

Basisboek Statistiek met SPSS



Noordhoff

**Ben Baarda, Cor van Dijkum
& René van Vianen**

7^e editie

Basisboek Statistiek met SPSS

Ben Baarda
Cor van Dijkum
René van Vianen

Zevende editie
Noordhoff Uitgevers Groningen/Utrecht

Ontwerp omslag: G2K (Groningen-Amsterdam)

Omslagillustratie: Shutterstock - 231007846

Eventuele op- en aanmerkingen over deze of andere uitgaven kunt u richten aan:
Noordhoff Uitgevers bv, Afdeling Hoger onderwijs, Antwoordnummer 13, 9700 VB
Groningen of via het contactformulier op www.mijnnoordhoff.nl.

De informatie in deze uitgave is uitsluitend bedoeld als algemene informatie. Aan deze informatie kunt u geen rechten of aansprakelijkheid van de auteur(s), redactie of uitgever ontlenen.



Wij maken bij Noordhoff sinds 2009 bewust keuzes om onze impact op het klimaat te verkleinen. We volgen de klimaatdoelen van het Klimaatakkoord van Parijs. Scan de QR-code voor meer informatie over ons duurzaamheidsbeleid.

Zo is dit product gemaakt van FSC®-gecertificeerd papier en ander gecontroleerd materiaal. Wij gebruiken plantaardige inkt en onze lijmen zijn op waterbasis en afbreekbaar. Hierdoor is dit product goed te recyclen na gebruik.



0 / 26

© 2026 Baarda, Van Vianen & Van Dijkum, vertegenwoordigd door Noordhoff Uitgevers bv Groningen/Utrecht, The Netherlands.

Deze uitgave is beschermd op grond van het auteursrecht. Wanneer u (her)gebruik wilt maken van de informatie in deze uitgave, dient u vooraf schriftelijke toestemming te verkrijgen van Noordhoff Uitgevers bv. Meer informatie over collectieve regelingen voor het onderwijs is te vinden op www.onderwijsauteursrecht.nl.

This publication is protected by copyright. Prior written permission of Noordhoff Uitgevers bv is required to (re)use the information in this publication.

ISBN(ebook) 978-90-01-07996-3

ISBN 978-90-01-07995-6

NUR 916

Voorwoord bij de zevende editie

In deze nieuwe versie van het *Basisboek Statistiek met SPSS* vind je aanwijzingen voor het verwerken en analyseren van onderzoeksgegevens met behulp van SPSS. Oorspronkelijk staat SPSS voor 'Statistical Products for the Social Sciences'. Sinds SPSS in 2009 een onderdeel van IBM Business Analytics is geworden, kent het een veel bredere toepassing en wordt het niet alleen door sociale wetenschappers, maar ook door andere onderzoekers gebruikt zoals economen, marktonderzoekers, medische wetenschappers en technici.

Een jaarlijkse licentie SPSS is voor studenten en medewerkers in het hoger onderwijs via SURFSPOT tegen gereduceerd tarief te verkrijgen, of maakt soms onderdeel uit van de softwarepakketten die het hoger onderwijs aanbiedt. Bij het schrijven van dit boek hebben wij gebruikgemaakt van SPSS versie 30.0. De eerdere versies van SPSS verschillen in de basis nauwelijks van de meest recente versie. De in dit boek beschreven handelingen vormen de basis en zijn navolgbaar voor eerdere versies.

In de nieuwe editie van deze SPSS-handleiding zijn een aantal verbeteringen aangebracht:

- Alle instructies voor het verwerken en analyseren van gegevens met behulp van SPSS staan stapsgewijs uitgelegd, met de stappen links in beeld en de schermafbeeldingen van de instructie rechts.
- Bij een aantal handelingen verwijzen QR-codes naar instructievideo's, waarbij de stappen nogmaals uitgelegd worden.
- Elk hoofdstuk bevat leerdoelen voor dat betreffende hoofdstuk. De hoofdstukken bevatten opdrachten waar voor de meeste opdrachten een koppeling wordt gelegd naar de data van de student (of docent).
- In hoofdstuk 2 is aandacht besteed aan het exporteren van gegevens van een online vragenlijst naar SPSS en het kunnen omzetten van tekstuele antwoorden naar waarden (getallen). Dit sluit aan bij de trend om gegevens te verzamelen via online vragenlijsten.
- De hoofdstukken 4, 5 en 6 waarin de analyse van de data centraal staat, hebben een inleiding gekregen met een overzicht waarin je kan zien bij welke meetniveaus je welke technieken kunt gebruiken.

In de vorige editie van het boek is gewerkt aan de hand van een voorbeeld 'Geld, Geluk en Gezondheid'. Dit voorbeeld is gehandhaafd voor deze versie.

Met dit boek willen we je al doende leren vertrouwd te raken met de statistische procedures in SPSS, en ook met statistiek. We nodigen je uit om aan de hand van een databestand over de relatie tussen geld, geluk en gezondheid aan de slag te gaan.

Na een algemene inleiding en kennismaking met de basisregels van SPSS, kun je direct beginnen met het invoeren en analyseren van gegevens in de computer. Om verantwoord statistische procedures met SPSS te kunnen uitvoeren, is niet alleen kennis van SPSS, maar ook statistische kennis nodig. We bespreken in de eerste hoofdstukken statistische basisbegrippen als meetniveau, normaalverdeling, kans, significantie, een- en tweezijdige toetsing en effectgrootte. In de laatste hoofdstukken bespreken we technieken die je gebruikt om gegevens te beschrijven, om verschillen te toetsen en om samenhang na te gaan.

We leggen zo veel mogelijk in woorden en met voorbeelden de essentie uit van iedere techniek die we bespreken, zonder daar diep wiskundig op in te gaan. Ook geven we aan wanneer je die techniek wel kunt gebruiken en wanneer niet. Uiteraard wordt uitgelegd hoe je die techniek in SPSS uitvoert. Via schermafdrucken en stapsgewijze instructies die we daarbij geven, laten wij duidelijk zien wat je moet doen. Voor iedere besproken techniek laten we een voorbeeld van de uitvoer zien. Aan de hand hiervan leggen we uit hoe je het moet lezen en wat het betekent. Ook geven we aan hoe je over het resultaat kunt rapporteren.

De databestanden waarop we de analyses uitvoeren, zijn te vinden op internet via studiemeister.noordhoff.nl. Deze databestanden zijn nieuw, er is onder andere een gezondheidsvariabele toegevoegd.

In deze handleiding werken we met drie informatieve bijlagen:

1. Het controleren van de homogeniteit van samengestelde scores.
2. De samenhang tussen meerdere nominale variabelen.
3. De samenhang tussen meerdere onafhankelijke variabelen met één afhankelijke variabele.

Het *Basisboek Statistiek met SPSS* komt optimaal tot zijn recht in samenhang met *Dit is onderzoek!*.

Voorjaar 2026
Ben Baarda
Cor van Dijkum
René van Vianen

Inhoud

- 1 Hoe bereid ik mij voor op statistiek met SPSS? 11**
 - 1.1 Analyseren van gegevens in de onderzoekscyclus 12
 - 1.2 Geld, Geluk en Gezondheid; toelichting op de gebruikte casus 13
 - 1.3 Hoe analyseer ik mijn data? Een gebruikswijzer 16
 - 1.4 Enkele algemene statistische begrippen 21
 - 1.5 Hoe werkt SPSS onder Windows? 24
 - Samenvatting 32

- 2 Hoe breng ik mijn gegevens in? 35**
 - 2.1 Hoe voer ik mijn gegevens in? De datamatrix 36
 - 2.2 Werken met identificatienummers of respondentnummers 45
 - 2.3 Hoe kan ik met de Variabele View werken? 48
 - 2.4 Hoe hercodeer ik tekstuele waarden? 59
 - 2.5 Hoe bewaar ik mijn ingevoerde gegevens? 63
 - Samenvatting 65

- 3 Hoe verander en combineer ik gegevens? 67**
 - 3.1 Enkele handige tips 68
 - 3.2 Hoe hercodeer ik gegevens? 70
 - 3.3 Hoe deel ik een variabele in in klassen? 75
 - 3.4 Hoe combineer ik mijn gegevens? 78
 - 3.5 Hoe houd ik bij het combineren van gegevens rekening met ontbrekende gegevens? 82
 - Samenvatting 89

- 4 Hoe analyseer ik mijn gegevens bij een frequentie-onderzoeksvraag 91**
 - 4.1 Hoe beschrijf ik gegevens bij frequenties? 92
 - 4.2 Hoe maak ik een frequentieverdeling met SPSS? 95
 - 4.3 Hoe geef ik frequenties grafisch weer? 98
 - 4.4 Hoe vergelijk ik (sub)groepen? 108
 - 4.5 Wat doe ik met multiple responsen? 117
 - 4.6 Hoe maak ik variabelen vergelijkbaar? 120
 - Samenvatting 123

