

**Ben Baarda** **Cor van Dijkum**  
**Martijn de Goede**



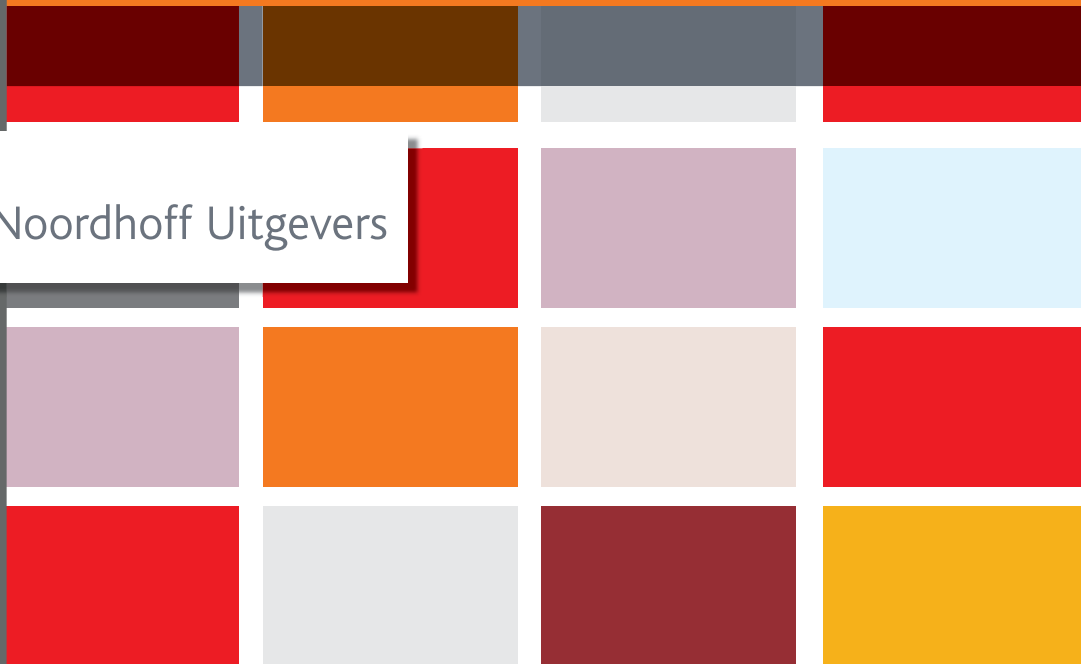
# Basisboek Statistiek met SPSS

Handleiding voor het verwerken en analyseren van  
en rapporteren over (onderzoeks)gegevens

Vijfde druk

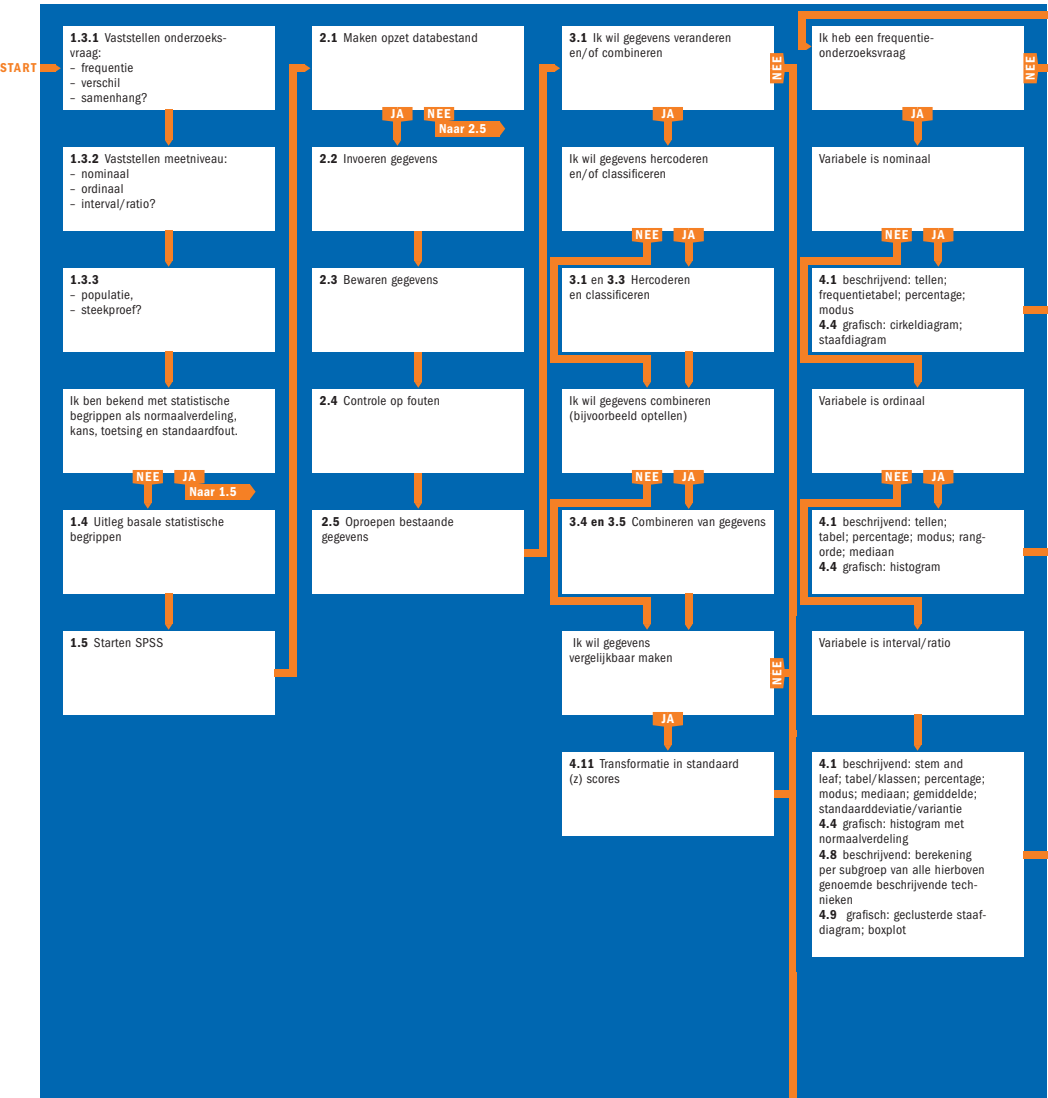


Noordhoff Uitgevers





# Statistiek met SPSS



**5**

Analyse bij verschil-  
onderzoeksvraag

**6**

Analyse bij samenhang-  
onderzoeksvraag

**Bijlage**

Controle homogeniteit  
samengestelde scores

Ik heb een verschil-  
onderzoeksvraag

JA

Splitsingsvariabele is nominaal;  
testvariabele is nominaal

NEE JA

**5.2** Kruistabel  
- Populatie  
- Steekproef

Splitsingsvariabele is nominaal;  
testvariabele is ordinaal

NEE JA

**5.3** Verschil in gemiddelde  
rangorde:  
- Populatie  
- Steekproef gekoppeld  
- twee: Wilcoxon  
- meer: Friedman  
- Steekproef niet-gekoppeld:  
- twee: Mann-Whitney  
- meer: Kruskal-Wallis

Splitsingsvariabele is nominaal;  
testvariabele is interval/ratio

**5.4** Verschil in gemiddelde:  
- Populatie  
- Steekproef gekoppeld:  
- twee: t-test  
- meer: MANOVA (niet in  
dit boek)  
- Steekproef niet-gekoppeld:  
- twee: t-test  
- meer: variantieanalyse

NEE

Ik heb een samenhang-  
onderzoeksvraag

JA

Variabelen zijn nominaal  
(ga uit van de variabelen  
met het laagste meetniveau)

NEE JA

**6.2** Cramér's V  
Bij steekproef: significantie  
vaststellen met Chi-kwadraat

Variabelen zijn ordinaal  
(ga uit van de variabelen  
met het laagste meetniveau)

