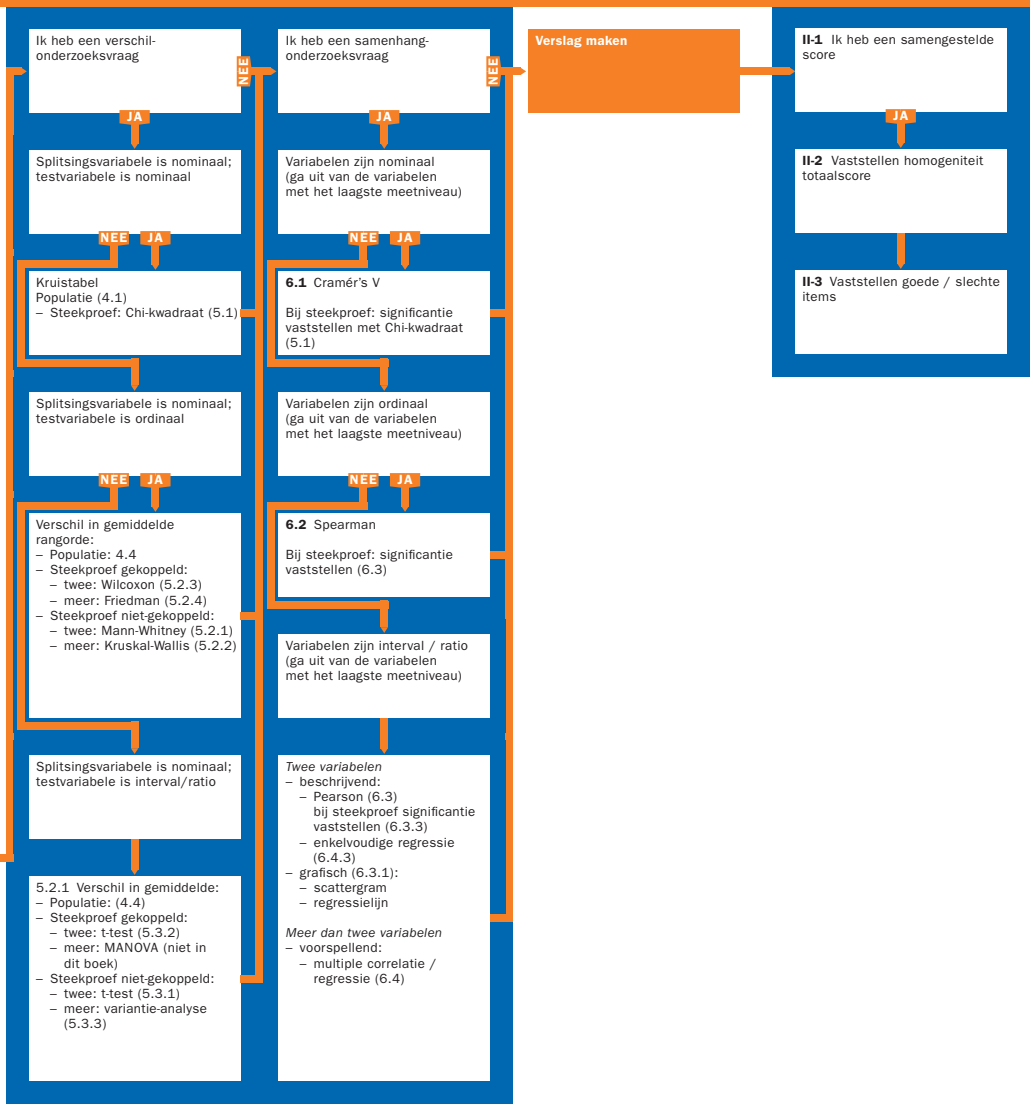
Ben Baarda
René van Vianen

Basisboek
Statistiek met Excel

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

Eerste druk

Noordhoff Uitgevers


Ik heb een verschil-onderzoeksvraag
NEE
JA
Splitsingsvariabele is nominaal; testvariabele is nominaal
NEE
JA
Kruistabel
Populatie (4.1)
– Steekproef: Chi-kwadraat (5.1)
Splitsingsvariabele is nominaal; testvariabele is ordinaal
NEE
JA
Verschil in gemiddelde rangorde:
– Populatie: 4.4
– Steekproef gekoppeld:
– twee: Wilcoxon (5.2.3)
– meer: Friedman (5.2.4)
– Steekproef niet-gekoppeld:
– twee: Mann-Whitney (5.2.1)
– meer: Kruskal-Wallis (5.2.2)
Splitsingsvariabele is nominaal; testvariabele is interval/ratio
5.2.1 Verschil in gemiddelde:
– Populatie: (4.4)
– Steekproef gekoppeld:
– twee: t-test (5.3.2)
– meer: MANOVA (niet in dit boek)
– Steekproef niet-gekoppeld:
– twee: t-test (5.3.1)
– meer: variantie-analyse (5.3.3)
Ik heb een samenhang-onderzoeksvraag
NEE
JA
Variabelen zijn nominaal (ga uit van de variabelen met het laagste meetniveau)
NEE
JA
6.1 Cramér's V
Bij steekproef: significantie vaststellen met Chi-kwadraat (5.1)
Variabelen zijn ordinaal (ga uit van de variabelen met het laagste meetniveau)
NEE
JA
6.2 Spearman
Bij steekproef: significantie vaststellen (6.3)
Variabelen zijn interval / ratio (ga uit van de variabelen met het laagste meetniveau)
Twee variabelen
– beschrijvend:
– Pearson (6.3) bij steekproef significantie vaststellen (6.3.3)
– enkelvoudige regressie (6.4.3)
– grafisch (6.3.1):
– scattergram
– regressielijn
Meer dan twee variabelen
– voorspellend:
– multiple correlatie / regressie (6.4)
Verslag maken
II-1 Ik heb een samengestelde score
JA
II-2 Vaststellen homogeniteit totaalscore
II-3 Vaststellen goede / slechte items

Basisboek Statistiek met Excel

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

Ben Baarda
René van Vianen

Eerste 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

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 1 2 3 4 5 / 15 14 13 12 11

ISBN (ebook) 978-90-01-84404-2
ISBN 978-90-01-79637-2
NUR 916

Voorwoord bij Basisboek *Statistiek met Excel*

In dit boek *Basisboek Statistiek met Excel* vind je aanwijzingen voor het verwerken en analyseren van onderzoeksgegevens met behulp van Excel. Doordat Excel standaard op vrijwel alle computers staat, kun je vrijwel overal je databestanden analyseren. Je hebt daar dus geen dure statistische software voor nodig, tenzij je zeer geavanceerde statistische technieken wilt gebruiken. Als je de 'Analysis Toolpak' invoegtoepassing van Excel installeert, uitleg in paragraaf 1.6, beschik je over de meest gangbare statistische analyseprocedures; ook over multivariate technieken als multivariate variantieanalyse.