5	Hoe analyseer ik mijn gegevens bij een verschil-onderzoeksvraag?	125
5.1	Wat is een verschilvraag?	126
5.2	Verschilvraag bij een nominale test- en splitsingsvariabele. Wat is een kruistabel en chikwadraat?	129
5.3	Verschilvraag bij ordinale testvariabelen en nominale splitsingsvariabelen. Wat zijn niet-parametrische toetsen?	139
5.4	Verschilvraag bij interval/ratio testvariabelen en nominale splitsingsvariabele; t-toets of variantieanalyse?	155
	Samenvatting	173
6	Hoe analyseer ik een vraag over samenhang?	175
6.1	Wat is een samenhangonderzoeksvraag?	176
6.2	Samenhang tussen twee nominale variabelen: Cramér's V	178
6.3	Samenhang tussen twee ordinale variabelen: Spearman's rangcorrelatie!	182
6.4	Samenhang tussen twee interval- of ratiovariabelen: Pearson's productmoment-correlatiecoëfficiënt!	186
6.5	Samenhang tussen twee interval- of ratiovariabelen: enkelvoudige regressieanalyse	194
6.6	Samenhang bij twee of meer variabelen op interval- of rationiveau: multiple correlatie- en regressieanalyse	200
6.7	Samenhang tussen twee interval- of ratiovariabelen, gecontroleerd voor het effect een derde variabele: partiële correlatie	203
	Samenvatting	208
	Bijlage I Hoe controleer ik de homogeniteit van samengestelde scores?	210
	Bijlage II De samenhang tussen meerdere nominale variabelen	217
	Bijlage III De samenhang van meerdere onafhankelijke variabelen met één afhankelijke variabele	225
	Illustratieverantwoording	231
	Register	232
	Over de auteurs	235

For I don't care too much for money, money can't buy me love

John Lennon & Paul McCartney

Voorkennis

Er wordt geen specifieke voorkennis vereist. Wel wordt aanbevolen om naast dit boek het *Basisboek Methoden en Technieken* (7e druk, 2021) of *Dit is onderzoek!* (4e druk, 2024) te gebruiken.



Hoe bereid ik mij voor op statistiek met SPSS?

Leerdoelen of -uitkomsten:

- Je weet in welke fase van jouw onderzoek je SPSS kunt inzetten met als uiteindelijk doel het analyseren van gegevens.
- Je weet het verschil tussen de verschillende meetniveaus en je kunt de consequenties van de verschillende meetniveaus voor je analyse benoemen.
- Je weet wat een normaalverdeling is en kent een aantal belangrijke statistische basisbegrippen, zoals standaardfout, zekerheid of kans, significantie en relevantie, effectgrootte en vrijheidsgraden.
- Je weet op welke wijze SPSS werkt en hoe je gebruikmaakt van de statistische functies.

1

- 1.1 Analyseren van gegevens in de onderzoekscyclus 12
- 1.2 Geld, Geluk en Gezondheid: toelichting op de gebruikte casus 13
- 1.3 Hoe analyseer ik mijn data? Een gebruikswijzer 16
- 1.4 Enkele algemene statistische begrippen 21
- 1.5 Hoe werkt SPSS onder Windows? 24

1.1 Analyseren van gegevens in de onderzoekscyclus

Om antwoord te geven op je deelvragen heb je voor je onderzoek kwantitatieve data verzameld door vragenlijsten af te nemen, observatieformulieren te gebruiken, bestaande data als kassabestanden te benutten, dossieranalyses te doen of uitslagen van toetsen of tentamens in te voeren. Het resultaat van de dataverzameling tijdens een onderzoek is meestal een berg gegevens. Dit zijn allemaal data die je kunt gebruiken voor het analyseren van onderzoeksgegevens. Als je onderzoek doet, is het kiezen van de juiste statistische techniek om de verzamelde gegevens te analyseren een belangrijke schakel in de lange keten van beslissingen die je moet nemen. Het uiteindelijke doel is het beantwoorden van de onderzoeksvraag of -vragen.

Om een duidelijk beeld te krijgen van de rol van data-analyse binnen de gehele onderzoekscyclus, staan hieronder de gebruikelijke fasen van een onderzoek voor je op een rij. Elke fase van de onderzoekscyclus is opgenomen in de vorm van een vraag:

- 1 Wat is/zijn mijn onderzoeksvra(a)g(en) en wat is de doelstelling van mijn onderzoek?
- 2 Hoe zoek ik informatie (onder meer literatuurstudie)?
- 3 Wat voor type onderzoek ga ik doen?
- 4 Hoe ziet mijn onderzoeksontwerp eruit?
- 5 Betrek ik de populatie in mijn onderzoek, of trek ik een steekproef?
- 6 Welke dataverzamelmethode ga ik gebruiken?
- 7 *Hoe prepareer ik mijn data voor de analyse?*
- 8 *Hoe analyseer ik mijn data?*
- 9 *Hoe rapporteer en evalueer ik mijn onderzoek?*
- 10 Welke conclusies kan ik op basis van het onderzoek te trekken?
- 11 Welke aanbevelingen kan ik formuleren?

In dit boek staan de fasen 7 en 8 en een deel van fase 9 centraal: preparatie, analyse en beschrijving van de onderzoeksgegevens die je met behulp van SPSS kunt analyseren. Deze analyse levert de basis voor de conclusies.

Aan de hand van een onderzoek naar de relatie tussen geld, geluk en gezondheid (zie casus 1.1), krijg je stap voor stap antwoord op de volgende vragen:

- Hoe prepareer je onderzoeksgegevens om deze te kunnen invoeren in SPSS?
- Hoe kun je gegevens met SPSS aanpassen en veranderen? Voordat je aan de analyse begint, moet je vaak eerst van bepaalde gegevens de waarden ompolen, oftewel hercoderen, of je moet de waarden van gegevens combineren tot een nieuwe score en die vervolgens controleren op hun betrouwbaarheid.
- Hoe kies je de juiste analysetechniek? Om de juiste analysetechniek te kiezen, stel je eerst vast wat het karakter is van je onderzoeksvraag. Gaat

het om frequenties, om verschillen of om samenhangen? Vervolgens ga je na wat het meetniveau van je gegevens is. Tot slot stel je vast of het in je onderzoek om een steekproef of een populatie gaat. Je kunt dan aan de hand van het blokschema 'Hoe analyseer ik mijn data?' vaststellen welke analysetechniek het best past bij jouw onderzoeksvraag. Je vindt het schema voor in dit boek.

- Hoe voer je de analyse met SPSS uit en hoe interpreteer je de resultaten? Bij iedere statistische techniek is aangegeven hoe je dit doet. Ook zie je hoe je verslag kunt doen van de uitdraaien van SPSS.

In dit hoofdstuk leer je een aantal belangrijke statistische begrippen, zoals normaalverdeling, significantie, kans en standaardfout. Ook zie je hoe je SPSS kunt starten.

1.2 Geld, Geluk en Gezondheid; toelichting op de gebruikte casus

Bij de meeste paragrafen (en analysetechnieken) maak je gebruik van het fictieve voorbeeldonderzoek 'Geld, Geluk en Gezondheid' (zie casus 1.1). De vragen uit casus 1.1 zijn voorgelegd aan een representatieve steekproef van 490 Nederlandse vrouwen en 500 Nederlandse mannen van 25 tot en met 60 jaar. Er is bewust voor deze leeftijdsgrenzen gekozen. Veel jonge mensen studeren nog en hebben daardoor geen vast inkomen. Bij mensen boven de 60 is er vaak al sprake van (gedeeltelijke) uittreding uit het arbeidsproces, waardoor zij in een andere financiële situatie komen. De data van dit onderzoek vind je op studiemeister.noordhoff.nl onder de naam 'data_geld_geluk_gezondheid'.

CASUS 1.1 | Voorbeeldonderzoek Geld, Geluk en Gezondheid

Ja, geld maakt gelukkig

Geld. In de aanloop naar de verkiezingen gaat het er vaak over. De meeste mensen willen er graag wat meer van. Maar ben je ook tevreden als je uiteindelijk meer euro's hebt? Maakt geld gelukkig?

Alle tegeltjeswijsheden ten spijt, de grote lijn in het onderzoek luidt: ja, geld maakt gelukkig.

Meer inkomen maakt mensen gelukkiger. Meer vermogen en minder schulden doen dat ook. Belangrijke kanttekening: grote blijdschap over een loonsverhoging of een geldprijs houdt niet lang aan. Maar onze gevoelens van geluk en welbevinden

worden wel degelijk beïnvloed door onze financiële situatie op de langere termijn.

Extra blij

De relatie tussen geluk en geld is best complex. Een voorbeeld: mensen die al heel gelukkig zijn, reageren anders op een stijging van hun inkomsten dan ongelukkige mensen. Dat blijkt uit recent onderzoek van psychologen Matthew Killingsworth, Daniel Kahneman en Barbara Mellers onder tienduizenden Amerikanen.

Ongelukkige mensen worden wel wat gelukkiger van meer inkomen. Maar de toename

van hun geluk vlakt af bij een bruto jaarinkomen tussen de 60.000 en 90.000 dollar. En komt tot stilstand bij 100.000. Opmerkelijk genoeg blijken de meest gelukkige mensen juist blij met elke extra dollar. Ook die boven de 100.000.

(Nuttig weetje: volgens de OESO was in 2022 het gemiddelde brutojaarinkomen in de VS 77.463 dollar. In Nederland was dat, gerekend in dollars: 63.225.)

Zelfverdiend maakt gelukkiger

Allerlei factoren bepalen hoe blij je wordt van geld. Ik noem er twee.