NEE JA

**6.3** Spearman  
Bij steekproef: significantie  
vaststellen

Variabelen zijn interval/ratio  
(ga uit van de variabelen  
met het laagste meetniveau)

**6.4 en 6.5** Twee variabelen  
- beschrijvend:  
- Pearson  
bij steekproef significantie  
vaststellen  
- enkelvoudige regressie  
- grafisch  
- scattergram  
- regressielijn

**6.6** Meer dan twee variabelen  
- voorspellend:  
- multiple correlatie/  
regressie  
- controlerend:  
- partiële correlatie

NEE

Ik heb een samengestelde score

JA

Vaststellen homogeniteit  
totaalscore

Vaststellen goede/slechte  
items

NEE

Verslag maken



# Basisboek Statistiek met SPSS

**Ben Baarda**  
**Cor van Dijkum**  
**Martijn de Goede**

---

Vijfde geheel herziene druk

Noordhoff Uitgevers Groningen/Houten

Ontwerp omslag: Studio Frank en Lisa, Groningen

Omslagillustratie: iStock

Eventuele op- en aanmerkingen over deze of andere uitgaven kunt u richten aan:  
Noordhoff Uitgevers bv, Afdeling Hoger Onderwijs, Antwoordnummer 13, 9700 VB  
Groningen, e-mail: [info@noordhoff.nl](mailto:info@noordhoff.nl)

*Aan de totstandkoming van deze uitgave is de uiterste zorg besteed. Voor informatie die desondanks onvolledig of onjuist is opgenomen, aanvaarden auteur(s), redactie en uitgever geen aansprakelijkheid. Voor eventuele verbeteringen van de opgenomen gegevens houden zij zich aanbevolen.*

0 / 14

Deze uitgave is gedrukt op FSC-papier.

© 2014 Baarda, Van Dijkum & De Goede, Noordhoff Uitgevers bv Groningen/Houten,  
The Netherlands.

Behoudens de in of krachtens de Auteurswet van 1912 gestelde uitzonderingen mag niets uit deze uitgave worden veelevoudigd, opgeslagen in een geautomatiseerd gegevensbestand of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of enige andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever. Voor zover het maken van reprografische veelevoudigingen uit deze uitgave is toegestaan op grond van artikel 16h Auteurswet 1912 dient men de daarvoor verschuldigde vergoedingen te voldoen aan Stichting Reprorecht (postbus 3060, 2130 KB Hoofddorp, [www.reprorecht.nl](http://www.reprorecht.nl)). Voor het overnemen van gedeelte(n) uit deze uitgave in bloemlezingen, readers en andere compilatiewerken (artikel 16 Auteurswet 1912) kan men zich wenden tot Stichting PRO (Stichting Publicatie- en Reproductierechten Organisatie, postbus 3060, 2130 KB Hoofddorp, [www.stichting-pro.nl](http://www.stichting-pro.nl)).

*All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.*

ISBN (ebook) 978-90-01-85282-5

ISBN 978-90-01-83437-1

NUR 916

# Woord vooraf bij de vijfde druk

In het *Basisboek Statistiek met SPSS* vind je aanwijzingen voor het verwerken en analyseren van onderzoeksgegevens met behulp van SPSS. SPSS staat voor 'Statistical Products and Service Solutions' en is een van de meest gebruikte softwarepakketten voor de invoer en statistische analyse van gegevens. SPSS is een onderdeel van IBM geworden en heet daarom officieel IBM SPSS Statistics. In dit boek maken we je al doende vertrouwd met het SPSS-pakket, maar ook met de statistiek. We nodigen je uit om aan de hand van een databestand over de relatie tussen geld en geluk aan de slag te gaan.

Bij het schrijven van dit boek hebben wij gebruikgemaakt van SPSS, versie 21.0.

Er is een aantal wijzigingen ten opzichte van de vierde druk:

- In de schermafdrukken zijn genummerde post-its opgenomen die de te volgen SPSS-procedure voor de statistische analyse stapsgewijs aangeven. De tekst sluit daarop aan.
- In bijlage II wordt uitgelegd hoe je bij meer dan twee nominale variabelen, bijvoorbeeld geslacht, leefsituatie en al dan niet parttime werken, een loglineaire analyse kunt gebruiken.
- In bijlage III wordt uitgelegd hoe je een variantieanalyse kunt uitvoeren wanneer er sprake is van meer dan één splitsingsvariabele, bijvoorbeeld zowel geslacht als de leefsituatie.

*Het Basisboek Statistiek met SPSS* komt optimaal tot zijn recht in samenhang met het *Basisboek Methoden en Technieken of Dit is onderzoek!*, waarin wij een praktische handleiding geven voor het opzetten en uitvoeren van onderzoek. Aanvullend kan ook het *Statistiekkwartetspel* gebruikt worden, waarin de belangrijkste statistische begrippen en toetsen op speelse wijze worden uitgelegd.

*Het Basisboek Statistiek met SPSS* is uiteraard ook los daarvan goed te gebruiken.

Ben Baarda  
Cor van Dijkum  
Martijn de Goede  
Winter 2014



# Inleiding

Vele duizenden organisaties over de hele wereld maken gebruik van IBM® SPSS® Statistics, zowel in het bedrijfsleven, binnen de overheid als in de academische wereld. De software helpt ze het antwoord te vinden op belangrijke bedrijfs- en onderzoeksvragen en maakt het ze gemakkelijk om onderzoeksgegevens te benaderen en beheren, verschillende analyses uit te voeren en de uitkomsten met anderen te delen.

IBM SPSS Statistics bestaat uit verschillende oplossingen die nauw met elkaar samenwerken. Ze ondersteunen het volledige analytische proces – van het plannen van onderzoek en het verzamelen en analyseren van data, tot het rapporteren van de onderzoeksresultaten. Op basis van geavanceerde analyses en modellen is het mogelijk om patronen in gegevens te herkennen en met een grote nauwkeurigheid ontwikkelingen en trends te voorspellen.

Oorspronkelijk werd IBM SPSS Statistics vooral gebruikt in de sociale wetenschappen, bijvoorbeeld om onderzoek te doen op het gebied van politicologie, sociologie en psychologie. Tegenwoordig wordt de software ook veel ingezet in het kader van marketing, marktonderzoek, beleidsonderzoek en fraudedetectie.

Zo maken allerlei soorten bedrijven gebruik van IBM SPSS Statistics om klantgestuurd te werken. Ze maken klantprofielen en segmentaties, en brengen het (toekomstige) gedrag en de behoeften van klanten in kaart om er doelgericht op in te spelen. Ook onderzoeken ze wie hun beste klanten of meest interessante doelgroepen zijn, voor optimaal marketingrendement. Tevens zetten organisaties de software in om nieuwe klanten te werven en te behouden.

Zo gebruiken mediabedrijven IBM SPSS Statistics bijvoorbeeld om uit te vinden wat het sentiment is over bepaalde tv-programma's op de sociale media, zodat ze aanpassingen kunnen doen om de kijkcijfers te verhogen. Sportorganisaties zetten analytics in om het risico op blessures te voorspellen en tegen te gaan. De politie past de software toe om criminaliteit te voorspellen en voorkomen, terwijl gemeenten onder meer uitkeringsfraude beter bestrijden. Ziekenhuizen kunnen met behulp van analytics de meest effectieve behandelmethodes bepalen.

Op al deze en andere manieren zijn organisaties in de meest uiteenlopende sectoren met IBM SPSS Statistics in staat om betere beslissingen te nemen en hun resultaten te verhogen. In dit boek ontdek je hoe dat in zijn werk gaat, zodat ook jij de software met goed gevolg in de praktijk kunt inzetten.