In dit boek gebruiken wij de 2007 versie. Wil je meer weten over Excel kijk dan eens op de Wikipedia pagina over Excel (http://en.wikipedia.org/wiki/Microsoft_Excel). Daar vind je niet alleen info over wat er mogelijk is met Excel, maar ook de geschiedenis.

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

Na een algemene inleiding en kennismaking met de basisregels van Excel kun je direct beginnen met het invoeren en analyseren van gegevens in de computer. Om verantwoord statistische procedures met Excel te kunnen uitvoeren is niet alleen kennis van Excel, maar ook statistische kennis nodig. We bespreken in de eerste hoofdstukken statistische basisbegrippen als meetniveau, normaalverdeling, kans, significantie, één- en tweezijdige toetsing, power 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 de techniek wel kunt gebruiken en wanneer niet. Uiteraard wordt uitgelegd hoe je die techniek in Excel uitvoert. Soms maken we ook uitstapjes naar internet, omdat sommige statistische procedures makkelijker via interactieve internetprogramma's zijn uit te voeren. Wel bereiden we het databestand altijd voor in Excel. Via schermafdrukken en de instructies die we daarbij geven, laten wij letterlijk 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. Tevens geven we aan hoe je over het resultaat kunt rapporteren.

De databestanden waarop we de analyses uitvoeren, zijn te vinden op internet via www.basisboekstatistiekmetexcel.noordhoff.nl

Het *Basisboek Statistiek met Excel* komt optimaal tot zijn recht in samenhang met het *Basisboek Methoden en Technieken* of *Dit is onderzoek*! Aanvullend kan ook het *Statistiekkwartetspel* gebruikt worden, waarin op een speelse wijze de belangrijkste statistische begrippen en toetsen worden uitgelegd. *Basisboek Statistiek met Excel* is uiteraard ook los daarvan goed te gebruiken.

November 2010
Ben Baarda
René van Vianen

Inhoud

Effectief studeren

Opbouw boek

Hoofdstuk 1
Openen Excel
+
Statistische toelichting

▼

Hoofdstuk 2
Invoer van gegevens

▼

Hoofdstuk 3
Het veranderen en
combineren van gegevens

▼ ▼ ▼

Hoofdstuk 4	Hoofdstuk 5	Hoofdstuk 6
Frequenties	Verschillen	Samenhang

Bijlage Controle op homogeniteit samengestelde scores.

Opbouw hoofdstuk

Voorkennis
We gaan ervan uit dat:
- Je bekend bent met de verschillende meetniveaus (1.3.2).
- Je Excel kunt starten (1.5).
- Je een databestand kunt oproepen (2.5).
- Je weet wat continue en discontinue variabelen zijn (1.3.2).
- Je weet wat een normaalverdeling is (1.3).

Dit is wat je moet weten als je aan het hoofdstuk begint.