- Wonen in een rijk land maakt mensen gelukkiger. Geluksonderzoeker Edward Diener en zijn collega's stelden vast dat een hoger nationaal inkomen leidt tot een positievere evaluatie van je leven. Dat geldt zowel voor mensen met een hoger als met een lager individueel inkomen.
- Het maakt uit hoe je aan je geld komt. Experimenteel onderzoek van de Chinese psychologen Bihui Jin en Jing Li laat zien dat studenten gelukkiger worden wanneer ze geld uitgeven dat ze zelf hebben verdiend dan wanneer ze geld uitgeven dat ze van hun ouders krijgen.

Volgens de onderzoekers komt zelfverdiend geld tegemoet aan twee belangrijke menselijke behoeftes: autonomie en competentie.

Echt geluk?

Zit écht geluk niet in heel andere dingen dan geld? Ons antwoord op die vraag blijkt deels afhankelijk van ons inkomen.

Onderzoeker Rhia Catapano en haar collega's ploegden door de data van ruim een half miljoen mensen in meer dan honderd landen. Hun conclusie: mensen die meer te besteden hebben, zijn doorgaans iets gelukkiger. Maar mensen met minder geld ontlenen hun geluk in grotere mate aan een gevoel van betekenis: de overtuiging dat hun leven een doel, waarde en richting heeft. En dat leidt dan weer tot meer welzijn op de lange termijn.

Ben je rijk en eigenlijk een beetje jaloers op dat gevoel van betekenis? Dan helpt het, volgens de onderzoekers, om geld te geven aan goede doelen en dingen te kopen voor anderen.

Grappig. Echt, duurzaam geluk blijkt te koop, zolang je je geld niet aan jezelf besteedt.

Bron: NRC, rubriek Zo werkt dat, Ben Tichelaar, 2 november 2023

Maakt geld gelukkig en gezond?

Een onderzoeker wil weten of er een verband bestaat tussen geld, geluk en gezondheid. Hij vraagt zich af of geld gelukkig en gezond maakt. Zijn centrale onderzoeksvraag luidt dan ook: 'Is er een verband tussen de mate waarin iemand over geld beschikt en de mate waarin hij/zij zich gelukkig en gezond voelt?'

Het begrip 'geld' heeft de onderzoeker als volgt gedefinieerd: geld is de hoeveelheid financiële middelen waarover iemand kan beschikken. Deze definitie is met opzet ruim gekozen. Hierdoor worden niet alleen het inkomen en het vermogen, maar ook andere financiële bronnen waarover iemand kan beschikken in het onderzoek betrokken.

Geluk wordt als volgt gedefinieerd: geluk is de mate waarin iemand tevreden is met het leven dat hij/zij leidt. Gezondheid wordt gemeten met de vraag hoe gezond iemand zichzelf voelt.

De onderzoeker heeft deze begrippen meetbaar gemaakt, door ze om te zetten in vragen. Dit wordt operationaliseren genoemd. Zowel voor het meten van het begrip geld, als voor het begrip geluk, heeft hij vijf uitspraken of items gemaakt. Het begrip gezondheid wordt gemeten met één vraag.

Geld is gemeten met de items:

1 Ik kan een nieuwe elektrische auto aanschaffen.	ja/nee
2 Ik woon in een koopwoning.	ja/nee
3 Ik bezit meer dan € 50.000 aan eigen vermogen.	ja/nee
4 Ik krijg zorgtoeslag.	ja/nee
5 Ik heb geen geld om op vakantie te gaan.	ja/nee

Geluk is gemeten met de volgende vragen:

1 Als ik mijn leven over zou mogen doen, dan zou ik het op dezelfde manier doen.

absoluut niet niet ten dele
wel/niet wel absoluut wel

2 De meeste anderen hebben het beter dan ik.

absoluut niet niet ten dele
wel/niet wel absoluut wel

3 Ik heb het naar mijn zin.

absoluut niet niet ten dele
wel/niet wel absoluut wel

4 Het leven is zwaar.

absoluut niet niet ten dele
wel/niet wel absoluut wel

5 Ik voel mij eenzaam.

absoluut niet niet ten dele
wel/niet wel absoluut wel

Gezondheid is gemeten met de volgende vraag:

Hoe is uw gezondheid?

Zeer goed Goed Voldoende Slecht Zeer slecht

Omdat geluk niet alleen van financiële middelen afhangt, maar ook van andere zaken, vraagt de onderzoeker ook naar een aantal gemakkelijk te meten kenmerken, zoals geslacht, leeftijd, opleidingsniveau en leefsituatie van de personen die hij enquêteert.

<i>Geslacht</i>	<input type="checkbox"/> vrouw <input type="checkbox"/> man <input type="checkbox"/> non-binair
<i>Leeftijd in jaren</i>	...
<i>Leefsituatie</i>	<input type="checkbox"/> alleen <input type="checkbox"/> met partner <input type="checkbox"/> met partner en kind(eren)
<i>Hoogste afgeronde opleiding</i>	<input type="checkbox"/> lager (beroeps)onderwijs (lagere school, lager technisch onderwijs, lager voortgezet onderwijs) <input type="checkbox"/> middelbaar (beroeps)onderwijs (middelbaar technisch onderwijs, middelbaar voortgezet onderwijs, mavo, mbo, vmbo, enz.) <input type="checkbox"/> hoger (beroeps)onderwijs (universiteit, hbo, vwo, havo)

OPDRACHT

- Met welke concepten (zoals geld, geluk of gezondheid) werk jij binnen jouw onderzoeksvragen?
- Op welke wijze heb je deze concepten geoperationaliseerd (in één of meer vragen)?
- Welke gegevens heb je (aanvullend) verzameld (zoals geslacht en leeftijd) om eventuele uitspraken te kunnen doen tussen deze kenmerken en uit je onderzoek afkomstige concepten?

1.3 Hoe analyseer ik mijn data? Een gebruikswijzer

Voor de keuze van een statistische analysetechniek zijn de antwoorden op de volgende vragen van belang:

- Gaat het in de vraagstelling om frequenties (hoe vaak/in welke mate zijn mensen gelukkig?), om een verschil (zijn vrouwen gelukkiger dan mannen?) of om een samenhang (is er een relatie tussen leeftijd en gezondheid)? Of gaat het om een combinatie daarvan?
- Wat is het meetniveau (nominaal, ordinaal of interval-/rationiveau) van de gegevens die je hebt verzameld?
- Gaat het om een steekproef of om een populatie?

Het blokschema (te vinden voorin dit boek) is ontworpen aan de hand van deze vragen. In de volgende subparagrafen gaan we hierop in.

1.3.1 Om wat voor specifieke onderzoeksvragen gaat het in mijn onderzoek?

Bij het beantwoorden van de vraag welke statistische techniek je kunt gebruiken, vormt de *onderzoeksvraag* of vormen de *onderzoeksvragen* het uitgangspunt.

Het onderzoek omvat altijd één of meer specifieke onderzoeksvragen waarop een antwoord moet worden gegeven. Globaal zijn er drie soorten onderzoeksvragen te onderscheiden:

Frequentie	1 Vragen waarbij het erom gaat hoe vaak of in welke mate iets voorkomt. Een voorbeeld daarvan is: 'In welke mate zijn Nederlanders gelukkig?' of 'Hoeveel procent van de Nederlanders is in het bezit van een auto?'
Verskil	2 Vragen waarbij het gaat om een verschil. Bijvoorbeeld: 'Zijn vrouwen gelukkiger dan mannen?'
Samenhang	3 Vragen waarbij het gaat om een samenhang. Bijvoorbeeld: 'Is er een samenhang tussen leeftijd en gezondheid?'

Hoofdstuk 4 behandelt een voorbeeld van de analyse van gegevens bij frequentievraagstellingen. De analyse van gegevens bij de verschillende vragen wordt behandeld in hoofdstuk 5 en in hoofdstuk 6 vind je voorbeelden van data-analyse bij een samenhangonderzoeksvraag.

OPDRACHT

- Met welke specifieke onderzoeksvragen heb jij te maken?
- Geef aan of het gaat om vragen over frequenties, verschillen of samenhangen?

1.3.2 Wat is het meetniveau van mijn gegevens?

Als je hebt bepaald om welk type onderzoeksvragen het in je onderzoek gaat (zie de eerste kolom in het blokschema), dan moet je vervolgens nagaan op welk meetniveau de variabele(n) is/zijn gemeten. Zie daarvoor in het genoemde blokschema de cellen onder frequentie, verschil of samenhang of scan de code voor een uitleg.



Per onderzoeksvraag moet je aangeven wat het meetniveau van de betreffende variabelen is. Bij de onderzoeksvraag naar het verschil tussen vrouwen en mannen in de mate waarin ze zich gelukkig voelen, is het meetniveau van de betreffende variabelen namelijk anders (en lager) dan het meetniveau van de variabelen in bijvoorbeeld de onderzoeksvraag naar de samenhang tussen leeftijd en gezondheid. Bij de variabele *geslacht* zijn er maar drie categorieën of waarden, namelijk 'man', 'vrouw' of 'non-binair'. Hierbij is alleen sprake van een verschil. Een man is anders dan een vrouw of een non-binair persoon, maar niet meer of minder. Hetzelfde geldt voor *leefsituatie*: iemand is gehuwd, samenwonend of alleenstaand. Bij dit type

Nominaal
meetniveau

antwoordmogelijkheden gaat het om een *nominaal meetniveau*. Je kunt deze gegevens tellen (hoeveel respondenten wonen samen, zijn gehuwd of zijn alleenstaand), maar je kunt geen onderscheid maken. Er is geen ordening mogelijk tussen samenwonend, gehuwd of alleenstaand. Het een is niet meer of minder dan het ander.