Veel succes!  
Laila Fettah  
IBM Nederland

# Inhoud

Effectief studeren 10

## **1 Hoe bereid ik mij voor op statistiek met SPSS 21.0? 13**

- 1.1 Inleiding 14
  - 1.2 Geld en geluk; toelichting op de casus 15
  - 1.3 Hoe analyseer ik mijn data? Een gebruikswijzer! 16
  - 1.4 Enkele algemene statistische begrippen 19
  - 1.5 Hoe werkt SPSS onder Windows? 21
- [Samenvatting 26](#)

## **2 Hoe breng ik mijn gegevens in de computer? 29**

- 2.1 Hoe maak ik de opzet voor het SPSS-bestand met Variable View? 30
  - 2.2 Hoe voer ik mijn gegevens in de computer in? 36
  - 2.3 Hoe bewaar ik mijn ingevoerde gegevens? 38
  - 2.4 Hoe controleer ik of ik fouten heb gemaakt bij de invoer van mijn gegevens? 40
  - 2.5 Hoe roep ik de ingevoerde gegevens weer op? 41
- [Samenvatting 42](#)

## **3 Hoe verander en combineer ik gegevens? 45**

- 3.1 Hoe hercodeer ik gegevens? 46
  - 3.2 Hoe bewaar ik gehercodeerde gegevens? 48
  - 3.3 Hoe kan ik een variabele indelen in klassen? 48
  - 3.4 Hoe kan ik gegevens combineren? 50
  - 3.5 Hoe kan ik bij het combineren van gegevens rekening houden met ontbrekende waarden? 51
- [Samenvatting 55](#)

## **4 Hoe analyseer ik mijn gegevens bij een frequentie-onderzoeksvraag? 57**

- 4.1 Wat is een frequentieverdeling? 58
- 4.2 Wanneer gebruik ik een frequentieverdeling? 60
- 4.3 Hoe maak ik een frequentieverdeling met SPSS? 61
- 4.4 Hoe geef ik frequenties grafisch weer? 62

- 4.5 Hoe moet ik de uitvoer van de opdracht Frequencies lezen? [65](#)
- 4.6 Hoe kan ik een frequentieverdeling presenteren? [66](#)
- 4.7 Hoe moet ik een frequentieverdeling interpreteren en erover rapporteren? [67](#)
- 4.8 Hoe bereken ik frequenties voor subgroepen? [67](#)
- 4.9 Hoe vergelijk ik (sub)groepen? [68](#)
- 4.10 Wat doe ik met multiple responsen? [72](#)
- 4.11 Hoe maak ik variabelen vergelijkbaar? [74](#)  
[Samenvatting 76](#)

## **5 Hoe analyseer ik mijn gegevens bij een verschil- onderzoeksvraag? [79](#)**

- 5.1 Wat is een verschilvraag? [80](#)
- 5.2 Verschilvraag bij nominale test- en splitsingsvariabele: kruistabel en chi-kwadraat [81](#)
- 5.3 Verschil bij ordinale test- en nominale splitsingsvariabelen: niet-parametrische toetsen [88](#)
- 5.4 Verschil bij interval/ratio testvariabele en nominale splitsingsvariabele; t-toets of variantieanalyse? [100](#)  
[Samenvatting 114](#)

## **6 Hoe analyseer ik mijn gegevens bij een samenhang- onderzoeksvraag? [117](#)**

- 6.1 Wat is een samenhangonderzoeksvraag? [118](#)
- 6.2 Samenhang tussen twee nominale variabelen: Cramér's V [120](#)
- 6.3 Samenhang tussen twee ordinale variabelen: Spearman's rangcorrelatie! [123](#)
- 6.4 Samenhang tussen twee interval- of ratiovariabelen: Pearson's productmoment-correlatiecoëfficiënt! [127](#)
- 6.5 Samenhang tussen twee interval- of ratiovariabelen: enkelvoudige regressieanalyse [133](#)
- 6.6 Samenhang bij twee of meer variabelen op interval- of ratio-niveau: multiple correlatie- en regressieanalyse [137](#)
- 6.7 Samenhang tussen twee interval- of ratiovariabelen, gecontroleerd voor het effect van een derde variabele: partiële correlatie [141](#)  
[Samenvatting 144](#)

Bijlage I Hoe controleer ik de homogeniteit van samengestelde scores? [146](#)

Bijlage II De samenhang tussen meerdere nominale variabelen [152](#)

Bijlage III De samenhang van meerdere onafhankelijke variabelen met één  
afhankelijke variabele [158](#)

Illustratieverantwoording [163](#)

Register [164](#)

Over de auteurs [166](#)

# Effectief studeren

In dit *Basisboek Statistiek met SPSS* willen we je leren al doende vertrouwd te raken met het pakket. Na een algemene inleiding en kennismaking met de basisregels van SPSS kun je direct beginnen met het invoeren van gegevens in de computer en het analyseren daarvan.

We nodigen de lezer uit om aan de hand van een databestand over de relatie tussen geld en geluk aan de slag te gaan. Dit databestand omvat gefingeerde vragenlijstgegevens van vijfhonderd mannen en vijfhonderd vrouwen. Daarbij volgen we de gebruikelijke gang van zaken in een onderzoek:

- We beginnen bij het invoeren van de gegevens (hoofdstuk 2).
- Vervolgens leggen we uit hoe je gegevens kunt veranderen en combineren met SPSS (hoofdstuk 3).

Na dit voorbereidende werk volgt de analyse. Welke analysetechniek gebruikt moet worden is onder meer afhankelijk van je specifieke onderzoeksvraag. Je vraagstelling bevat altijd een of meer specifieke onderzoeksvragen waarop je via onderzoek een antwoord probeert te vinden. We maken daarbij een onderscheid tussen:

- frequentie-onderzoeksvragen (hoe vaak of in welke mate komt iets voor; hoofdstuk 4)
- verschil-onderzoeksvragen (vaststellen en toetsen van verschillen op één kenmerk tussen twee of meer groepen; hoofdstuk 5)
- samenhang-onderzoeksvragen (vaststellen en toetsen van een samenhang tussen twee kenmerken; hoofdstuk 6)

In de bijlagen bespreken we een aantal technieken, zoals de homogeniteitsanalyse, de loglineaire analyse en de univariate variantieanalyse, die wat verder gaan dan de basale statistische technieken.

Om verantwoord met SPSS te kunnen werken is niet alleen kennis van SPSS, maar ook statistische kennis nodig. In het eerste hoofdstuk bespreken we basisbegrippen als meetniveau, normaalverdeling, kans, significantie, een- en tweezijdige toetsing, power en effectgrootte. In de hoofdstukken 4, 5 en 6 wordt uitgelegd dat de keuze voor een statistische techniek niet alleen afhankelijk is van het type onderzoeksvraag maar ook van het meetniveau van de betreffende kenmerken, en of het om een populatie of om een steekproef gaat.

De stof wordt verder gepresenteerd aan de hand van de volgende vragen die je je als gebruiker stelt:

- Welke statistische analysetechniek moet ik gezien mijn onderzoeksvraag gebruiken?
- Wat houdt die techniek in, wat zijn de voorwaarden en hoe voer ik de techniek in SPSS uit?

- 
- Als het gelukt is de gekozen analysetechniek met SPSS uit te voeren, hoe lees ik dan de computer-output? Wat staat er precies?
  - Hoe moet ik de resultaten interpreteren? Wat betekenen de resultaten voor mijn onderzoeksvraag?
  - Hoe presenteer ik de resultaten in mijn onderzoeksrapport?
  - Hoe schrijf ik mijn conclusie(s) op?

Voor elke besproken techniek laten we een voorbeeld van de SPSS-uitvoer zien. Aan de hand van dit voorbeeld leggen we uit hoe je het moet lezen en wat het betekent. Om dit laatste te illustreren, laten we telkens zien hoe je over het resultaat kunt rapporteren.