Dit zijn de vragen die in het hoofdstuk worden behandeld.

81

5
Hoe analyseer ik mijn gegevens bij een verschilonderzoeksvraag?

In hoofdstuk 5 worden de volgende vragen beantwoord en de volgende begrippen behandeld:
5.1 Verschil bij een nominale test- en splitsingsvariabele?
5.2 Verschilvragen bij ordinale testvariabelen en nominale splitsingsvariabelen
5.3 Verschilvraag bij interval- / ratiotestvariabele en nominale splitsingsvariabele?.

- Kruistabel
- Chikwadraattoets
- Waargenomen frequentie
- Verwachte frequentie
- Vrijheidsgraden
- Splitsingsvariabele
- Testvariabele
- Rijpercentages
- Kolompercentages
- Mann-Whitney U-toets
- Rangscore
- U-waarde
- Niet-gekoppelde of onafhankelijke steekproeven
- T-toets - Kruskall-Wallis-toets
- Afhankelijke steekproeven
- Effectgrootte
- T-waarde
- Wilcoxon Signed-ranks-toets
- Friedman-toets
- Ongelijke variantie
- Cohen's d


11

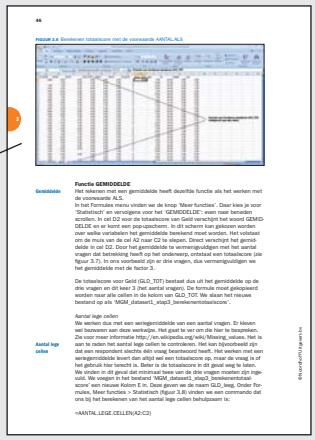
1
Hoe bereid ik mij voor op statistiek met Excel?

In hoofdstuk 1 worden de volgende vragen beantwoord en de volgende begrippen behandeld:
1.1 Inleiding: welke onderwerpen worden in dit boek behandeld?
1.2 Geld en geluk: toelichting op de gebruikte casus
1.3 Hoe analyseer ik mijn data? Een gebruiksaanwijzer!
1.4 Enkele algemene statistische begrippen
1.5 Hoe werkt Excel?
1.6 Installeren van de statistische functies van Excel.

- Frequentie
- Verschil
- Samenhang
- Nominaal
- Ordinaal
- Interval
- Ratio
- Nulpunt
- Scale
- Continu
- Discreet
- Beschrijvende statistiek
- Inductieve/inferentiële statistiek
- Normaalverdeling
- Standaardfout
- Significantie
- Eén- of tweezijdig toetsen
- Relevantie
- Effectgrootte
- Vrijheidsgraden
- Formule

© Noordhoff Uitgevers bv

Dit zijn de begrippen die in een hoofdstuk worden behandeld.


46

FIGUUR 2.2 Berekenen totaalscore met de voorwaarde AANTAL.ALS

Gemiddelde

Functie GEMIDDELDE
Het rekenen met een gemiddelde heeft dezelfde functie als het werken met de voorwaarde ALS.
In het Formules menu vinden we de knop 'Meer functies'. Daar kies je voor 'Statistisch' en vervolgens voor het 'GEMIDDELDE': een naar beneden scrollen. In cel D2 voor de totaalscore van Geld verschijnt het woord GEMIDDELDE en er komt een pop-upscherm. In dit scherm kan gekozen worden over welke variabelen het gemiddelde berekend moet worden. Het volstaat om de muis van de cel A2 naar C2 te slepen. Direct verschijnt het gemiddelde in cel D2. Door het gemiddelde te vermenigvuldigen met het aantal vragen dat betrekking heeft op het onderwerp, ontstaat een totaalscore (zie figuur 3.7). In ons voorbeeld zijn er drie vragen, dus vermenigvuldigen we het gemiddelde met de factor 3.

De totaalscore voor Geld (GLD_TOT) bestaat dus uit het gemiddelde op de drie vragen en dit keer 3 (het aantal vragen). De formule moet gekopieerd worden naar alle cellen in de kolom van GLD_TOT. We slaan het nieuwe bestand op als 'MGM_dataset1_stap3_berekenentotaalscore'.

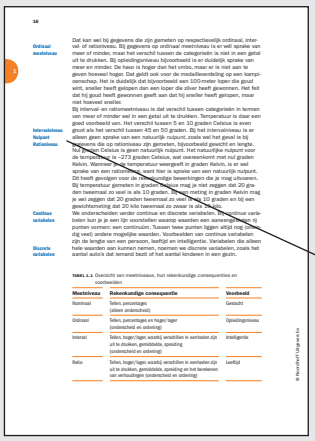
Aantal lege cellen
We werken dus met een wetenschappelijke van een aantal vragen. Er kleven wel bezwaren aan deze werkwijze. Het gaat te ver om die hier te bespreken. Zie voor meer informatie http://en.wikipedia.org/wiki/Missing_values. Het is aan te raden het aantal lege cellen te controleren. Het kan bijvoorbeeld zijn dat een respondent slechts één vraag beantwoord heeft. Het werken met een aantal gemiddelde levert dan altijd wel een totaalscore op, maar de vraag is of het gebruik hier terecht is. Beter is de totaalscore in dit geval weg te laten. We vinden in dit geval dat minimaal twee van de drie vragen moeten zijn ingevuld. We voegen in het bestand 'MGM_dataset1_stap3_berekenentotaalscore' een nieuwe Kolom E in. Deze geven we de naam GLD_weg. Onder Formules, Meer functies > Statistisch (figuur 3.8) vinden we een commando dat ons bij het berekenen van het aantal lege cellen behulpzaam is:

