

# 1. Starten met het onderzoeksproces

## Casus

Suxes is marktleider op de Nederlandse cateringmarkt. Het bedrijf is succesvol in meerdere marktsegmenten. Voor het hogeronderwijssegment heeft het een concept ontwikkeld dat goed aanslaat in onderwijsland. Uitgangspunten hierbij zijn de wensen en behoeften van de onderwijsinstelling en die van de grootste groep gebruikers: de studenten. Dit heeft bij Suxes geleid tot een formule die gekarakteriseerd kan worden als 'value for money' in een sfeervol campusrestaurant. In het midden van Nederland is de Pandion Hogeschool sinds anderhalf jaar gehuisvest in een nieuw gebouw. Alle opleidingen zijn in dit nieuwe gebouw gesitueerd. Er is sprake van een groot en splinternieuw campusrestaurant. Bij het tot stand komen van het nieuwe onderwijscomplex is besloten de catering uit te besteden aan Suxes. Bij Pandion biedt Suxes naast het basisassortiment (broodjes, soep, koffie, thee, melk) ook een wat luxer assortiment aan. Dit luxe assortiment bestaat onder meer uit producten uit de saladebar, rijk gevulde soepen en extra luxe broodjes. Nu, anderhalf jaar nadat Suxes begonnen is met de cateringactiviteiten bij Pandion, wil de Raad van Bestuur onder de studenten een tevredenheidsonderzoek uit laten voeren.

## 1.1 Het onderzoeksproces

Het uitvoeren van het hiervoor genoemde tevredenheidsonderzoek betekent dat er een **onderzoeksproces** doorlopen moet worden. In dit proces kunnen de verschillende fases schematisch weergegeven worden:

1. *Formuleren van de probleemstelling.*
2. *Bepalen van de exacte informatiebehoefte: de onderzoeksdeelvragen.*
3. *Bepalen van de onderzoeksopzet.*
4. *Uitvoeren van desk research.*

5. *Uitvoeren van field research: kwalitatief en/of kwantitatief.*
6. *Verwerken van gegevens.*
7. *Analyseren van gegevens.*
8. *Interpreteren van gegevens, trekken van conclusies, formuleren van aanbevelingen.*
9. *Schrijven van het onderzoeksverslag.*

Dit boek richt zich met name op het verwerken, analyseren en interpreteren van enquêtegegevens met behulp van het computerpakket SPSS. Het past niet in de opzet van het boek om alle fases in het onderzoek afzonderlijk en uitgebreid te beschrijven. Hiervoor verwijzen we naar uitstekende boeken van collega-auteurs.

Nadat de probleemstelling voor het onderzoek is vastgesteld (fase 1) zal daarna (fase 2) op detailniveau worden bepaald wat de exacte informatiebehoefte is. Een deel van de geformuleerde onderzoeksvragen zal beantwoord kunnen worden met desk research of kwalitatieve field research in de vorm van bijvoorbeeld diepte-interviews of focus-groepgesprekken. Op de onderzoeksvragen die beantwoord kunnen worden via het uitvoeren van kwantitatieve field research zal veelal een antwoord worden verkregen door een enquête op te stellen, af te nemen en vervolgens de resultaten te analyseren.

Voor het tevredenheidsonderzoek van deze casus zal men dus onder studenten van Pandion gegevens verzamelen via een vragenlijst. Deze vragenlijst is gebaseerd op de volgende probleemstelling en onderzoeksvragen:

### **Probleemstelling**

Op welke wijze kan Suxes de klanttevredenheid verhogen door beter in te spelen op wensen en behoeftes van de studenten van Pandion?

### **Onderzoeksdeelvragen (field research)**

- (1) Wat is het aantal dagen per week dat men gebruikmaakt van het restaurant?
- (2) Welk bedrag besteedt men per week in het restaurant?
- (3) Hoe beoordeelt men het restaurant op diverse aspecten?
  - a. Keuzemogelijkheden basisassortiment
  - b. Keuzemogelijkheden luxe assortiment
  - c. Klantvriendelijkheid van het personeel
  - d. Prijs/kwaliteitsverhouding
  - e. Hygiëne
  - f. Presentatie van het productaanbod.

- (4) Welke producten koopt men in het restaurant?
- (5) Is er verschil tussen mannelijke en vrouwelijke studenten wat betreft de aangekochte producten?
- (6) Wie is de klant (studierichting, aantal jaren dat de student is ingeschreven, geslacht, woonsituatie)?
- (7) Wat is het totaaloordeel over de catering uitgedrukt in een rapportcijfer?
- (8) Is er een verschil in waardering van het restaurant tussen studenten van de verschillende studierichtingen?
- (9) Is er een verschil in waardering van het restaurant tussen mannelijke en vrouwelijke studenten?
- (10) Is er een verschil in bestedingsbedrag tussen thuiswonende en uitwonende studenten?
- (11) Is er een verschil in bestedingsbedrag tussen studenten van de verschillende studierichtingen?
- (12) Is er een verschil in totaaloordeel tussen studenten van de verschillende studierichtingen?
- (13) Is er een samenhang tussen totaaloordeel dat studenten geven en het bestedingsbedrag?
- (14) ...
- (15) ...

Op basis van de onderzoeksdeelvragen voor de field research wordt een vragenlijst (enquête) opgesteld. Het handmatig verwerken van de enquêteresultaten is hierbij vaak geen optie. Meestal zal de hoeveelheid gegevens zodanig zijn dat we de hulp inroepen van een statistisch programma: SPSS.