Er is veel aandacht besteed aan de didactische opbouw van het boek. Elk hoofdstuk kent de volgende indeling:

- *voorkennis*: de kennis die nodig is om het betreffende hoofdstuk te lezen
- *vragen*: de vragen die in het hoofdstuk beantwoord worden
- *begrippen*: de belangrijkste begrippen die in het hoofdstuk behandeld worden
- *inhoudelijke informatie*: de specifieke stof voor het in het hoofdstuk behandelde onderwerp
- *samenvatting*

Bij het schrijven van dit boek hebben wij gebruikgemaakt van SPSS, versie 21.0. Dit programma is op de computers van de meeste hogescholen en universiteiten geïnstalleerd. Je kunt het als student met korting bestellen bij [www.surfspot.nl](http://www.surfspot.nl).

De databestanden waarop we de analyses uitvoeren, zijn te vinden op internet via <http://basisboekstatistiekmetpss.noordhoff.nl>

# 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* (5e herziene druk, 2012) of *Dit is onderzoek!* (1e druk, 2010) en *het Statistiekkwartetspel* (1e druk, 2010) te gebruiken.

## 1

# Hoe bereid ik mij voor op statistiek met SPSS 21.0?

In dit hoofdstuk worden de volgende vragen en begrippen behandeld:

- Welke onderwerpen worden in dit boek behandeld (1.1)?
- Waarover gaat de casus 'Geld en geluk' in dit boek (1.2)?
- Voor de analyse is het van belang om te weten (1.3):
  - Gaat het in de onderzoeksvraag om een frequentie, verschil, of om een samenhang (1.3.1)?
  - Wat is het meetniveau (1.3.2)?
  - Gaat het om een steekproef of een populatie (1.3.3)?
- Wat houden de statistische begrippen normaalverdeling, standaardfout, betrouwbaarheid, zekerheid, significantie, een- of tweezijdige toetsing, relevantie, effectgrootte, vrijheidsgraden in (1.4)?
- Hoe werkt SPSS-21 onder Windows (1.5)?

---

Frequenties 16

Verskil 16

Samenhang 16

Nominaal meetniveau 17

Ordinaal meetniveau 17

Intervalmeetniveau 17

Ratiomeetniveau 17

Nulpunt 17

Continue variabele 18

Discrete variabele 18

Beschrijvende statistiek 18

Inductieve/inferentiële  
statistiek 19

Normaalverdeling 19

Standaardfout 20

Significantie 20

Een- of tweezijdig toetsen 20

Effectgrootte 21

Vrijheidsgraden 21

Variable Viewscherf 23

Data Editorscherf 23



## 1.1 Inleiding

Als je onderzoek doet, is het kiezen van de juiste statistische techniek om de verzamelde gegevens te analyseren een belangrijke schakel in de lange reeks van beslissingen die je neemt. Het uiteindelijke doel is het beantwoorden van de onderzoeksvraag of -vragen.

Om de plaats van de data-analyse in de onderzoekscyclus als geheel duidelijk te maken, zetten we de gebruikelijke fasen van een onderzoek op een rij. Elke fase van de onderzoekscyclus is in de vorm van een vraag opgenomen:

- 1 Wat is de onderzoeksvraag of -vragen 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?

In dit boek staan de fasen 7 en 8 en een deel van fase 9 centraal: preparatie, analyse en beschrijving van de onderzoeksgegevens die met SPSS geanalyseerd worden. Aan de hand van een onderzoek naar de relatie tussen geld en geluk (zie figuur 1.1), behandelen we stap voor stap:

- Hoe prepareer je onderzoeksgegevens om ze te kunnen invoeren (hoofdstuk 2)?
- 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 (hercoderen), of de waarden van gegevens combineren tot een nieuwe score (hoofdstuk 3) en vervolgens controleren op hun betrouwbaarheid (paragraaf 3.4 en de bijlage).
- Hoe kies je de juiste analysetechniek? Om de juiste analysetechniek te kiezen stel je eerst vast wat het karakter is van je onderzoeksvraag (subparagraaf 1.3.1). Gaat het om frequenties (hoofdstuk 4), verschillen (hoofdstuk 5) of om samenhangen (hoofdstuk 6). Vervolgens ga je na wat het meetniveau van je gegevens is (subparagraaf 1.3.2). 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 op de binnenkant van de kaft en als los inlegvel in dit boek.
- Hoe voer je de analyse met SPSS uit en hoe interpreteer je de resultaten? Dit geven we telkens aan bij elke statistische techniek die we behandelen. We bespreken hoe je verslag kunt doen van de resultaten.

In dit hoofdstuk bespreken we verder een aantal belangrijke statistische begrippen zoals normaalverdeling, significantie, kans en standaardfout (paragraaf 1.4). In de laatste paragraaf van dit hoofdstuk (paragraaf 1.5) leggen we uit hoe je SPSS kunt starten.

## 1.2 Geld en geluk; toelichting op de casus

Voordat we ingaan op het gebruik van SPSS, introduceren we eerst het voorbeeldonderzoek 'Geld en geluk' (zie figuur 1.1). De vragen uit figuur 1.1 zijn fictief voorgelegd aan een voor de Nederlandse volwassen bevolking representatieve steekproef van vijfhonderd Nederlandse mannen en vijfhonderd Nederlandse vrouwen van 25 tot en met 55 jaar. Er is bewust voor deze leeftijdsgrenzen gekozen. Veel jonge mensen studeren nog en hebben daardoor geen vast inkomen. Bij mensen ouder dan 55 jaar is er vaak al sprake van (gedeeltelijke) uittreding uit het arbeidsproces waardoor zij in een andere financiële situatie verkeren. De data van dit onderzoek vind je op de website <http://basisboekstatistiekmetpspp.noordhoff.nl> onder de naam 'data1'.

FIGUUR 1.1 Voorbeeldonderzoek 'Geld en geluk'

### Maakt geld gelukkig?

Een onderzoeker wil weten of er een verband bestaat tussen geld en geluk. Hij vraagt zich af of geld gelukkig maakt. Zijn centrale onderzoeksvraag luidt dan ook: 'Is er een positief verband tussen de mate waarin iemand over geld beschikt en de mate waarin hij/zij zich gelukkig voelt?' Het begrip 'geld' heeft de onderzoeker als volgt gedefinieerd: geld is de hoeveelheid financiële middelen waar iemand over kan beschikken. Deze definitie is met opzet ruim gekozen. Hierdoor worden niet alleen het inkomen en het vermogen, maar ook andere financiële bronnen waar iemand over 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. De onderzoeker heeft deze twee begrippen geoperationaliseerd in een vragenlijst. Zowel voor het meten van het begrip geld, als voor het begrip geluk heeft hij vijf uitspraken of items gemaakt.

Geld is gemeten met de items:

- |                                    |               |
|------------------------------------|---------------|
| 1 Ik ben in het bezit van een auto | <i>ja/nee</i> |
| 2 Ik heb een koopwoning/flat       | <i>ja/nee</i> |
| 3 Ik bezit een tablet              | <i>ja/nee</i> |
| 4 Ik krijg zorgtoeslag             | <i>ja/nee</i> |
| 5 Ik krijg huurtoeslag             | <i>ja/nee</i> |

5 Ik voel mij eenzaam.

- |                       |                       |                       |                       |                       |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| <i>absoluut</i>       | <i>niet</i>           | <i>ten dele</i>       | <i>wel</i>            | <i>absoluut</i>       |
| <i>niet</i>           |                       | <i>wel/niet</i>       |                       | <i>wel</i>            |

Geluk is gemeten met de volgende vragen:

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

- |                       |                       |                       |                       |                       |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| <i>absoluut</i>       | <i>niet</i>           | <i>ten dele</i>       | <i>wel</i>            | <i>absoluut</i>       |
| <i>niet</i>           |                       | <i>wel/niet</i>       |                       | <i>wel</i>            |

- 2 De meeste anderen hebben het beter dan ik.