Aantal lege cellen

=AANTAL.LEGE.CELLEN(A2:C2)

© Noordhoff Uitgevers bv

Dit is een voorbeeld van een schermafdruk om uit te leggen hoe je Excel moet gebruiken.


16

Ordinaal meetniveau
Intervalniveau
Nulpunt
Rationeau
Continue variabelen
Discrete variabelen

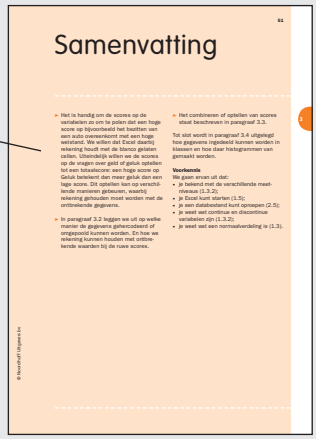
Dat laat wel bij gegevens die zijn gemeten op respectievelijk ordinaal, interval- of rationiveau. Bij gegevens op ordinaal meetniveau is er wél sprake van meer of minder, maar het verschil tussen de categorieën is niet in een getal uit te drukken. Bij opleidingsniveau bijvoorbeeld is er duidelijk sprake 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 kompinioenschap. 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.
Bij interval- en rationeelniveau 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 intervalniveau is er echter geen sprake van een natuurlijk nulpunt, zoals wel het geval is bij lengtes die op rationeelniveau zijn gemeten, bijvoorbeeld gewicht en lengte. 0 Graden Celsius is geen natuurlijk nulpunt. Het natuurlijke nulpunt voor de temperatuur is –273 graden Celsius, wat overeenkomt met nul graden Kelvin. Wanneer de temperatuur weergeeft in graden Kelvin, is er wel sprake van een rationeelniveau, want hier is sprake van een natuurlijk nulpunt. Dit heeft gevolgen voor de conclusies die bewerkingen die je mag uitvoeren. Bij temperatuur gemeten in graden Celsius mag je niet zeggen dat 20 grs den 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.
We onderscheiden verder continue en discrete variabelen. Continue variabelen kun je je een lijn voorstellen waarop waarden een aaneengesloten rij punten vormen: een continuüm. Tussen twee punten liggen altijd nog (theoretisch) weer andere mogelijke waarden. Voorbeelden van continue variabelen zijn de lengte van een persoon, leeftijd en intelligentie. Variabelen die alleen hele waarden aan kunnen nemen, noemen we discrete variabelen, zoals het aantal auto's dat iemand bezit of het aantal kinderen in een gezin.

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

© Noordhoff Uitgevers bv

Margewoorden om de belangrijkste begrippen te benadrukken.


51

Samenvatting

- Het is handig om de scores op de variabelen zo om te polen dat een hoge score op bijvoorbeeld het bezitten van een auto overeenkomt met een hoge welstand. We willen dat Excel daarbij rekening houdt met de blanco gelaten cellen. Uiteindelijk willen we de scores op de vragen over geld of geluk optellen tot een totaalscore: een hoge score op Geluk betekent dan meer geluk dan een lage score. Dit optellen kan op verschillende manieren gebeuren, waarbij rekening gehouden moet worden met de ontbrekende gegevens.
- In paragraaf 3.2 leggen we uit op welke manier de gegevens gehercodeerd of omgepoold kunnen worden. En hoe we rekening kunnen houden met ontbrekende waarden bij de nieuw scores.
- Het combineren of optellen van scores staat beschreven in paragraaf 3.3.

Tot slot wordt in paragraaf 3.4 uitgelegd hoe gegevens ingedeeld kunnen worden in klassen en hoe daar histogrammen van gemaakt worden.

Voorkennis
We gaan ervan uit dat:
- je bekend met de verschillende meetniveaus (1.3.2);
- je Excel kunt starten (1.5);
- je een databestand kunt opnepen (2.5);
- je weet wat continue en discontinue variabelen zijn (1.3.2);
- je weet wat een normaalverdeling is (1.3).

© Noordhoff Uitgevers bv

Dit is een voorbeeld van een samenvatting die aan het eind van ieder hoofdstuk staat.