Ordinaal
meetniveau

Ordering kan wel bij gegevens die zijn gemeten op respectievelijk ordinaal, interval- of rationiveau. Bij gegevens op *ordinaal meetniveau* is wél sprake van meer of minder, maar het verschil tussen de categorieën is niet in een getal uit te drukken. Bij de variabele *opleidingsniveau* bijvoorbeeld is er duidelijk sprake van meer en minder. De havo is hoger dan het vmbo, maar er is niet aan te geven hoeveel hoger. Dat geldt ook voor de medailleverdeling op een kampioenschap. Het is duidelijk dat bijvoorbeeld een 100-meterloper die goud wint sneller heeft gelopen dan een looper die zilver heeft gewonnen. Het feit dat hij goud heeft gewonnen, geeft aan dat hij sneller heeft gelopen, maar niet hoeveel sneller.

Intervalniveau
Nulpunt
Rationiveau

Bij interval- en ratiomeetniveau is dat verschil tussen categorieën in termen van meer of minder wel in een getal uit te drukken. Temperatuur is daarvan een goed voorbeeld. Het verschil tussen 5 en 10 graden Celsius is even groot als het verschil tussen 25 en 30 graden. Bij het *intervalniveau* is er alleen geen sprake van een natuurlijk *nulpunt*, zoals wel het geval is bij gegevens die op *rationiveau* zijn gemeten, bijvoorbeeld gewicht en lengte. Nul graden Celsius is geen natuurlijk nulpunt. Het natuurlijke nulpunt voor de temperatuur is -273 graden Celsius, wat overeenkomt met nul graden Kelvin. Wanneer je de temperatuur weergeeft in graden Kelvin, is er wel sprake van een ratiometing, want hier is sprake van een natuurlijk nulpunt. Figuur 1.1 is een illustratie van verschillende meetniveaus.

Figuur 1.1
Verschillende
meetniveaus



Ik ben Iwan:

- *professioneel gewichtheffer (nominaal)*
- *vrijgezel (nominaal)*
- *ik ben tweede geworden op het wereldkampioenschap (ordinaal)*
- *mijn IQ is 110 (interval)*
- *ik weeg 140 kilo (ratio)*

Het meetniveau heeft gevolgen voor de rekenkundige bewerkingen die je mag uitvoeren. Bij temperatuur gemeten in graden Celsius mag je niet zeggen dat 20 graden tweemaal zoveel is als 10 graden. Bij een meting in graden Kelvin mag je wel zeggen dat 20 graden tweemaal zoveel is als 10 graden en bij een gewichtsmeting dat 20 kilo tweemaal zo zwaar is als 10 kilo.

Het voorgaande wordt schematisch weergegeven in tabel 1.1.

Tabel 1.1 Overzicht van meetniveaus, hun rekenkundige consequenties en voorbeelden

Meetniveau	Rekenkundige consequentie	Voorbeeld
Nominaal	Tellen, percentages (alleen onderscheid)	Geslacht
Ordinaal	Tellen, percentages en hoger/lager (onderscheid en ordening)	Opleidingsniveau
Interval	Tellen, hoger/lager, waarbij verschillen in eenheden zijn uit te drukken, gemiddelde, spreiding (onderscheid en ordening)	Intelligentie
Ratio	Tellen, hoger/lager, waarbij verschillen in eenheden zijn uit te drukken, gemiddelde, spreiding en het berekenen van verhoudingen (onderscheid en ordening)	Leeftijd

Continue variabelen

We onderscheiden verder continue en discontinue of discrete variabelen. Bij *continue variabelen* kun je je een lijn voorstellen waarop waarden een aaneengesloten rij punten vormen: een continuüm. Tussen twee punten liggen altijd nog (oneindig veel) andere mogelijke waarden. Voorbeelden van continue variabelen zijn de lengte en het gewicht van een persoon. Variabelen die alleen hele waarden kunnen aannemen, noem je *discrete variabelen*, zoals het aantal auto's dat iemand bezit, of het aantal kinderen in een gezin.

Discrete variabelen

OPDRACHT

- Vul tabel 1.1 in met jouw gegevens. Welke variabelen heb jij gemeten (in de laatste kolom) en welk meetniveau hebben jouw variabelen?
- Check deze meetniveaus aan de hand van de wijze waarop jij de variabelen geoperationaliseerd hebt.

1.3.3 Gaat het om een populatie of om een steekproef?

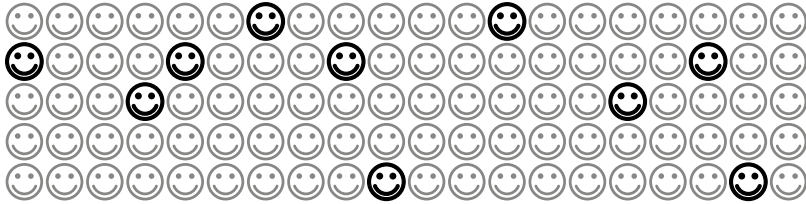
Beschrijvende statistiek

Er zijn twee vormen van statistiek: de beschrijvende en de inductieve statistiek. *Beschrijvende statistiek* gebruik je wanneer je onderzoek doet bij een *populatie*. Er is sprake van een populatie wanneer alle eenheden waarover je uitspraken wilt doen in je onderzoek worden betrokken. Dus

wanneer je bijvoorbeeld alle werknemers van een bedrijf enquêteert om hun arbeidssatisfactie vast te stellen. Om kosten te besparen, kun je ook een deel van de werknemers enquêteren (*steekproef*) die je aselect uit het totale bestand van werknemers trekt (zie figuur 1.2).

Figuur 1.2

Voorbeeld van een populatie en een aselecte steekproef



N=100 hele populatie van werknemers

n=10 aselect getrokken steekproef uit populatie van werknemers

Inductieve/inferentiële statistiek

Bij een steekproef is het de bedoeling dat je uitspraken doet over de totale populatie. Denk bijvoorbeeld aan verkiezingsonderzoek waarbij een onderzoeksbureau niet alle stemgerechtigden ondervraagt, maar slechts een steekproef vraagt op wie ze gaan stemmen. In dit geval moet je gebruikmaken van de *inductieve of inferentiële statistiek*. Een steekproef is namelijk altijd enigszins afhankelijk van toeval. Dit hangt onder andere af van de grootte van de steekproef. Je wilt immers op basis van een specifiek geval (een steekproef) algemene uitspraken doen (over de populatie). Raadpleeg een methoden- of statistiekboek, zoals het *Basisboek Methoden en Technieken*, voor meer informatie. Ook op Wikipedia vind je uitgebreide informatie over het trekken van verschillende soorten steekproeven (zoek bijvoorbeeld op *Sampling (statistics)*).

Voordat je begint met het analyseren van je gegevens, moet je jezelf de vraag stellen over welke eenheden (wie of wat) je uitspraken wilt doen. Als je alleen uitspraken wilt doen over de personen of zaken zijn die betrokken zijn in je onderzoek, dan is er sprake van een *populatieonderzoek*. Wil je ook uitspraken doen over de personen of zaken die niet betrokken zijn bij je onderzoek, maar wel vertegenwoordigd worden door de onderzoekseenheden die je geselecteerd hebt, dan gaat het om een *steekproefonderzoek*. In de paragraaf hierna vind je een beknopte uitleg van enkele statistische termen die je regelmatig tegenkomt bij het toetsen of de resultaten. Deze termen helpen je te bepalen of je bevindingen op toeval berusten of, met een bepaalde marge aan onzekerheid, kunnen worden gegeneraliseerd naar de populatie waaruit de steekproef is getrokken.

OPDRACHT

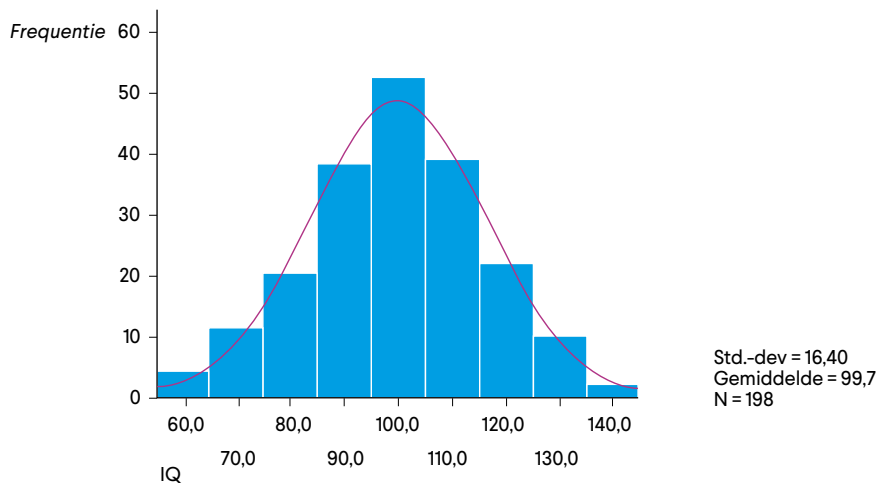
- Heb jij je gegevens verzameld bij een populatie of heb je een steekproef getrokken?
- Is er bij jouw onderzoek dan sprake van inductieve of beschrijvende statistiek?