- |                       |                       |                       |                       |                       |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| <i>absoluut</i>       | <i>niet</i>           | <i>ten dele</i>       | <i>wel</i>            | <i>absoluut</i>       |
| <i>niet</i>           |                       | <i>wel/niet</i>       |                       | <i>wel</i>            |

- 3 Ik heb het ... naar mijn zin.

- |                       |                       |                       |                       |                       |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| <i>absoluut</i>       | <i>niet</i>           | <i>ten dele</i>       | <i>wel</i>            | <i>absoluut</i>       |
| <i>niet</i>           |                       | <i>wel/niet</i>       |                       | <i>wel</i>            |

- 4 Het leven is zwaar.

- |                       |                       |                       |                       |                       |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| <i>absoluut</i>       | <i>niet</i>           | <i>ten dele</i>       | <i>wel</i>            | <i>absoluut</i>       |
| <i>niet</i>           |                       | <i>wel/niet</i>       |                       | <i>wel</i>            |

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

Geslacht  *man*  *vrouw*

Leeftijd in jaren ...

Leefsituatie  *alleen*

*met partner*

*met partner en kinderen*

Hoogste afgeronde opleiding

*lager (beroeps)onderwijs (basisschool, lager technisch onderwijs, lager voortgezet onderwijs)*

*middelbaar (beroeps)onderwijs (middelbaar technisch onderwijs, middelbaar voortgezet onderwijs, mavo, mulo, vmbo, enz.)*

*hoger (beroeps)onderwijs (universiteit, hbo, vwo, havo)*

De onderzoeker heeft de vraagstelling uitgewerkt in een aantal specifieke onderzoeksvragen:

- 1 Hoeveel Nederlanders bezitten respectievelijk een auto, koopwoning, tablet?
- 2 Hoeveel Nederlanders krijgen zorgtoeslag?
- 3 Hoeveel Nederlanders ontvangen huurtoeslag?
- 4 Hoe tevreden zijn Nederlanders over het leven dat zij leiden?
- 5 Hoe eenzaam voelen Nederlanders zich?
- 6 Zijn er verschillen tussen mannen en vrouwen wat betreft het hebben van een auto, tablet en een spelcomputer; zorgtoeslag ontvangen; het ontvangen van huursubsidie?
- 7 Is er een verschil in tevredenheid over het leven dat men leidt tussen:
  - mensen met en zonder partner
  - mensen met en zonder kinderen?
- 8 Is er een verschil in eenzaamheid tussen mannen en vrouwen?
- 9 Is er een verschil in tevredenheid over het leven dat men leidt tussen mannen en vrouwen?
- 10 Is er een relatie tussen de mate van tevredenheid over het leven en de leeftijd?
- 11 Bestaat er een relatie tussen geld en geluk?

### 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:

- 1 Gaat het in de vraagstelling om frequenties (hoe vaak/in welke mate), om een verschil of om een samenhang? Of gaat het om een combinatie daarvan?
- 2 Wat is het meetniveau (nominaal, ordinaal of interval-/rationiveau) van de gegevens die je hebt verzameld?
- 3 Gaat het om een steekproef of om een populatie?

Het blokschema 'Hoe analyseer ik mijn data' (te vinden aan de binnenzijde van de kaft van dit boek en als los inlegvel) is ontworpen aan de hand van deze vragen. In de volgende subparagrafen gaan we daar nader op 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 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:

- vragen waarbij het erom gaat hoe vaak of in welke mate (**frequentie**) iets voorkomt; bijvoorbeeld: 'In welke mate zijn Nederlanders gelukkig?' of 'Hoeveel procent van de mensen is in het bezit van een auto?'
- vragen waarbij het gaat om een **verschil**; bijvoorbeeld: 'Zijn mannen gelukkiger dan vrouwen?'
- vragen waarbij het gaat om een **samenhang**; bijvoorbeeld: 'Is er een samenhang tussen geld en geluk?'

Het is duidelijk dat de voorbeeldonderzoeksvragen 1 tot en met 5 (paragraaf 1.2) zijn te karakteriseren als frequentieonderzoeksvragen. Het gaat bijvoorbeeld bij onderzoeksvraag 1 om het aantal Nederlanders dat onder andere een auto bezit. In hoofdstuk 4 geven we een voorbeeld van de analyse van gegevens bij dit type vraagstelling. De onderzoeksvragen 6 tot en met 9 in dezelfde paragraaf zijn verschilonderzoeksvragen. De analyse van gegevens bij dit type onderzoeksvraag behandelen we in hoofdstuk 5. De onderzoeksvragen 10 en 11 tot slot zijn samenhangonderzoeksvragen. In hoofdstuk 6 behandelen we voorbeelden van data-analyse bij een samenhangonderzoeksvraag.

### 1.3.2 Wat is het meetniveau van mijn gegevens?

Als je hebt bepaald om wat voor type onderzoeksvragen het in je onderzoek gaat (zie de eerste kolom in het blokschema 'Hoe analyseer ik mijn data'), dan ga je vervolgens na op welk meetniveau de variabele(n) is/zijn gemeten. Zie daarvoor in het genoemde blokschema de cellen onder frequentie, verschil of samenhang. Per onderzoeksvraag geef je aan wat het meetniveau is van de betreffende variabelen. Bij onderzoeksvraag 9 (het verschil tussen mannen en vrouwen 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 onderzoeksvraag 11 (de samenhang tussen geld en geluk). Bij de variabele geslacht zijn er maar twee categorieën of waarden, namelijk 'man' en 'vrouw'. Hierbij is alleen sprake van een verschil, maar niet van meer of minder. Een man is anders dan een vrouw maar niet meer of minder. Hetzelfde geldt voor leefsituatie; iemand is gehuwd, samenwonend of alleenstaand. Bij dit type antwoordmogelijkheden gaat het om een *nominaal meetniveau*. Je kunt alleen zeggen hoeveel mannen of vrouwen een auto hebben, maar niet dat iemand meer 'mans' of meer 'vrouws' is. Je bent man of vrouw. Je kunt niet een beetje man of veel vrouw zijn.

Nominaal  
meetniveau

Dat kan wel bij gegevens die zijn gemeten op respectievelijk ordinaal, interval- of rationiveau. Bij gegevens op *ordinaal meetniveau* is er wél sprake van meer of minder, maar het verschil tussen de categorieën is niet in een getal uit te drukken. Bij 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-meter loper die goud wint sneller heeft gelopen dan een loper die zilver heeft gewonnen. Het feit dat hij goud heeft gewonnen geeft aan dat hij sneller heeft gelopen, maar niet hoeveel sneller.

Ordinaal  
meetniveau

Bij interval- en ratiomeetniveau is dat verschil tussen categorieën in termen van meer of minder wel in een getal uit te drukken. Temperatuur is daar een goed voorbeeld van. Het verschil tussen 5 en 10 graden Celsius is even groot als het verschil tussen 45 en 50 graden. Bij het *intervalmeetniveau* is er alleen geen sprake van een natuurlijk *nulpunt*, zoals wel het geval is bij gegevens die op ratiomeetniveau 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 er een natuurlijk nulpunt.

Interval-  
meetniveau  
Ratio-  
meetniveau  
Nulpunt

Dit heeft gevolgen voor de rekenkundige bewerkingen die je mag uitvoeren. Bij temperatuur gemeten in graden Celsius mag je niet zeggen dat 20 graden tweemaal zo veel is als 10 graden. Bij een meting in graden Kelvin

mag je wel zeggen dat 20 graden tweemaal zo veel is als 10 graden en bij een gewichtsmeting dat 20 kilo tweemaal zo zwaar is als 10 kilo. Overigens vat SPSS het interval- en ratiomeetniveau samen onder het begrip *Scale*. De beide andere meetniveaus worden *Nominal* en *Ordinal* genoemd. We onderscheiden verder continue en 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 van een persoon, leeftijd en intelligentie. Variabelen die alleen hele waarden kunnen aannemen, noemen we *discrete variabelen*, zoals het aantal auto's dat iemand bezit of het aantal kinderen in een gezin.

