

Basiscursus SPSS

© 2020, Uitgeverij Koninklijke Van Gorcum BV, Postbus 43, 9400 AA Assen.

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

NUR 916

ISBN folioboek: 978 90 232 5719 6

ISBN ebook: 978 90 232 5720 2

1e druk 2004

2e herziene druk 2006

3e herziene druk 2009

4e herziene druk 2011

5e herziene druk 2013

6e herziene druk 2016

7e herziene druk 2020

Ofschoon iedere poging is ondernomen volgens de auteurswet rechthebbenden van het in dit boek opgenomen illustratiemateriaal te traceren, is dit in enkele gevallen niet mogelijk gebleken. In het onderhavige geval verzoekt de uitgever rechthebbenden contact met hem op te nemen.

Uitgave: Uitgeverij Koninklijke Van Gorcum, Assen

Grafische verzorging: Richard Bos

Omslagontwerp: Kim Boeren, Viesrood grafisch & interactief ontwerp, Zwolle

Druk: Drukkerij Van Gorcum, Raalte



BASISCURSUS SPSS

Pieter van Groenestijn

Michael Savelkoul

7e herziene druk

VOORWOORD	7
Voorwoord bij de zevende druk	7
1 HET PROGRAMMA SPSS	8
1.1 Wat is SPSS?	9
1.2 Waarvoor dient SPSS?	9
1.3 Structuur van het boek	9
1.4 De fictieve gegevens in de voorbeelden	10
1.5 Vensters, balken en menu's	11
1.5.1 Het datavenster	11
1.5.2 De titelbalk	12
1.5.4 De knoppenbalk	13
1.5.5 De statusbalk	13
1.6 Werken met vensters	13
1.6.1 Het datavenster: Data View en Variable View	13
1.6.2 Het outputvenster	15
1.6.3 Het syntaxvenster	15
1.7 Bestandstypen	16
2 SPSS-BESTANDEN	17
2.1 Inleiding	18
2.2 Het openen van bestanden	18
2.3 Het aanmaken van bestanden	19
2.3.1 Het definiëren van variabelen in SPSS	20
2.3.2 Variabelen voorzien van labels	21
2.3.3 Variabelen categorienamen geven (value labels)	22
2.3.4 Het definiëren van ongeldige scores (missing values)	24
2.3.5 Het invoeren van gegevens	26
2.4 Het openen van Excel-bestanden	29
2.5 Het opslaan van SPSS-bestanden	31
2.6 Opdrachten	32
3 BEWERKEN VAN VARIABELEN	35
3.1 Inleiding	36
3.2 Het hercoderen van variabelen	36
3.3 Het berekenen van nieuwe variabelen	40
3.4 Het selecteren van analyse-eenheden (cases)	42
3.5 Het opsplitsen van bestanden	45
3.6 Opdrachten	47

4	BESCHRIJVENDE STATISTIEK	49
4.1	Inleiding	50
4.2	Numerieke analyse	50
4.2.1	Univariate numerieke analyse: frequentieverdeling	51
4.2.2	Univariate numerieke analyse: centrum- en spreidingsmaten	52
4.2.3	Bivariate numerieke analyse: kruistabel	53
4.3	Grafische analyse	57
4.3.1	Univariate grafische analyse: staafdiagram	57
4.3.2	Univariate grafische analyse: cirkeldiagram	58
4.3.3	Univariate grafische analyse: boxplot	60
4.3.4	Bivariate grafische analyse	62
4.4	Uitvoer naar tekstverwerkingsprogramma's	64
4.5	Opdrachten	65
5		66
5.1	Inleiding	67
5.2	Univariate toets: t-toets op één gemiddelde	67
5.3	Bivariate toetsen	69
5.3.1	Toetsen met kruistabel	69
5.3.2	T-toets voor verschil tussen twee gemiddelden	72
5.3.3	Variantieanalyse	76
5.3.4	Correlatie: Pearson	79
5.4	Multivariate toets: meervoudige regressieanalyse	80
5.5	Opdrachten	83
	TOT SLOT	84
	IBM SPSS	85