Opbouw paragraaf

- Wat is ...?
- Wanneer gebruik ik ...?
- Hoe voer ik ... uit?
- Hoe lees ik de uitvoer van ...?
- Hoe rapporteer ik over ...?

Voorkennis

Er wordt geen specifieke voorkennis vereist. Wel wordt aanbevolen om naast dit boek het *Basisboek Methoden en Technieken* (4e herziene druk, 2006) of *Dit is onderzoek!* te gebruiken.

1 Hoe bereid ik mij voor op statistiek met Excel?

In hoofdstuk 1 worden de volgende vragen beantwoord en de volgende begrippen behandeld:

1.1 Inleiding: welke onderwerpen worden in dit boek behandeld?
1.2 Geld en geluk; toelichting op de gebruikte casus
1.3 Hoe analyseer ik mijn data? Een gebruikswijzer!
1.4 Enkele algemene statistische begrippen
1.5 Hoe werkt Excel?
1.6 Installeren van de statistische functies van Excel.

- Frequentie 17
- Verschil 17
- Samenhang 17
- Nominaal meetniveau 17
- Ordinaal meetniveau 18
- Intervalniveau 18
- Rationiveau 18
- Nulpunt 18
- Continue variabelen 18
- Discrete variabelen 18
- Beschrijvende statistiek 19
- Inductieve/Inferentiële statistiek 19
- Normaalverdeling 19
- Standaardfout 20
- Significantie 21
- Eén- of tweezijdig toetsen 21
- Relevantie 21
- Effectgrootte 21
- Vrijheidsgraden 21
- Formule 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 keten van beslissingen die je moet nemen. Het uiteindelijke doel is het beantwoorden van de onderzoeksvraag of -vragen.
Om de plaats van de data-analyse in de onderzoekscyclus als geheel voor ogen te houden, zetten we de gebruikelijke fasen van een onderzoek hier op een rij. Iedere fase van de onderzoekscyclus is in de vorm van een vraag opgenomen:

1 Wat is mijn onderzoeksvraag of vragen en wat mijn 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 dataverzamelingsmethode 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 9 centraal: preparatie, analyse en beschrijving van de onderzoeksgegevens die met Excel geanalyseerd worden. Aan de hand van een onderzoek naar de relatie tussen geld en geluk (zie figuur 1.1), behandelen we stap voor stap de volgende punten:

- Hoe je onderzoeksgegevens moet prepareren om ze te kunnen invoeren (hoofdstuk 2).
- Hoe je gegevens met Excel kunt aanpassen en veranderen. Voordat je aan de analyse begint moet je vaak eerst van bepaalde gegevens de waarden ompolen of hercoderen, of de waarden van gegevens combineren tot een nieuwe score (hoofdstuk 3).
- Hoe je de juiste analysetechniek kiest. Om de juiste analysetechniek te kunnen kiezen moet je eerst vaststellen wat het karakter van je onderzoeksvraag is (paragraaf 1.3.1). Gaat het om frequenties (hoofdstuk 4), verschillen (hoofdstuk 5) of om samenhangen (hoofdstuk 6)? Vervolgens moet je nagaan wat het meetniveau van je gegevens is (paragraaf 1.3.2). Tot slot moet je vaststellen 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 beste bij jouw onderzoeksvraag past. Je vindt het schema op de binnenkant van de kaft en als los inlegvel in dit boek.
- Hoe je de analyse met Excel moet uitvoeren en hoe je de resultaten moet interpreteren. Dit geven we telkens aan bij iedere 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 als normaalverdeling, significantie, effectgrootte, kans en standaardfout (paragraaf 1.4). In paragraaf 1.5 leggen we uit hoe je Excel kunt starten en in de laatste paragraaf (1.6) tonen we je hoe je de ‘Analysis Tool Pak’ invoegtoepassing installeert. Hierdoor kun je allerlei statistische analysetechnieken, zoals de variantieanalyse, gebruiken.

1.2 Geld en geluk; toelichting op de gebruikte casus

Voordat we ingaan op het gebruik van Excel, introduceren we eerst het voorbeeldonderzoek Geld en Geluk (zie figuur 1.1). De vragen uit figuur 1.1 zijn voorgelegd aan een representatieve steekproef van 500 Nederlandse mannen en 500 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 boven de 55 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 de website http://basisboekstatistiekmetexcel.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	ja/nee
2	Ik heb een koopwoning/flat	ja/nee
3	Ik bezit een spelcomputer	ja/nee
4	Ik krijg zorgtoeslag	ja/nee
5	Ik krijg huursubsidie	ja/nee

Geluk is gemeten met de volgende vragen:

