Ga dan naar opdracht 6.



**Let op!** Er zijn nog meer verschillende analyses, maar in dit lespakket beperken we ons tot correlaties en drie van de vele type t-toetsen. Op het moment dat je een andere analyse wilt uitvoeren, kun je dit het beste overleggen met je profielwerkstukbegeleider.

## Opdracht 5: Oefenen met Correlaties

Gebruik voor deze opdrachten Excel. Vul de gegevens uit de onderstaande tabel in. In kolom 1 vul je de leeftijden in en in kolom 2 vul je de salarissen in. Bereken daarna de correlatie door de volgende formule =CORRELATIE(A2:A8;B2:B8) in te toetsen in de functiebalk (fx).

### Onderzoeksvraag:

"In hoeverre bestaat er een verband tussen de leeftijd en de hoogte van het salaris van dierenartsen tussen de 20-40 jaar?"

De onderzoeksgegevens:

LEEFTIJD	SALARIS
30	52000
35	58000
28	51000
40	65000
22	40000
33	55000
25	45000

Correlatietabel:

CORRELATIE	BETEKENIS
<b>0.00 – 0.30</b>	nauwelijks of geen correlatie
<b>0.30 – 0.50</b>	lage of zwakke correlatie
<b>0.50 – 0.70</b>	middelmatige correlatie
<b>0.70 – 0.90</b>	hoge of sterke correlatie
<b>0.90 – 1.00</b>	zeer hoge of zeer sterke correlatie

**A** Wat is de correlatie?

.....  
.....

**B** Wat is de betekenis van de correlatie?

.....  
.....

**C** Wat zou het betekenen als er een minteken (-) voor de correlatie zou staan?  
Geef hier een voorbeeld bij.

.....  
.....

**D** Doe dit op dezelfde manier voor je eigen onderzoeksgegevens.

## Opdracht 6: Oefenen met de T-toets

Lees de theorie over het significantieniveau.

### Onderzoeksvraag:

*'Hebben leerlingen die 8 uur slapen, een sneller reactievermogen dan leerlingen die slechts 4 uur slaap hebben gehad gedurende meerdere dagen?'*

Je meet de reactietijd van beide groepen met behulp van een online reactietijdtest. En je hebt een significantieniveau van 0,05 vastgesteld. (5%)

- A** Welke soort t-toets past het beste bij deze onderzoeksvraag?

.....

.....

.....

- B** Eerst verzamel je de reactietijden van beide groepen. Deze staan in de tabel onder deze opdracht. Zet de reactietijden in Excel. (groep A in kolom A, groep B in kolom B).

- C** Voer een independent sample t-toets uit om te bepalen of er een significant verschil is in reactietijden tussen de groep die 8 uur slaap heeft gehad en de groep die 4 uur slaap heeft gehad.

→ Dit doe je met behulp van de volgende formule =T.TEST(A1:A5;B1:B5;2;2)

Deze formule vergelijkt de gegevens in kolom A met die in kolom B, waarbij 2 aangeeft dat het om een tweezijdige t-toets gaat. De uitvoer van de formule geeft de kans (p-waarde) dat het verschil in reactietijden tussen de twee groepen puur op toeval berust. Als de p-waarde lager is dan het gekozen significantieniveau (bijvoorbeeld 0,05), kun je concluderen dat er een significant verschil is tussen de twee groepen.

- D** Voer de T-toets uit. Kun je concluderen dat het verschil in slaapduur een significant effect heeft gehad op de reactietijden van de twee groepen en waarom?

.....

.....

.....

**p-waarde  $\leq$  0,05:** Als de p-waarde kleiner is dan of gelijk is aan 0,05, kun je concluderen dat er een significant verschil is tussen de reactietijden van de twee groepen. Met andere woorden, je hebt voldoende bewijs om de nulhypothese te verwerpen en te stellen dat er waarschijnlijk een echt verschil is.

**p-waarde  $>$  0,05:** Als de p-waarde groter is dan 0,05, zou je de nulhypothese niet verwerpen. In dit geval heb je onvoldoende bewijs om te concluderen dat er een significant verschil is tussen de reactietijden van de twee groepen.

GROEP A (8 UUR SLAAP)	GROEP B (4 UUR SLAAP)
12	18
15	20
14	22
13	17
16	21



**Let op!** de betrouwbaarheid van beide opdrachten is niet hoog, omdat de steekproef eigenlijk te klein is. Dit is enkel om te oefenen!

Nu je je data geanalyseerd hebt, kun je verder met de afronding van je verslag. Zorg daarbij dat je goed naar je bronnen verwijst.