## Continue variabelen

## Discrete variabelen

**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

### 1.3.3 Gaat het om een populatie of om een steekproef?

**FIGUUR 1.2** Voorbeeld van een populatie en een aselechte steekproef



N=100 hele populatie van werknemers

n=10 aselekt getrokken steekproef uit populatie van werknemers

## Beschrijvende statistiek

Er zijn twee vormen van statistiek: de beschrijvende en de inductieve of inferentiële statistiek. *Beschrijvende statistiek* gebruik je wanneer je onderzoek doet bij een populatie. Er is sprake van een populatie wanneer je alle eenheden waarover je uitspraken wilt doen in je onderzoek betreft. 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 aselekt uit het

totale bestand van werknemers trekt. Bij een steekproef is de opzet dat je uitspraken doet over de totale populatie van werknemers. In dit geval moet je gebruikmaken van de *inductieve of inferentiële statistiek*. Je wilt namelijk op grond van een speciaal geval (een steekproef) algemene uitspraken doen (over de populatie). Raadpleeg een methoden- of statistiekboek voor meer informatie. Ook op Wikipedia vind je uitgebreide informatie over het trekken van steekproeven: [http://en.wikipedia.org/wiki/Sampling\\_\(statistics\)](http://en.wikipedia.org/wiki/Sampling_(statistics)). Voordat je aan de analyse van je gegevens begint, stel je jezelf de vraag over welke eenheden (wie of wat) je uitspraken wilt doen. Wanneer dat alle personen of zaken zijn die in je onderzoek zijn betrokken, 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 als het ware gerepresenteerd worden door de onderzoekseenheden die je geselecteerd hebt, dan gaat het om een steekproefonderzoek. In de volgende paragraaf behandelen we in het kort enkele statistische termen die je steeds weer tegenkomt bij het toetsen of de resultaten die je in een steekproef vindt op toeval berusten, of met een bepaalde marge aan onzekerheid kunnen worden gegeneraliseerd naar de populatie waaruit de steekproef is getrokken.

Inductieve/  
inferentiële  
statistiek

1

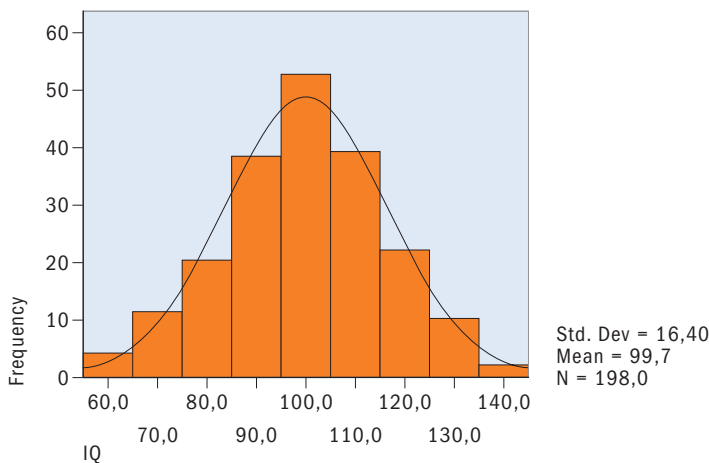
## 1.4 Enkele algemene statistische begrippen

Zoals in de voorgaande paragraaf is aangegeven, zijn er twee vormen van statistiek: de beschrijvende en de inductieve, of inferentiële statistiek. Het doel van de beschrijvende statistiek is op een inzichtelijke en overzichtelijke manier je gegevens te presenteren. Als je van bijna duizend werknemers de arbeidssatisfactie hebt vastgesteld, heeft het weinig zin al die gegevens los te presenteren. Meestal vat je ze samen in percentages of een gemiddelde (paragraaf 4.1) of in een histogram (subparagraaf 4.1.1). Je beschrijft je gegevens in een gereduceerde en daardoor overzichtelijke vorm.

Als je gegevens grafisch weergeeft, is het resultaat nogal eens de zogenoemde *normaalverdeling*. In figuur 1.3 zie je daar een (fictief) voorbeeld van.

Normaal-  
verdeling

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



De figuur geeft de verdeling van scores op een intelligentietest weer die is afgenomen bij 198 werknemers van de firma Arbeid. Die verdeling neemt nagenoeg de vorm van een normaalverdeling aan. Ter vergelijking is de normaalverdeling ingetekend. Deze verdeling wordt ook wel de verdeling van *Gauss* of *Gausskromme* genoemd. Het karakteristieke kenmerk van de normaalverdeling is de vorm van een klok en de symmetrie van de linker- en de rechterhelft.

Met SPSS kun je controleren of de verdeling van jouw gegevens de normaalverdeling benadert (paragraaf 4.4).

Wanneer de 198 werknemers een aselechte steekproef vormen uit het totale bestand van werknemers van de firma Arbeid ( $N = 2213$ ), moeten we gebruikmaken van de inductieve of inferentiële statistiek. De vraag is dan in welke mate 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 wanneer je inderdaad de gehele populatie in je onderzoek zou kunnen betrekken? Die kans is uiteraard niet zo groot. Want het gemiddelde dat je in de steekproef hebt gevonden is afhankelijk van de toevallige samenstelling van deze steekproef. Wanneer we opnieuw een aselechte steekproef trekken en nog een keer et cetera, dan zal het gemiddelde IQ waarschijnlijk steeds iets hoger of mogelijk iets lager zijn. De gevonden waarden zullen wel iets van elkaar afwijken, maar waarschijnlijk ook weer niet erg veel. In SPSS kun je de standaardfout laten berekenen (paragraaf 4.3). De *standaardfout* geeft aan in hoeverre het gevonden steekproefgemiddelde een *betrouwbare* schatting is van het populatiegemiddelde. De standaardfout wordt bepaald door de steekproefgrootte en de homogeniteit van de steekproef. De standaardfout is groter naarmate de steekproef kleiner is en het verschil in IQ binnen de groep groter is. Op basis van de standaardfout kun je bijvoorbeeld met minstens 95% zekerheid aangeven dat het populatiegemiddelde ligt tussen het gevonden steekproefgemiddelde minus tweemaal de standaardfout en het gemiddelde plus tweemaal de standaardfout.

Het begrip *zekerheid* of *kans* speelt een belangrijke rol in de inductieve statistiek. Ook wanneer je de gemiddelden van twee steekproeven vergelijkt, is het de vraag wat de kans is dat je een gevonden verschil in gemiddelden terugvindt in de populatie. Stel dat de steekproef van werknemers van de firma Arbeid uit 99 vrouwen en 99 mannen bestaat. Je vindt dat het gemiddelde IQ van de vrouwen 102,2 is en dat van de mannen 98,2. Kun je dan stellen dat de vrouwelijke werknemers van de firma Arbeid gemiddeld intelligenter zijn dan de mannelijke werknemers? Of dat verschil 'significant' is, kun je toetsen. In hoofdstuk 5 leggen we uit hoe je dat doet voor verschillen en in hoofdstuk 6 hoe je dat doet bij samenhangen.

Wanneer spreek je nu van *significantie*? Men houdt over het algemeen de regel aan dat er van significantie sprake is als de overschrijdingskans kleiner is dan 5% of bij grotere steekproeven ( $> 1000$ ) kleiner is dan 1%. Vaak staat er in de SPSS-uitdraai ook bij of er *een- of tweezijdig* getoetst is. Je toetst eenzijdig (*one-tailed*) wanneer je een hypothese of verwachting hebt. Als je een theorie hebt op grond waarvan je kunt verwachten dat de vrouwelijke werknemers intelligenter zijn, dan kun je eenzijdig toetsen. Heb je echter geen idee of er sprake is van een verschil en zeker niet van de richting van dat verschil, dan toets je tweezijdig (*two-tailed*).

Het bepalen van de significantie is gebaseerd op enkele kenmerken van de steekproef. Dat zijn vaak de omvang en de homogeniteit van de steekproef.