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

☐	☐	☐	☐	☐
absoluut niet	niet	ten dele wel/niet	wel	absoluut

2 De meeste anderen hebben het beter dan ik.

☐	☐	☐	☐	☐
absoluut niet	niet	ten dele wel/niet	wel	absoluut

3 Ik heb het ... naar mijn zin.

☐	☐	☐	☐	☐
absoluut niet	niet	ten dele wel/niet	wel	absoluut

4 Het leven is zwaar.

☐	☐	☐	☐	☐
absoluut niet	niet	ten dele wel/niet	wel	absoluut

5 Ik voel mij eenzaam.

☐	☐	☐	☐	☐
absoluut niet	niet	ten dele wel/niet	wel	absoluut

FIGUUR 1.1 Voorbeeldonderzoek Geld en Geluk (vervolg)

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 (lagere-school, 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 zijn er in het bezit van respectievelijk een auto, koopwoning, spelcomputer?
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, koopwoning en spelcomputer; het ontvangen van zorgtoeslag; het ontvangen van huurtoeslag?
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 Bestaat er een verschil in eenzaamheid tussen mannen en vrouwen?
9 Is er tussen mannen en vrouwen een verschil in tevredenheid over het leven dat men leidt?
10 Is er een relatie tussen de mate van tevredenheid over het leven en 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 paragrafen 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 waar een antwoord op moet worden gegeven. Globaal zijn er drie soorten onderzoeksvragen te onderscheiden:

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?' **Frequentie**
2 Vragen waarbij het gaat om een *verschil*. Bijvoorbeeld: 'Zijn mannen gelukkiger dan vrouwen?' **Verschil**
3 Vragen waarbij het gaat om een *samenhang*. Bijvoorbeeld: 'Is er een samenhang tussen geld en geluk?' **Samenhang**

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 zijn verschilonderzoeksvragen. De analyse van gegevens bij dit type onderzoeksvraag behandelen we in hoofdstuk 5. De onderzoeksvragen 10 en 11 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 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. Per onderzoeksvraag moet je aangeven wat het meetniveau van de betreffende variabelen is. 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. 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**

Ordinaal 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.

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 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 *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. 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.

Continue variabelen

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.

Discrete variabelen

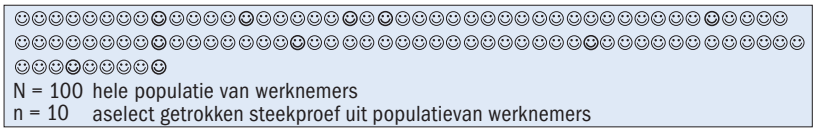
Variabelen die alleen hele waarden aan kunnen nemen, noemen we *discrete variabelen*, zoals het aantal auto's dat iemand bezit of het aantal kinderen in een gezin.

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 aselecte steekproef



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 waar je uitspraken over 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. Bij een steekproef is het wel de opzet dat je uitspraken doet over de totale populatie van werknemers. In dit geval zul je gebruik moeten maken van de *inductieve of inferentiële statistiek*: je wilt 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)

Beschrijvende statistiek

Inductieve of inferentiële statistiek

Voordat je aan de analyse van je gegevens begint moet je jezelf de vraag stellen over welke eenheden (wie of wat) je uitspraken wilt doen. Wanneer dat alleen de 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 paragraaf 1.4 behandelen we in het kort enkele statistische termen die je steeds 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.

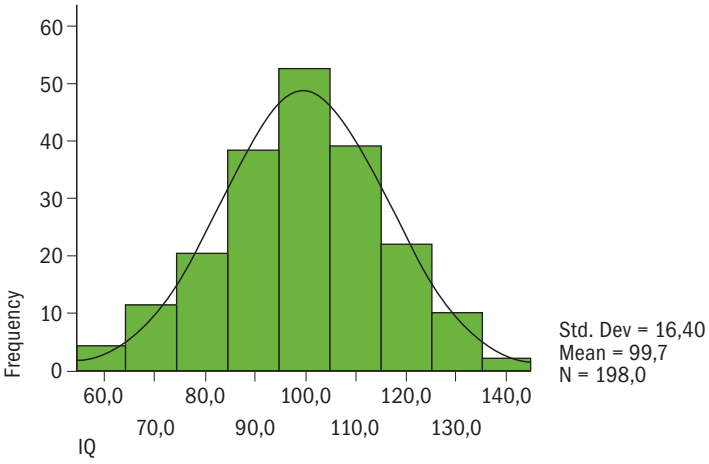
1.4 Enkele algemene statistische begrippen

Zoals in paragraaf 1.3 is aangegeven, zijn er twee vormen van statistiek: de beschrijvende en de inductieve of inferentiële statistiek. Het doel van de beschrijvende statistiek is om 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 bijvoorbeeld een histogram (paragraaf 4.1.4) of in de vorm van percentages of een gemiddelde (paragraaf 4.2). 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.

Normaalverdeling

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



Dit is de verdeling van de scores op een intelligentietest 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 linkerhelft met de rechterhelft.