1.4 Enkele algemene statistische begrippen

In deze paragraaf komen een aantal belangrijke statistische begrippen aan bod, zoals normaalverdeling, standaardfout, zekerheid of kans, significantie, relevantie, effectgrootte en vrijheidsgraden. Het doel van de beschrijvende statistiek is om je gegevens op een inzichtelijke en overzichtelijke manier te presenteren. Als je bijvoorbeeld de arbeidssatisfactie van bijna duizend werknemers hebt gemeten, heeft het weinig zin al die gegevens los te presenteren. Meestal vat je ze samen in bijvoorbeeld een histogram of in de vorm van percentages of een gemiddelde. Door je gegevens op een gereduceerde manier weer te geven, creëer je overzicht. Als je gegevens grafisch weergeeft, resulteert dit vaak in een zogenaemde *normaalverdeling*. In figuur 1.3 zie je hiervan een (fictief) voorbeeld.

Normaalverdeling

Figuur 1.3
Het intelligentieniveau van de 198 werknemers van de firma Arbeid



In figuur 1.3 zie je de verdeling van de scores op een intelligentietest, afgenomen bij 198 werknemers van de firma 'Arbeid'. Deze verdeling lijkt sterk op een normaalverdeling. Ter vergelijking is de normaalverdeling in figuur 1.3 ingetekend. Deze verdeling noem je de *verdeling van Gauss* of *Gausskromme*. Het kenmerkende van de normaalverdeling is de klokvorm en de symmetrie tussen de linker- en rechterhelft.

Wanneer de 198 werknemers een aselechte steekproef vormen uit het totale bestand van werknemers van de firma 'Arbeid' (N = 2213), moet je gebruikmaken van de inductieve of inferentiële statistiek. De vraag is dan in hoeverre het gevonden gemiddelde IQ van 99,7 representatief is voor de totale populatie van werknemers. Met andere woorden: wat is de kans op een gemiddeld IQ van 99,7 in de populatie, als je daadwerkelijk de gehele populatie in je onderzoek zou kunnen betrekken? Die kans is uiteraard niet zo groot, omdat het gemiddelde dat je in de steekproef hebt gevonden, afhankelijk is van de toevallige samenstelling van deze steekproef en bij elke steekproeftrekking iets kan verschillen.

Wanneer je opnieuw een aselechte steekproef trekt en nog een keer (enzovoort), zal het gemiddelde IQ waarschijnlijk steeds iets verschillen, soms iets hoger en soms iets lager. De gevonden waarden zullen van elkaar afwijken, maar waarschijnlijk niet veel.

Standaardfout

In SPSS kun je de standaardfout laten berekenen. De *standaardfout* geeft aan hoe *betrouwbaar* het gevonden steekproefgemiddelde is als schatting van het populatiegemiddelde. De standaardfout is groter naarmate het verschil in IQ binnen de groep groter en de steekproef kleiner is. De standaardfout wordt dus bepaald door de steekproefgrootte en de homogeniteit van de steekproef.

Op basis van de standaardfout kun je met minimaal 95% zekerheid aangeven dat het populatiegemiddelde ligt tussen het gevonden steekproefgemiddelde minus tweemaal de standaardfout en plus tweemaal de standaardfout. Bij een steekproefgrootte van 10 en een spreiding van 20 is het 95% betrouwbaarheidsinterval bij een gevonden gemiddeld IQ van 100 tussen de 85,69 en 114,31 (zie tabel 1.2), wat betekent dat je niet veel kunt zeggen over de gemiddelde intelligentie in de populatie. Wanneer de steekproefgrootte 1.000 is, wordt het betrouwbaarheidsinterval veel kleiner, namelijk 98,76 - 101,24. Met deze gegevens kun je wél zinnige uitspraken doen over de gemiddelde intelligentie van de populatie.

Tabel 1.2 Voorbeeld invloed spreiding en steekproefgrootte op het betrouwbaarheidsinterval om een gemiddelde te schatten in de populatie op basis van een gevonden steekproefgemiddelde van 100

	<i>Kleine steekproef, grote onderlinge verschillen</i>	<i>Kleine Steekproef, kleine onderlinge verschillen</i>	<i>Grote steekproef, grote onderlinge verschillen</i>	<i>Grote steekproef, kleine onderlinge verschillen</i>
Steekproefgemiddelde	100	100	100	100
Steekproefgrootte	10	10	1000	1000
Spreiding	20	5	20	5
Standaardfout	6,32	1,58	0,63	0,16
95% betrouwbaarheidsinterval	85,69 – 114,31	96,42 – 103,58	98,76 – 101,24	99,69 – 100,31

	<p>Je kunt het gemiddelde, de spreiding, de standaardfout en de betrouwbaarheidsintervallen berekenen door de data in te voeren in de tool GraphPad (www.graphpad.com/quickcalcs/CImean1.cfm).</p>
Kans	<p>Het begrip <i>zekerheid</i> of <i>kans</i> speelt een belangrijke rol in de inductieve statistiek. Ook wanneer je de gemiddelden van twee steekproeven vergelijkt, wil je weten wat de kans is dat het gevonden verschil in gemiddelden ook in de populatie voorkomt. Stel dat de steekproef van werknemers van de firma 'Arbeid' bestaat uit 99 vrouwen en 99 mannen. Je ontdekt dat het gemiddelde IQ van de vrouwen 102,2 is en dat van de mannen 98,2. Kun je dan concluderen dat de vrouwelijke werknemers gemiddeld intelligenter zijn dan de mannelijke? Of dit verschil 'significant' is, kun je toetsen.</p>
Significantie	<p>Wanneer spreek je van <i>significantie</i>? Over het algemeen geldt dat er sprake is van significantie als de overschrijdingskans kleiner is dan 5%, of bij grotere steekproeven (> 1.000) kleiner dan 1%. In de SPSS-uitdraai kun je vaak ook zien of er <i>een- of tweezijdig</i> getoetst is. Je toetst eenzijdig (<i>one-tailed</i>) wanneer je een hypothese of verwachting hebt geformuleerd. Als je bijvoorbeeld een theorie hebt die voorspelt dat de vrouwelijke werknemers intelligenter zijn, kun je eenzijdig toetsen. Heb je echter geen specifieke verwachting over een verschil of richting ervan, dan toets je tweezijdig (<i>two-tailed</i>).</p>
Een- of tweezijdig toetsen	<p>Het bepalen van de significantie is gebaseerd op enkele kenmerken van de steekproef. Dat zijn vaak de omvang en de homogeniteit van de steekproef. Naarmate de steekproef groter is, is de kans op toeval uiteraard kleiner. Ook als de groepen homogener zijn, is de kans kleiner dat verschillen tussen groepen op toeval berusten.</p>
Relevantie	<p>Als een verschil significant is, betekent dat niet automatisch dat het ook <i>relevant</i> is. Uit de analyse van het gemiddelde IQ van de werknemers van de firma Arbeid blijkt dat het verschil in gemiddeld IQ vier IQ-punten in het voordeel van de vrouwen is. Dat verschil is inderdaad significant, de kans dat het op toeval berust, is 2%, dus inderdaad kleiner dan 5%. Hoewel het verschil van vier IQ-punten significant is, is het niet erg relevant, omdat het verschil weinig zegt over de verschillen tussen mannelijke en vrouwelijke werknemers. In onderzoeksverslagen zie je tegenwoordig naast significantie steeds vaker dat ook een <i>effectgrootte</i> wordt vermeld. De gebruikelijke maat voor effectgrootte is <i>Cohen's d</i>. De Cohen's d in het voorbeeld is 0,17, wat een te verwaarlozen effect is. Pas als deze groter is dan 0,20 spreek je van een klein effect. Het geslacht verklaart hier namelijk minder dan 1% van de verschillen in IQ.</p>
Effectgrootte	
Vrijheidsgraden	<p>In de SPSS-uitdraai kom je vaak de term <i>vrijheidsgraden</i> (<i>df = degrees of freedom</i>) tegen. Vrijheidsgraden geven de mate aan waarin scores kunnen variëren. Als je bijvoorbeeld twee getallen hebt en er maar één kent (namelijk 36) met een gemiddelde van 40, dan moet het andere getal 44 zijn. Je hebt hier één vrijheidsgraad: als je namelijk het ene getal weet, dan weet je het andere ook. Bij veel toetsen – zoals de t-toets – is het aantal</p>

vrijheidsgraden het aantal steekproefelementen minus 1. Bij een kruistabel is het aantal vrijheidsgraden het aantal rijen minus 1, vermenigvuldigd met het aantal kolommen minus 1. Voor een 2×2 -tabel is het aantal vrijheidsgraden dus 1. Als de randtotalen van een 2×2 -kruistabel bekend zijn en je een van de celfrequenties weet, dan kun je de andere celfrequenties berekenen. Of een waargenomen verschil of samenhang in een steekproef significant is, hangt mede af van het aantal vrijheidsgraden. Dit aantal vrijheidsgraden is, behalve bij kruistabellen, vaak afhankelijk van de steekproefgrootte.