### Standaardfout

### Significantie

### Een- of tweezijdig toetsen

Naarmate de steekproef groter is, is de kans op toeval uiteraard kleiner. En naarmate de verschillen op een variabele binnen groepen kleiner zijn (homogene groepen), is de kans dat de verschillen tussen groepen op toeval berusten eveneens kleiner.

Als een verschil significant is, wil dat niet automatisch zeggen dat het ook *relevant* is. Je ziet in het voorbeeld 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%. Het verschil van vier IQ-punten is dan wel significant, maar niet erg relevant. Het zegt erg weinig over de verschillen tussen mannelijke en vrouwelijke werknemers. Je ziet tegenwoordig in onderzoeksverslagen dat naast de significantie steeds vaker een *effectgrootte* wordt vermeld. De gebruikelijke maat voor effectgrootte is *Cohen's d*. De *Cohen's d* in het voorbeeld is 0,17, wat een te verwaarlozen effect is. Pas als *d* groter is dan 0,20 wordt er van een klein effect gesproken. Het geslacht verklaart hier namelijk nog geen 1% van de verschillen in IQ. In hoofdstuk 5 leggen we uit hoe je de effectgrootte kunt berekenen.

In de SPSS-uitvoer kom je ook vaak de term *vrijheidsgraden* (*df = degrees of freedom*) tegen. Het aantal vrijheidsgraden geeft de mate aan waarin scores kunnen variëren. Als je van twee getallen er maar één kent (namelijk 36) en je weet dat het gemiddelde 40 is, dan moet het andere getal 44 zijn. Je hebt hier één vrijheidsgraad. Als je namelijk het ene getal weet, weet je het andere ook. Bij veel toetsen, zoals de t-toets (subparagraaf 5.4), is het aantal vrijheidsgraden het aantal steekproefelementen minus 1. Bij een kruistabel (zie paragraaf 5.2) 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 weet ook een van de celfrequenties, dan kun je de andere celfrequenties berekenen. Vrijheidsgraden zijn van belang als je op basis van een steekproef een schatting wilt maken van bijvoorbeeld het gemiddelde van de populatie. Of een waargenomen verschil of samenhang in een steekproef significant is, hangt dus mede af van het aantal vrijheidsgraden. Dit aantal vrijheidsgraden is, met uitzondering van kruistabellen, vaak afhankelijk van de grootte van de steekproef.

Effectgrootte

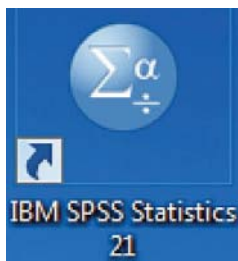
Vrijheidsgraden

1

## 1.5 Hoe werkt SPSS onder Windows?

Zoals in de inleiding al aangegeven, is SPSS een uitgebreid statistisch softwarepakket dat veel gebruikt wordt. Als SPSS is geïnstalleerd, vind je, als het goed is, het icoon op je scherm zoals afgebeeld in figuur 1.4.

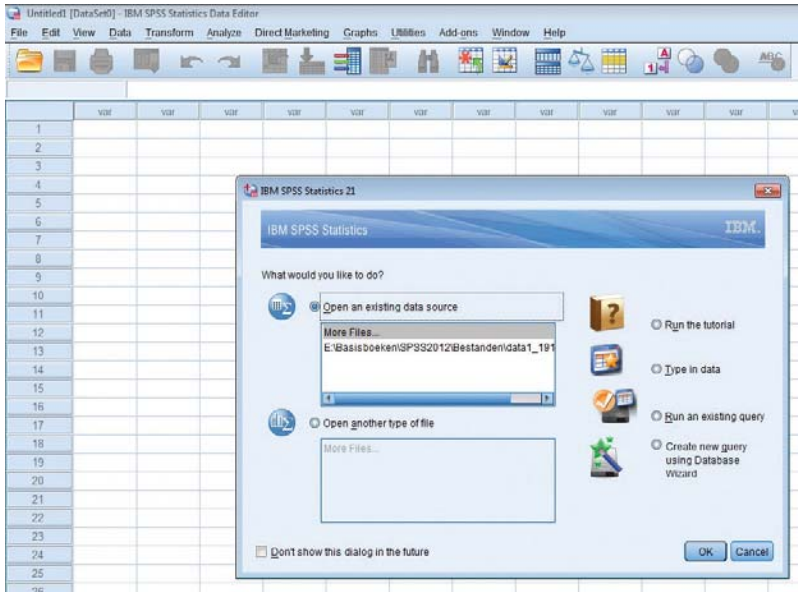
FIGUUR 1.4 Het SPSS-icoon





Je 'opent' SPSS door op het icoon te dubbelklikken. Als je het icoon niet kunt vinden dan kun je altijd nog op de startknop linksonder op je scherm klikken. Je kiest met de muisaanwijzer vervolgens 'Programma's'. Controleer dan of IBM SPSS Statistics voorkomt in het rijtje programma's dat op het scherm verschijnt. Als dat het geval is, kun je SPSS starten door daarop te dubbelklikken.

FIGUUR 1.5 Het openingsscherm van SPSS

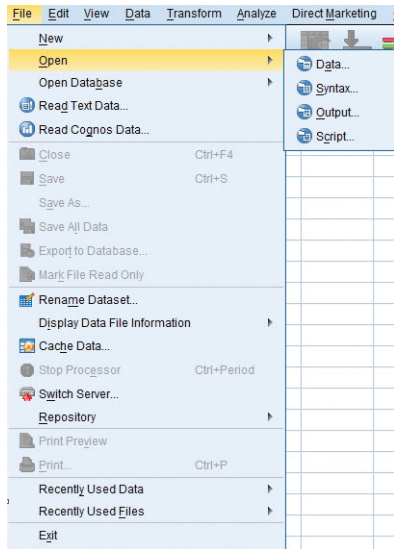


Je ziet dan het openingsscherm zoals dat in figuur 1.5 is weergegeven.

In het openingsscherm zijn verschillende mogelijkheden aangegeven:

- *Open an existing data source.*  
Als je verder wilt gaan met een bestand dat je al eerder gebruikt hebt, vind je het waarschijnlijk onder het kopje 'Open an existing data source'. Je klikt dan op het betreffende bestand en op 'ok'. SPSS heeft dit standaard aangevinkt.  
Als je alleen op 'ok' klikt, kun je aangeven waar een ander bestand staat dat je wilt gebruiken en kan SPSS het openen.  
Je kunt een bestand ook op de meer gebruikelijke manier openen met de file-knop die je in de Data Editor vindt (zie figuur 1.6). Je ziet daar dat je kunt kiezen uit verschillende soorten files. We concentreren ons hier op de data-file. De andere soorten files komen later aan bod.

FIGUUR 1.6 Het openen van een datafile via het file-menu



Het eerste wat je ziet als je een bestand opent, is het *Data Views*scherm; daar staan alle ingevoerde gegevens.