VOORWOORD

In 1968 startte in Californië op de Stanford University een project dat zou uitgroeien tot één van de meest populaire statistische programma's: SPSS (oorspronkelijk: **S**tatistical **P**ackage for the **S**ocial **S**ciences). Het programma is in een groot aantal leerboeken beschreven, vaak in combinatie met statistische theorie. Aan de Radboud Universiteit Nijmegen bleek er echter behoefte aan een relatief eenvoudige cursus SPSS om in korte tijd de basisbeginselen over te dragen. Daarom werd de uitgebreide uitleg van de statistiek geschrapt en kwam de nadruk te liggen op het zelfstandig uitvoeren van de juiste computerhandelingen. Voor dit doel werd een niet te omvangrijk dictaat geschreven dat na elk cursusjaar werd aangepast en verbeterd. Zo is uiteindelijk dit boekje ontstaan, waarmee in zes tot acht uur voldoende kennis, maar vooral vaardigheid wordt ontwikkeld om gangbare statistische bewerkingen en analyses zelfstandig uit te voeren.

Manfred te Grotenhuis en Anneke Matthijssen
Nijmegen, 2004

VOORWOORD BIJ DE ZEVENDE DRUK

De Basiscursus SPSS wordt inmiddels al ruim 15 jaar gebruikt op hogescholen en universiteiten in Nederland en België. Met het overlijden van Manfred te Grotenhuis verloor de Radboud Universiteit een van de grondleggers van het boek.

De behoefte aan een toegankelijke uitleg van de basishandelingen in het programma SPSS, zonder uitgebreid in te gaan op de statistische achtergronden, blijft echter onder de studenten aanwezig. Ook docenten richtten verzoeken aan ons om deze basiscursus te actualiseren. Met de komst van nieuwe versies van SPSS vonden de auteurs het daarom tijd voor een update van de Basiscursus SPSS.

De afbeeldingen en de instructies zijn afgestemd op de versies 25 en 26 van SPSS. De onderwerpen zijn deels hetzelfde gebleven in vergelijking met de vorige druk. De opbouw van het hoofdstuk over beschrijvende statistiek en het hoofdstuk over toetsende statistiek heeft wel een fundamentele wijziging ondergaan. Ook wordt er in deze versie gewerkt met kaders met aanvullende informatie voor lezers die op sommige punten extra diepgang zoeken.

Wij bedanken Anneke Matthijssen en de erven van Manfred te Grotenhuis voor het in ons gestelde vertrouwen om deze basiscursus ook in de toekomst voort te zetten. We hopen hiermee ook een bijdrage te kunnen leveren aan toekomstig statistiekonderwijs in Nederland en België.

Pieter van Groenestijn en Michael Savelkoul
Nijmegen, 2020

2

SPSS- BESTANDEN

2.1 INLEIDING

In SPSS hebben bestanden met gegevens een centrale rol. De gegevens (men spreekt ook wel van *data*) kunnen op allerlei manieren worden verzameld. Voor het gemak gaan we hier uit van ons eigen voorbeeld, namelijk de ingevulde evaluatieformulieren van een cursus statistiek. De gegevens van dit onderzoek kunnen op verschillende wijze worden aangeleverd:

- De gegevens zijn al verwerkt in een databestand; men hoeft alleen het bestand te openen om de statistische bewerkingen uit te voeren (paragraaf 2.2).
- Men heeft alleen de ingevulde vragenlijsten, dus de gegevens zijn nog niet gedigitaliseerd. Er moet daarom nog een SPSS-databestand worden gemaakt (paragraaf 2.3).
- De gegevens van de vragenlijsten zijn als rijtjes getallen in de computer ingevoerd (bijvoorbeeld met Excel); van dit bestand wordt een SPSS-databestand gemaakt (paragraaf 2.4).

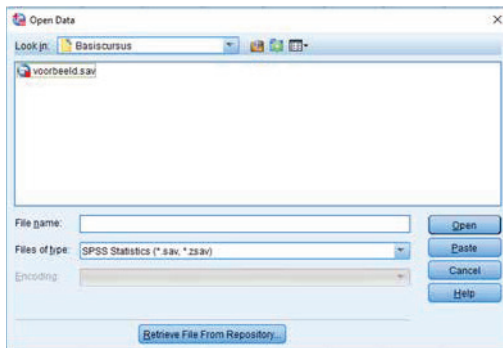
2.2 HET OPENEN VAN BESTANDEN

We gaan laten zien op welke wijze de beschreven bestandstypen (zie paragraaf 1.7) geopend worden.