OPDRACHT

- Voer een (beperkt) deel van je data voor één van de variabelen van minimaal intervalniveaus (zoals leeftijd) in GraphPad in. (graphpad.com/quickcalcs/CImean1.cfm).
- Wat is in deze uitdraai het gemiddelde?
- Wat is de spreiding?
- Wat is de standaardfout?
- Welke betrouwbaarheidsintervallen zijn er?
- Hoe interpreteert je deze data aan de hand van het voorafgaande? Wat valt jou op?

1.5 Hoe werkt SPSS onder Windows?

Zoals je in de inleiding al hebt gelezen, is SPSS een veelgebruikt en uitgebreid statistisch softwarepakket. Als student en docent in het hoger onderwijs kun je via SURFSPOT een campuslicentie aanschaffen tegen een redelijke prijs. Je krijgt dan een licentiecode waarmee je SPSS via SURFSPOT kunt downloaden. In dit boek maken we gebruik van de laatste versies van SPSS. Als je SPSS hebt geïnstalleerd, zie je op je scherm het icoon zoals afgebeeld in figuur 1.4.

Figuur 1.4
Het IBM-SPSS-icoon

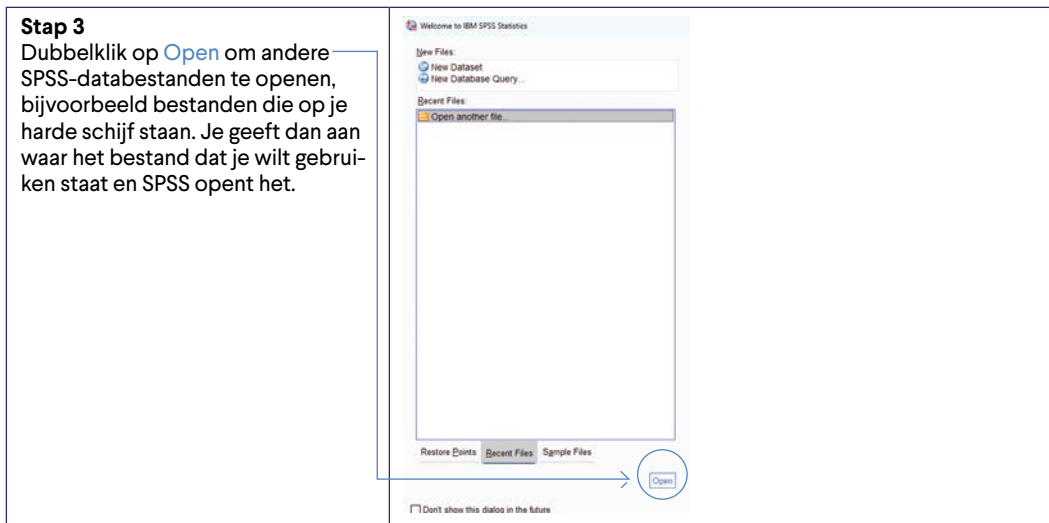


1.5.1 Hoe te werken met een datamatrix onder SPSS?

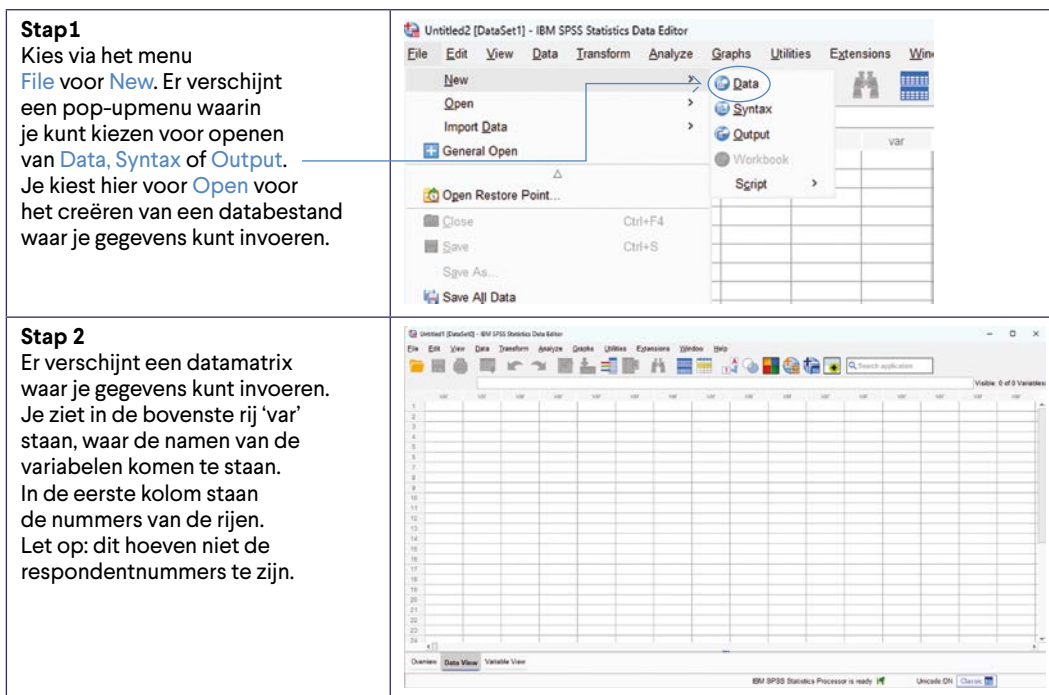
Klik je op het SPSS-icoon, dan kom je in het openingsscherm met het startmenu waarin SPSS eerder geraadpleegde bestanden weergeeft. Kies voor het maken van nieuwe datasets of maak gebruik van de voorbeelden (Sample Files).

Figuur 1.5 Het openingsscherm van SPSS

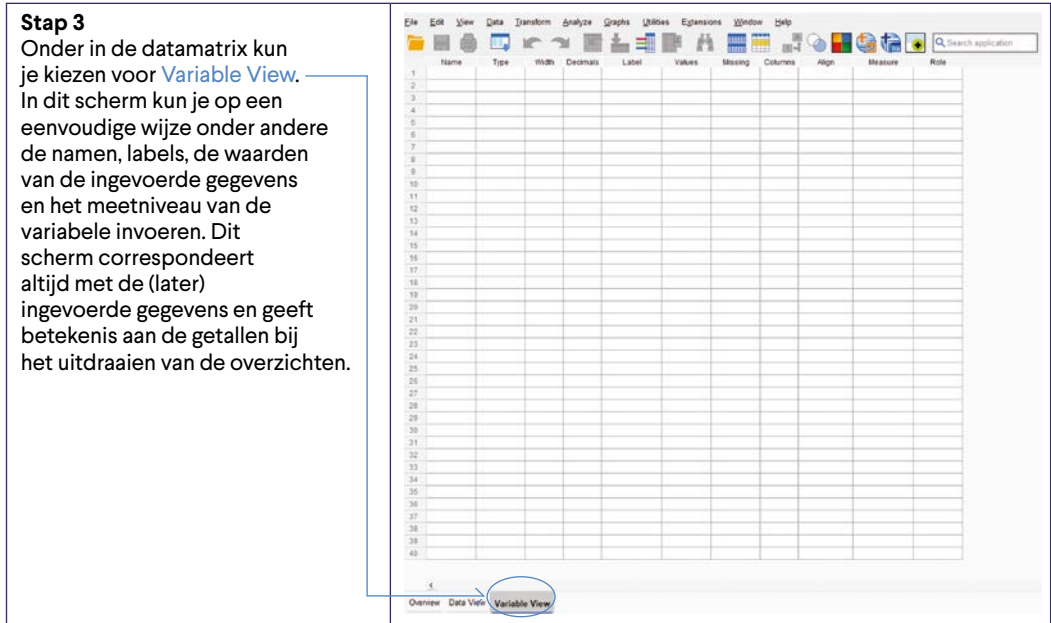
<p>Stap 1 Je ziet het volgende openingsscherm. Als eerste biedt SPSS je de mogelijkheid om te beginnen met het maken van een nieuw databestand. Als je op New Dataset dubbelklikt, opent SPSS een lege datasheet, die je kunt vullen met de gegevens van bijvoorbeeld je enquête.</p>	
<p>Stap 2 Als je verder wilt gaan met een bestand dat je al eerder gebruikt hebt, vind je het onder het kopje Recent files. Dubbelklik dan op het betreffende bestand en SPSS opent het. SPSS markeert het meest recente bestand.</p>	

Figuur 1.5 Het openingsscherm van SPSS (vervolg)

Wanneer je in SPSS een nieuw bestand opent, verschijnt er een nieuw scherm (figuur 1.6) met een lege datamatrix.