**Data  
Views**scherm

FIGUUR 1.7 Het Data Viewscherm als je Data1 opent

 A screenshot of the SPSS Data Views screen. The window title is '\*data1.sav [DataSet1] - IBM SPSS Statistics Data Editor'. The menu bar includes File, Edit, View, Data, Transform, Analyze, Direct Marketing, Graphs, Utilities, Add-ons, Window, and Help. The toolbar contains various icons for file operations and analysis. The data table below has 26 rows and 10 columns.
 

	ident	gid1	gid2	gid3	gid4	gid5	gik1	gik2	gik3	gik4
1	1	1	1	1	2	2	4	1	4	1
2	2	.	.	1	1	2	2	3	2	3
3	3	2	2	2	1	1	3	4	2	3
4	4	2	1	2	2	2	4	2	4	2
5	5	1	1	1	2	2	5	5	4	2
6	6	2	2	1	2	2	3	1	4	3
7	7	1	2	2	1	1	4	4	3	2
8	8	1	1	1	1	2	4	4	3	2
9	9	1	2	1	2	2	4	2	4	1
10	10	2	2	1	1	1	1	4	3	4
11	11	1	1	1	2	2	4	1	4	1
12	12	.	1	2	1	2	4	2	4	2
13	13	1	1	1	2	2	5	5	4	2
14	14	1	2	2	1	1	4	4	3	2
15	15	1	1	1	1	2	4	4	3	2
16	16	1	2	1	2	2	4	2	4	1
17	17	2	2	1	1	1	1	4	3	4
18	18	1	1	1	2	2	4	1	4	1
19	19	2	2	1	1	2	2	3	2	3
20	20	2	2	2	1	1	3	4	2	3
21	21	1	1	1	2	2	5	5	4	2
22	22	2	2	1	2	2	3	1	4	3
23	23	1	2	2	1	1	4	4	3	2
24	24	1	1	1	1	2	4	4	3	2
25	25	1	2	1	2	2	4	2	4	1
26	26	2	2	1	1	.	1	4	3	4

Als je links onderin op 'Variable View' klikt, kom je in het *variabelenschermbord*; daarin geeft SPSS aan welke variabelen in het bestand voorkomen en wat de eigenschappen van die variabelen zijn. In hoofdstuk 2 leggen we uit wat al die eigenschappen inhouden en hoe je ze kunt veranderen. Dit zie je in figuur 1.8. Je kunt nu een aantal dingen doen. Je kunt de data gaan bewerken. Je kunt er ook analyses op uitvoeren via 'Analyze', of grafieken maken met 'Graphs'. In de volgende hoofdstukken werken we deze mogelijkheden uit aan de hand van concrete voorbeelden.

FIGUUR 1.8 Het Variable Viewschermbord als je Data1 opent

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	ident	Numeric	8	0	identificatienum...	None	None	8	Right	Nominal	Input
2	geld1	Numeric	8	0	geld1: auto	{1, ja}...	None	8	Right	Nominal	Input
3	geld2	Numeric	8	0	geld2: koopwoni	{1, ja}	None	8	Right	Nominal	Input
4	geld3	Numeric	8	0	geld3: tablet	{1, ja}	None	8	Right	Nominal	Input
5	geld4	Numeric	8	0	geld4: ziekenfon	{1, ja}	None	8	Right	Nominal	Input
6	geld5	Numeric	8	0	geld5: huursubs	{1, ja}	None	8	Right	Nominal	Input
7	gk1	Numeric	8	0	gk1: zverdean	{1, absoluut...	None	8	Right	Ordinal	Input
8	gk2	Numeric	8	0	gk2: anderen b...	{1, absoluut...	None	8	Right	Ordinal	Input
9	gk3	Numeric	8	0	gk3: naar mijn zin	{1, absoluut...	None	8	Right	Ordinal	Input
10	gk4	Numeric	8	0	gk4: leven is zw...	{1, absoluut...	None	8	Right	Ordinal	Input
11	gk5	Numeric	8	0	gk5: eenzaam	{1, absoluut...	None	8	Right	Ordinal	Input
12	gesl	Numeric	8	0	geslacht	{1, man}	None	8	Right	Nominal	Input
13	leeft	Numeric	8	0	leeftijd in jaren	None	None	8	Right	Scale	Input
14	leefsit	Numeric	8	0	leefstusie	{1, alleen}	None	8	Right	Nominal	Input
15	opst	Numeric	8	0	opleiding	{1, laag}	None	8	Right	Ordinal	Input

- *Run the tutorial* biedt je de mogelijkheid om over allerlei onderwerpen (zie figuur 1.9) nadere uitleg te krijgen. Als je op 'All topics' klikt, krijg je het overzicht van de behandelde onderwerpen.

FIGUUR 1.9 De onderwerpen in het Tutorialonderdeel

Search:   [Search scope: All topics](#)

**Contents**

- [Help](#)
- Tutorial**
  - Introduction
  - Reading Data
  - Using the Data Editor
  - Working with Multiple Data Sources
  - Examining Summary Statistics for Individual Variables
  - Crosstabulation Tables
  - Creating and editing charts
  - Working with Output
  - Working with Syntax
  - Modifying Data Values
  - Time Saving Features
  - Customizing IBM SPSS Statistics
  - Automated production
  - Scoring data with predictive models
  - Getting Started with Custom Tables
- Case Studies**
- Statistics Coach**
- Add-ons**
  - Integration Plug-in for Python Help
  - Integration Plug-in for R Help
  - Integration Plug-in for Java User Guide
  - Integration Plug-In for Java API Reference
  - Integration Plug-in for Microsoft .NET User Guide
  - Working with R

- *Type in data.* Als je nieuwe gegevens hebt die je nog niet hebt ingetypt in SPSS of in een ander programma als Excel, kun je het beste voor deze optie kiezen. SPSS opent dan de Data Editor (figuur 1.7) voor je. In hoofdstuk 2 wordt uitgelegd hoe je in de nu geopende data-editor de gegevens voor jouw onderzoek invoert.
- *Run an existing query/ Create new query.* In SPSS kun je ook werken met gegevens uit andere dataverwerkingsprogramma's zoals Excel. Alleen moet je dan wel aangeven hoe SPSS de Excelgegevens moet lezen en dat doe je in een query. Je herkent de query aan de extensie .spq. Als je al een query hebt, klik je het derde rondje aan; moet je nog een query maken, dan klik je het vierde rondje aan. SPSS helpt je dan bij het maken van de query. Het gaat voor dit Basisboek te ver, om uit te leggen hoe dit precies in zijn werk gaat. Wil je meer weten over het importeren van Excel-gegevens in SPSS, kijk dan eens op de volgende site:  
<http://www.statutorials.com/SPSS/TUTORIAL-SPSS-Prepare-Data-Excel.htm>

Het SPSS-programma *sluit je af* door op het witte kruisje in de rechterbovenhoek te klikken. Bij afsluiten van het programma wordt je gevraagd of de (eventuele) wijzigingen die je in het bestand hebt aangebracht moeten worden opgeslagen. Als je aangeeft dat je het gewijzigde bestand wilt *opslaan*, word je verzocht het bestand een naam te geven. Als je een andere naam intypt, dan wordt de gewijzigde versie van het bestand bewaard naast de ongewijzigde versie. Laat je de naam ongewijzigd, dan wordt het gewijzigde bestand 'over' de eerdere versie van het bestand 'heen geschreven'. Het is daarom verstandig een bestand waarin je wijzigingen hebt aangebracht op te slaan onder een nieuwe naam, mocht je een fout hebben gemaakt dan heb je altijd nog de laatste versie van het bestand.

# Samenvatting

1

- ▶ In paragraaf 1.1 wordt aangegeven wat in dit boek besproken wordt aan de hand van de voorbeeldcasus 'Geld en geluk' (paragraaf 1.2).
- ▶ Het is het verstandig om wanneer je zelf een onderzoek doet voor je begint jezelf een aantal vragen te stellen:
  - Gaat het in mijn onderzoek om een frequentie, verschil of samenhang? (subparagraaf 1.3.1)
  - Wat is het meetniveau van mijn gegevens: nominaal, ordinaal, interval of ratio? (subparagraaf 1.3.2)
  - Betreft het een populatie of een steekproef, want in het geval van een steekproef moet je gebruikmaken van inductieve of inferentiële statistiek (subparagraaf 1.3.3)
- ▶ In paragraaf 1.4 worden kort de volgende statistische begrippen besproken: normaalverdeling, standaardfout, betrouwbaarheid, zekerheid, significantie, een- of tweezijdige toetsing, relevantie, effectgrootte, vrijheidsgraden.
- ▶ In paragraaf 1.5 wordt uitgelegd hoe je SPSS start, uit welke schermen (de hoofdschermen: 'Variable View' en 'Data View') het is opgebouwd en waar je binnen SPSS hulp kunt vinden (de 'Tutorials').