Wanneer de 198 werknemers een aselecte 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, in het geval 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 aselecte 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 Excel kun je de standaardfout laten berekenen (paragraaf 4.2.2). De *standaardfout* geeft aan in hoeverre het gevonden steekproefgemiddelde een *betrouwbare* schatting is 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 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.

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 Excel-uitdraai ook bij of er *één- of tweezijdig* getoetst is. Je toetst eenzijdig (*one-tailed*) wanneer je een hypothese of verwachting hebt geformuleerd. 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*).

Significantie

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. 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 twee procent, dus inderdaad kleiner dan vijf procent. In hoofdstuk 5 leggen we uit hoe je dat berekent. 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 één procent van de verschillen in IQ. Hoe je dit uitrekent leggen we uit in paragraaf 5.3.1.

Effectgrootte

In de Excel-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 een kent (namelijk 36) en je weet dat het gemiddelde 40 is, dan moet het andere getal 44 zijn. Je hebt hier 1 vrijheidsgraad. Als je namelijk het ene getal weet, weet je het andere ook. Bij veel toetsen, zoals de t-toets (paragraaf 5.3), is het aantal vrijheidsgraden het aantal steekproefelementen minus 1. Bij een kruistabel (zie paragraaf 5.1) 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.

Vrijheidsgraden

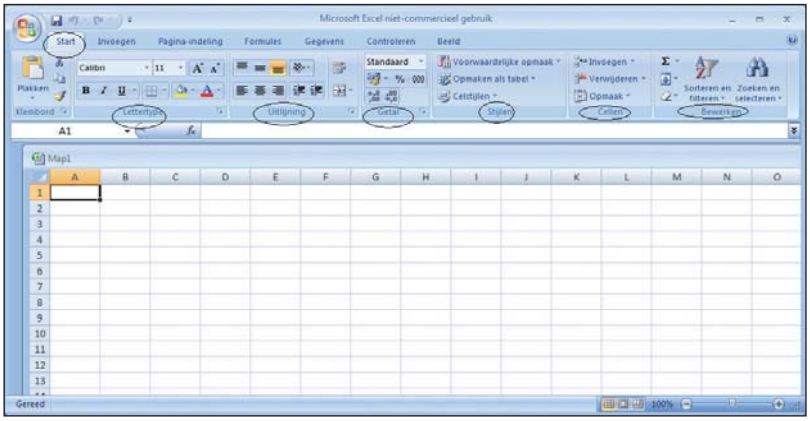
1.5 Hoe werkt Excel?

Zoals in de inleiding al aangegeven, is Excel niet alleen een middel om data op te slaan en te ordenen, maar biedt het ook veel mogelijkheden voor statistische analyse. Als Excel 2007 is geïnstalleerd vind je, als het goed is, het volgende icoon op je scherm (figuur 1.4):

FIGUUR 1.4 Het Microsoft Excel 2007-icoon

Je 'opent' Excel door op dit Microsoft Office Excel 2007 icoon te dubbelklikken. Als je het icoon niet kunt vinden, ga je naar de startknop linksonder en vervolgens naar programma's. Je krijgt dan een overzicht van de programma's die op je computer staan. Als Microsoft Office in het rijtje voorkomt, klik je daarop en zie je dat Excel een onderdeel is van Microsoft Office.
Je kunt Excel starten door er op te dubbelklikken.

FIGUUR 1.5 Het openingsscherm van Excel 2007



Je ziet dan het openingsscherm zoals dat in figuur 1.5 is weergegeven. Je bevindt je nu in het *startmenu*. Dit menu kent een aantal submenu's:

- *Lettertype*; hiermee kies je voor een lettertype, grootte, stijl, kleur en dergelijke.
- *Uitlijning*; hiermee kun je de informatie in de cel uitlijnen, bijvoorbeeld links en boven, of juist in het midden van de cel.
- *Getal*; dit staat standaard ingesteld op 'standaard', maar je kunt ook kiezen voor getal, valuta, datum en percentage, al naar gelang het karakter van je gegevens. Voor de meeste statistische analyses is het handig als de *getalinstelling* gebruikt wordt.
- *Stijlen*; hier kun je kiezen voor een specifieke opmaak van de cel, het is soms handig om onderscheid te maken tussen gegevens die zijn ingevoerd en gegevens die je berekend of bewerkt hebt. Door te kiezen voor twee verschillende celstijlen, zie je direct het verschil.

- *Cellen*; je kunt hiermee cellen invoegen, verwijderen of opmaken. In een cel kan je bijvoorbeeld rekenkundige formules zetten. Je kan zo twee cellen bij elkaar optellen en het resultaat daarvan komt in een derde cel. Het werken met formules komt uitgebreid aan bod. In bijlage I staan de gebruikte commando's voor de formules uitgewerkt.
- *Bewerken*; je kunt daar, zoals de titel al aangeeft gegevens zoeken, sorteren, bewerken, zoals het uitrekenen van een totaalscore en gegevens verwijderen. We komen op de meeste functies in de later hoofdstukken terug.