Figuur 1.6 Het openen van een datamatrix binnen het SPSS-menu

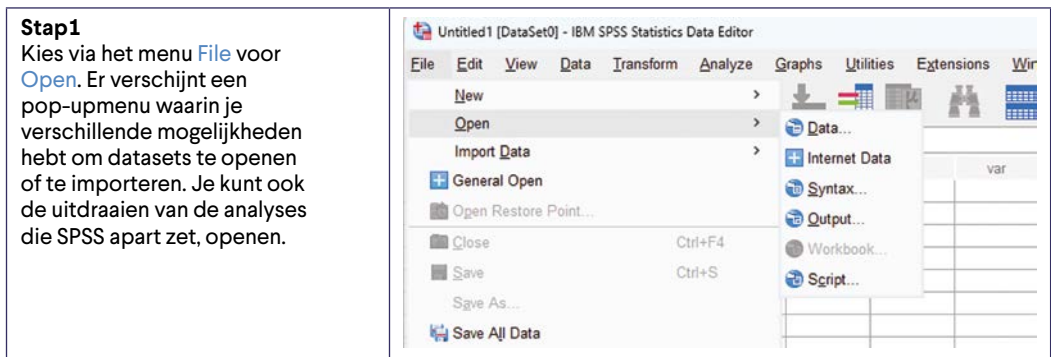
Figuur 1.6 Het openen van een datamatrix binnen het SPSS-menu (vervolg)



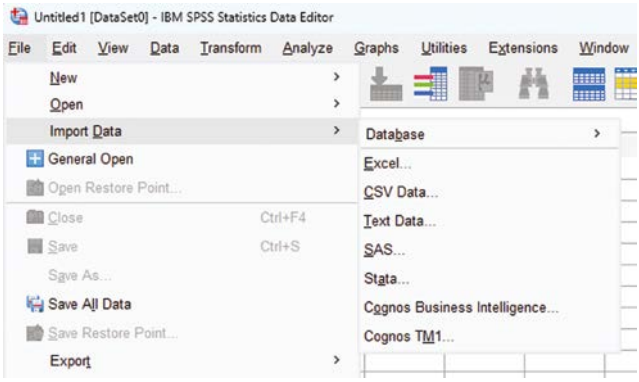
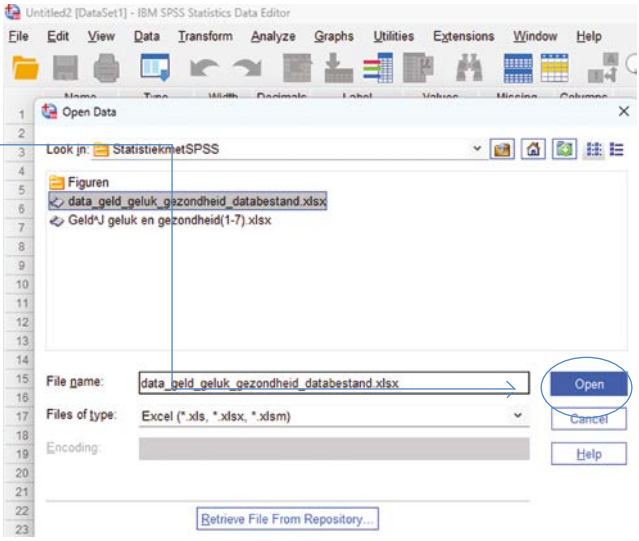
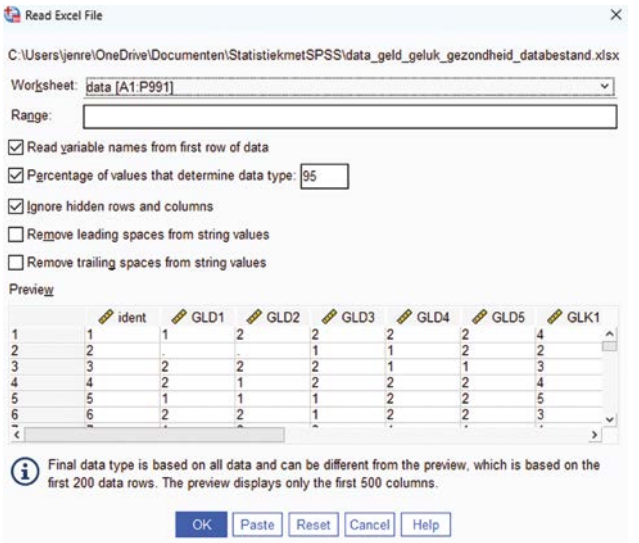
In de taakbalk van SPSS vind je verschillende functies die het programma biedt. De eerste functie is **File** waarmee je op de gebruikelijke manier bestanden kunt openen met de file-knop in de Data editor (zie figuur 1.6). Je kunt hier kiezen uit verschillende soorten files. Je hebt bijvoorbeeld de mogelijkheid om Excel-bestanden of andere soorten data te importeren in SPSS. We concentreren ons hier op het importeren van het Excel-bestand. Je kunt de dataset over 'Geld, Geluk en Gezondheid' dus op verschillende manieren openen. Je haalt het bestand van studiemeester.noordhoff.nl en downloadt het naar één van je mappen. Vervolgens zoek je door **File** en **Open** te kiezen dit bestand op in jouw map. Het dubbelklikken zorgt voor het openen van het bestand en je ziet het volgende overzicht van de data (figuur 1.8).



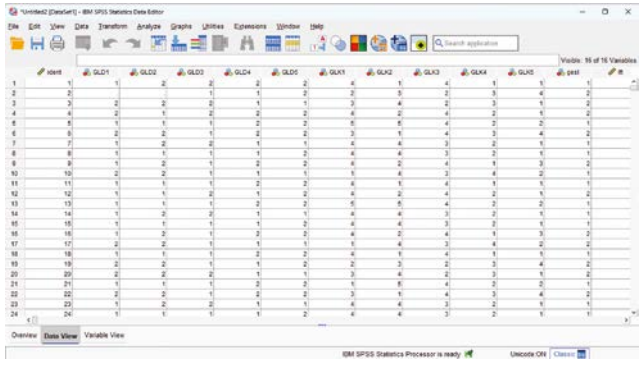
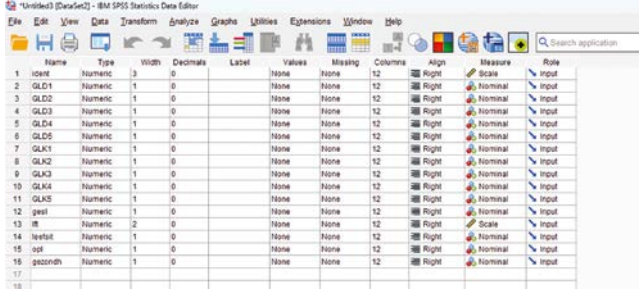

Figuur 1.7 Openen van verschillende datafiles met SPSS



Figuur 1.7 Openen van verschillende datafiles met SPSS (vervolg)

<p>Stap 2 Een andere mogelijkheid is om data die in andere vormen of programma's zijn opgeslagen te openen. SPSS kan bijvoorbeeld op een eenvoudige wijze de data van vragenlijsten die in Excel zijn opgeslagen, openen. SPSS helpt je bij het inlezen van data die voor SPSS 'vreemd' is. Het gaat voor dit Basisboek te ver om uit te leggen hoe dit precies in zijn werk gaat.</p>																																																									
<p>Stap 3 Kies voor Excel. Ga naar de map waar je bestand staat dat je wilt inlezen. Klik op het bestand en daarna op Open.</p>																																																									
<p>Stap 4 Er verschijnt een pop-upmenu waarin gevraagd wordt of je de data wilt inlezen en op welke wijze. Je ziet een beeld waarin SPSS aangeeft op welke wijze de data worden overgenomen. Klik op OK.</p>	 <table border="1" data-bbox="515 1428 1133 1574"> <thead> <tr> <th></th> <th>ident</th> <th>GLD1</th> <th>GLD2</th> <th>GLD3</th> <th>GLD4</th> <th>GLD5</th> <th>GLK1</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>1</td> <td>1</td> <td>2</td> <td>2</td> <td>2</td> <td>4</td> <td>4</td> </tr> <tr> <td>2</td> <td>2</td> <td>.</td> <td>.</td> <td>1</td> <td>1</td> <td>2</td> <td>2</td> </tr> <tr> <td>3</td> <td>3</td> <td>2</td> <td>2</td> <td>2</td> <td>1</td> <td>1</td> <td>3</td> </tr> <tr> <td>4</td> <td>4</td> <td>2</td> <td>1</td> <td>2</td> <td>2</td> <td>2</td> <td>4</td> </tr> <tr> <td>5</td> <td>5</td> <td>1</td> <td>1</td> <td>1</td> <td>2</td> <td>2</td> <td>5</td> </tr> <tr> <td>6</td> <td>6</td> <td>2</td> <td>2</td> <td>1</td> <td>2</td> <td>2</td> <td>3</td> </tr> </tbody> </table>		ident	GLD1	GLD2	GLD3	GLD4	GLD5	GLK1	1	1	1	2	2	2	4	4	2	2	.	.	1	1	2	2	3	3	2	2	2	1	1	3	4	4	2	1	2	2	2	4	5	5	1	1	1	2	2	5	6	6	2	2	1	2	2	3
	ident	GLD1	GLD2	GLD3	GLD4	GLD5	GLK1																																																		
1	1	1	2	2	2	4	4																																																		
2	2	.	.	1	1	2	2																																																		
3	3	2	2	2	1	1	3																																																		
4	4	2	1	2	2	2	4																																																		
5	5	1	1	1	2	2	5																																																		
6	6	2	2	1	2	2	3																																																		

Figuur 1.7 Openen van verschillende datafiles met SPSS (vervolg)

<p>Step 5</p> <p>Je ziet in de kolommen de verschillende variabelen staan. Deze variabelen komen overeen met de uitgezette vragenlijsten. De rijen vormen de respondenten en deze zijn opeenvolgend weergegeven (1-998). De cellen van deze datamatrix zijn gevuld met de scores van de respondenten op de variabelen.</p>	
<p>Step 6</p> <p>SPSS zet de dataset van Excel ook om in een toelichting op de variabelen (Variable View). Dit moet nog ingevuld worden. Hoe je dit doet leer je in hoofdstuk 2.</p>	
<p>Scan de QR-code voor een visuele instructie over het importeren van een databestand uit Excel.</p>	

Variabelenscherm

Het is belangrijk om te weten dat SPSS altijd werkt met twee gekoppelde schermen. In de **Data View** zie je de scores van de respondenten. Daarnaast is er het **Variabele View**-scherm. Als je linksonder op **Variabele View** klikt, kom je in het variabelenscherm. Hierin geeft SPSS aan welke variabelen in het bestand voorkomen en wat de eigenschappen van die variabelen zijn (zie figuur 1.7). Bijvoorbeeld, de variabele *gld1* gaat over het bezit van een elektrische auto. Deze twee schermen zijn aan elkaar gekoppeld. Wanneer je gegevens analyseert, zorgt SPSS ervoor dat de betekenis van de getallen duidelijk is, omdat je deze in het variabelenscherm hebt gedefinieerd.