- Start SPSS op, via het pictogram van SPSS op het bureaublad, of door op de pc ernaar te zoeken.
- Als er een SPSS-welkomvenster verschijnt met tal van mogelijkheden, sluit dit dan af met *close*.
- Kies na het verschijnen van het lege datavenster van SPSS in de menubalk: **File – Open – Data**.

Nu verschijnt het dialoogvenster Open Data (zie afbeelding 2.1)



Afbeelding 2.1 Het dialoogvenster Open Data



- Selecteer achter Look in: (of Zoeken in:) de juiste folder (hier: basiscursus) en open het bestand *voorbeeld.sav* door erop te dubbelklikken of door te klikken (selecteren) en te kiezen voor **Open** (zie afbeelding 2.1).

Op soortgelijke wijze kunnen outputbestanden worden geopend, dit zijn bestanden met daarin de uitkomsten van statistische analyses. Open nu dit bestand volgens de aanwijzingen in het volgende kader.



- Kies in de menubalk: **File – Open** en klik op **Output**. Het bijbehorende dialoogvenster verschijnt.
- Selecteer indien nodig de juiste schijf en de folder basiscursus.
- Open door te dubbelklikken het bestand *output.spv*.

In dit bestand is een overzicht te zien van de leeftijden van 37 studenten die het evaluatieformulier hebben ingevuld. De jongste vier studenten zijn 18 jaar oud en dat is 10,8% van het totaal (37) en de twee oudste studenten zijn 37 jaar (5,4% van het totaal). Deze tabel is eerder in hoofdstuk 1 ter sprake gekomen (afbeelding 1.4).

Tot slot gaan we op soortgelijke wijze een syntaxbestand openen.



- Kies in de menubalk: **File – Open** en klik op **Syntax**. Het bijbehorende dialoogvenster verschijnt.
- Selecteer indien nodig de juiste schijf en de folder basiscursus.
- Open door te dubbelklikken het bestand *syntax.spv*.
- Selecteer de tekst met de muis en klik op ► in de knoppenbalk.
- In de output verschijnt nu de eerder besproken frequentietabel van leeftijd (zie afbeelding 1.4).

2.3 HET AANMAKEN VAN BESTANDEN

Indien de onderzoeksgegevens op ingevulde papieren vragenlijsten staan, moet men er eerst een digitaal bestand van maken. Men begint dan eerst met een omschrijving van de in te voeren variabelen: de zogenaamde *datadefinitie*. Dit betekent dat de vragen uit de lijst worden omgezet in variabelen die een naam en eventueel een verdere omschrijving (label) hebben in het digitale bestand. Nadat

alle vragen een naam (en label) hebben gekregen, kan worden begonnen met het invoeren van de scores (waarden) die bij deze variabelen horen.

2.3.1 Het definiëren van variabelen in SPSS

Standaard opent SPSS met een leeg datavenster. Je hebt nu waarschijnlijk het bestand *voorbeeld.sav* open (zie paragraaf 2.2). Een nieuw leeg datavenster open je als volgt:



- Kies in de menubalk: **File – New – Data**.
- Afhankelijk van de vorige keer dat SPSS werd gebruikt, verschijnt nu of de Data View of de Variabele View. Zorg ervoor dat laatstgenoemde openstaat.
- Kies (indien nodig) voor **Variable View**.



Op soortgelijke wijze kan een leeg syntaxvenster worden geopend:
File – New – Syntax.

Het resultaat is dat je nu een bestand krijgt dat vergelijkbaar is met afbeelding 1.3, echter is er nu (nog) niets ingevuld. Voordat de waarden worden ingevoerd, is het gebruikelijk om eerst aan te geven (te definiëren) welke variabelen er in het bestand voorkomen. Dit wordt uitgelegd aan de hand van het onderstaande voorbeeld.

VOORBEELD

Men wil twee variabelen definiëren aan de hand van de enquêtevragen ‘Hoeveel tijd besteedt u per week thuis aan statistiekopgaven?’ en ‘Bent u een man of een vrouw?’ Kies als naam voor de variabelen respectievelijk huistijd en sekse.



- Klik op de lege cel onder Name.
- Voer vervolgens de naam van de variabele in (hier *huistijd*) en druk op **Enter** op het toetsenbord.
- Dubbelklik op de lege cel onder *huistijd* en vul *sekse* in en druk weer op **Enter**.

Het resultaat staat in afbeelding 2.2. Merk op dat voor beide variabelenamen alleen letters zijn gebruikt. In SPSS moet een naam voor een variabele altijd