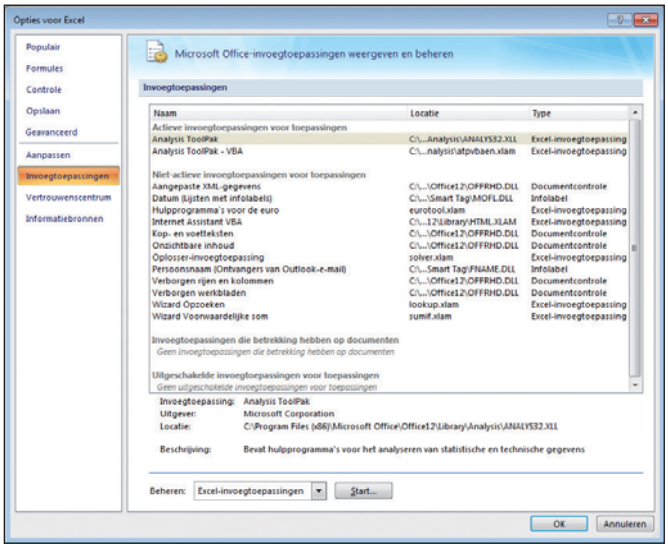
Een aantal van de specifieke Excel-functies, zoals de formulefunctie, zullen in dit boek uitgebreid aan bod komen. Een aantal algemene functies, zoals knippen, plakken, kopiëren, kolommen aanpassen, celeigenschappen veranderen, bespreken wij in hoofdstuk 2. Ben je erg onhandig met Excel, dan verdient het misschien aanbeveling om eerst een korte cursus Excel op internet te volgen. Je kunt dan bijvoorbeeld naar de website http://www.gratiscursus.be/excel_2007/ gaan.

1.6 Installeren van de statistische functies van Excel

Wanneer je ooit gekozen hebt voor de standaardconfiguratie van het pakket Office, dan worden de statistische mogelijkheden van Excel niet direct geïnstalleerd. Zij zijn wel onderdeel van Excel. Deze statistische functies heb je later nodig wanneer we gaan indelen in klassen, samenhangen gaan analyseren et cetera.

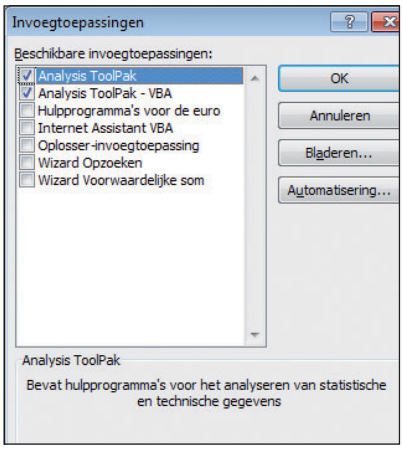
Onder de Officeknop  vind je rechtsonder het commando Opties voor Excel. Klik hier op en daarna in het menu dat verschijnt op 'Invoegtoepassingen' (figuur 1.6).

FIGUUR 1.6 Het invoegtoepassingenmenu in Excel



Klik dan op Start onderaan bij het beheren van de ‘Invoegtoepassingen’. Vink daarna de invoegtoepassingen ‘Analysis ToolPak’ en ‘Analysis ToolPak VBA’ aan (figuur 1.7).

FIGUUR 1.7 Het installeren van de ‘Analysis Toolpak’



Excel installeert vervolgens de benodigde statistische functies. In het tabblad Gegevens is nu een nieuw tabblad toegevoegd: Gegevensanalyse. Statistiek met Excel kan beginnen!

Samenvatting

- Het prepareren en analyseren van onderzoeksgegevens is een van de laatste fasen in de onderzoekscyclus. De eindfase, de rapportage, is hier voor een groot deel op gebaseerd.
- In paragraaf 1.2 lichten wij de casus toe, die wij gebruiken om het uitvoeren van statistische analyse uit te leggen.
- In paragraaf 1.3 geven wij aan dat je voordat je gaat analyseren, een aantal vragen moet beantwoorden. Gaat het bijvoorbeeld in je onderzoeksvraag om een frequentie, een verschil, of om een samenhang? Wat is het meetniveau van je variabelen? Gaat het om een steekproef of om een populatie? Afhankelijk van het antwoord op deze vragen kun je met behulp van het blokschema een keuze maken voor een statistische analysetechniek.
- In paragraaf 1.4 leggen wij enkele gangbare statistische begrippen uit. Het gaat daarbij om begrippen als standaardfout, betrouwbaarheid, kans, significantie, één- en tweezijdig toetsen, effectgrootte en vrijheidsgraden.
- In paragraaf 1.5 geven wij in het kort aan hoe Excel werkt.
- Voor de statistische analyses is het belangrijk dat je de statistische functies in Excel installeert. Hoe je dat doet, wordt uitgelegd in paragraaf 1.6.