Je kunt nu een aantal dingen doen. Je kunt de data gaan bewerken. Ook kun je er analyses op uitvoeren via **Analyze** of grafieken maken met **Graphs**. In de volgende hoofdstukken ga je aan de slag met het uitwerken van deze mogelijkheden aan de hand van concrete voorbeelden. Bij het afsluiten van het programma krijg je de vraag of de (eventuele) wijzigingen die je in het bestand hebt aangebracht moeten worden opgeslagen. Het is verstandig een bestand waarin je wijzigingen hebt aangebracht op te slaan onder een nieuwe naam, zodat je altijd nog de laatste versie van het bestand bij de hand hebt als je een fout hebt gemaakt.

OPDRACHT

Oefen met het openen van een databestand en sla deze op in SPSS.

1

1.5.2 Hoe te werken met de output van SPSS?

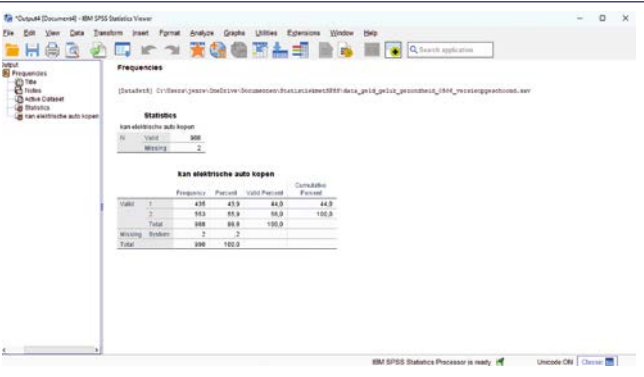
SPSS zet alle uitdraaien over gegevens (zoals frequenties) in een ander scherm en in een ander bestand: de Output. Deze Output is te zien als een verantwoording van de prompt die je hebt gegeven en de analyses of grafieken komen daarin te staan. Je kunt de analyses voor je verslag (tabellen of grafieken) eenvoudig kopiëren naar andere programma's.

Deze bestanden moet je apart bewaren en ze krijgen ook een andere extensie dan je databestanden, namelijk .spv. Ook hier geldt dat je deze bestanden regelmatig moet opslaan.

Figuur 1.8 Het werken met de Output van SPSS

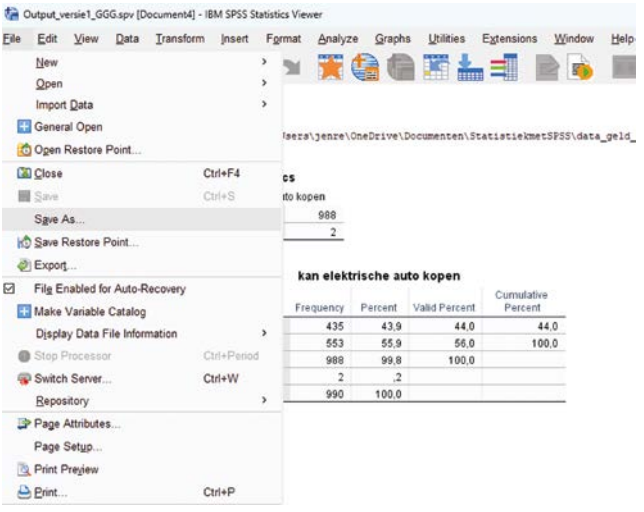
Step 1
SPSS zet alle uitdraaien over gegevens zoals deze frequenties in een ander scherm en in een ander bestand: de *Output*.

Je ziet hier een voorbeeld van een analyse van de vraag of de respondenten een elektrische auto kunnen kopen.

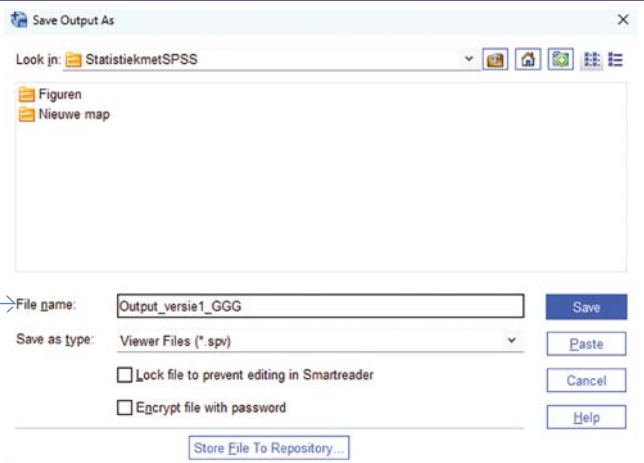
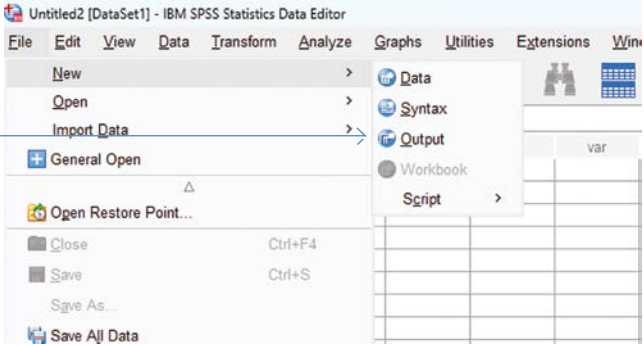
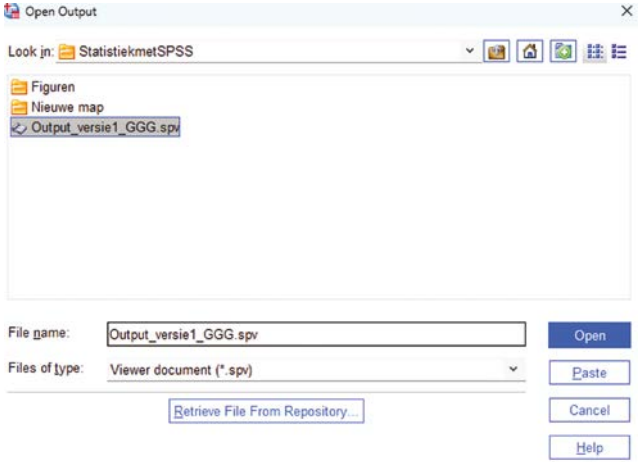


	Frequency	Percent	Valid Percent	Cumulative Percent
ja	435	43,9	44,0	44,0
nee	553	55,9	56,0	100,0
Total	988	99,8	100,0	
Missing	2	,2		
Total	990	100,0		

Step 2
Kies bij de Output voor **File** en **Save As...**



	Frequency	Percent	Valid Percent	Cumulative Percent
ja	435	43,9	44,0	44,0
nee	553	55,9	56,0	100,0
Total	988	99,8	100,0	
Missing	2	,2		
Total	990	100,0		

<p>Stap 3 Geef je bestand een naam en sla het op in de juiste map.</p>	
<p>Stap 4 Je kunt de output terugvinden door via het menu File te kiezen voor Open. Er verschijnt een pop-upmenu waarin je verschillende mogelijkheden hebt om datasets te openen of te importeren. Kies voor Output.</p>	
<p>Stap 5 Ga naar de map waarin jij je bestanden hebt opgeslagen en kies voor de opgeslagen versie van de Output.</p>	

Samenvatting

- SPSS is bruikbaar voor het analyseren van datasets.
- Weet dat er een verschil is tussen de meetniveaus van een variabele. Er is verschil tussen nominaal, ordinaal, interval en ratio meetniveau van de door jou verzamelde gegevens. De verschillende meetniveaus hebben rekenkundige consequenties voor de analyse van de gegevens.
- Zorg voor kennis van belangrijke begrippen, zoals de verdelingen van gegevens (normaal versus scheef) en begrippen als spreiding van de gegevens, de standaardfout, kans en zekerheid, significantie, relevantie, effectgrootte en vrijheidsgraden. Deze begrippen spelen een belangrijke rol bij het komen tot en het interpreteren van uitkomsten uit de verzamelde data.
- SPSS werkt altijd met twee schermen, één scherm waar de data in staan en een variabelenscherm waarin de variabelen en de betekenis van de scores gedefinieerd zijn. Deze twee schermen corresponderen met elkaar bij uitdraaien die SPSS maakt als je gegevens analyseert.