Steeds vaker zien we dat vragenlijsten via het internet worden afgenomen. Er zijn diverse firma's en websites die softwareproducten aanbieden waarmee je dit soort web-enquêtes kunt realiseren. Je kunt met deze software de vragenlijst ontwerpen, online zetten en via de mail de deelnemers uitnodigen. Tijdens het onderzoek kun je online van dag tot dag, of zelfs op continue basis, de voortgang volgen. Als ten slotte iedereen heeft deelgenomen, exporteer je alle gegevens eenvoudig naar SPSS.

Als voorbeeld zullen we de gang van zaken schetsen aan de hand van de volgende enquête, die de onderzoekers hebben opgesteld naar aanleiding van het tevredenheids-onderzoek bij Pandion.

1. Hoeveel dagen per week bezoek je (gemiddeld) het restaurant?

..... dagen per week (1, 2, 3, 4 of 5 invullen)

2. Welk bedrag besteed je gemiddeld per week in het restaurant?

..... euro per week

3. Geef een oordeel over de volgende zaken van het restaurant.

	1=Goed	2=Voldoende	3=Onvoldoende	4=Slecht
Keuzemogelijkheden basisassortiment	1	2	3	4
Keuzemogelijkheden luxe assortiment	1	2	3	4
Klantvriendelijkheid van het personeel	1	2	3	4
Prijs/kwaliteitsverhouding	1	2	3	4
Hygiene	1	2	3	4
Presentatie van de producten	1	2	3	4

4. Welke producten koop je wel eens in het restaurant?

(meerdere antwoorden mogelijk)

- Broodje ham/kaas
- Een ander broodje
- Zuivelproducten
- Koffie of thee
- Soep
- Salades

5. Als ik hogeschoolcateraar Suxes met een rapportcijfer (1 t/m 10) moet beoordelen dan geef ik het cijfer:

.....

Ten slotte graag een antwoord op een aantal persoonlijke vragen:

6. Ik doe bij Pandion een studie op het gebied van:

- Economie
- Zorg
- Techniek
- Anders

7. Hoeveel jaar ben je als student ingeschreven bij Pandion?

..... jaar

8. Wat is je geslacht?

- Vrouw
- Man

9. Wat is je woonsituatie?

- Thuiswonend
- Uitwonend

**Figuur 1.1** De vragenlijst voor het tevredenheidsonderzoek van Suxes onder studenten van Pandion

## 1.2 Soorten enquêtevragen

Wanneer de gegevens die verzameld zijn met het afnemen van een vragenlijst met SPSS worden verwerkt, is het noodzakelijk dat iedere vraag wordt omgezet in een variabele. Bij deze omzetting is het van belang dat we onderscheid maken tussen verschillende soorten enquêtevragen:

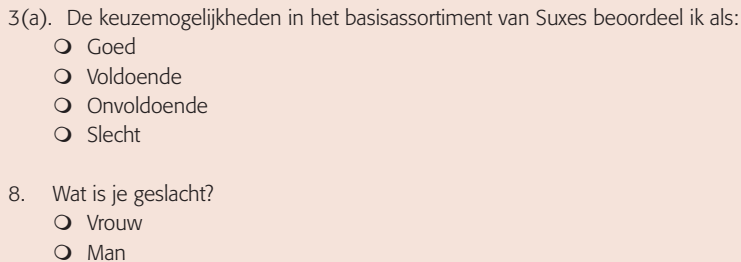
- gesloten vragen;
- open vragen met numerieke antwoorden;
- open vragen met tekstantwoorden;
- multiple response vragen.

In de volgende paragrafen bespreken we deze verschillende soorten enquêtevragen.

### 1.2.1 Gesloten vragen

**Gesloten vragen** zijn enquêtevragen waarop de respondent slechts één keuze mag maken uit een beperkt aantal antwoordmogelijkheden. De vertaling van dit soort vragen naar een SPSS-databestand is eenvoudig. Voor elke vraag van dit type wordt in SPSS één variabele gemaakt en de mogelijke antwoorden worden met *numerieke* codes gedefinieerd.

Voorbeelden uit de Pandion-enquête:

- 
- 3(a). De keuzemogelijkheden in het basisassortiment van Suxes beoordeel ik als:
- Goed
  - Voldoende
  - Onvoldoende
  - Slecht
8. Wat is je geslacht?
- Vrouw
  - Man

Bij enquêtevraag 3 worden de antwoordmogelijkheden als volgt gecodeerd:

- 1 = Goed
- 2 = Voldoende
- 3 = Onvoldoende
- 4 = Slecht

De data-invoer kunnen we zo beperken tot het invoeren van codes (1, 2, 3, 4) in plaats van het invoeren van de gegeven antwoorden (goed, voldoende, onvoldoende, slecht).

Voor enquêtevraag 8 kunnen we de volgende codes kiezen:

- 1 = Vrouw
- 2 = Man

### 1.2.2 Open vragen met numerieke antwoorden

Bij dit type enquêtevragen zijn er geen antwoordmogelijkheden aangegeven op het enquêteformulier. De respondent (of enquêteur) noteert een getal als antwoord.

Ook hier kan de enquêtevraag door middel van een variabele in SPSS weergegeven worden. Codering van de antwoorden is in dit geval natuurlijk niet nodig, want het ingevoerde antwoord (=getal) is de score en niet een code met een bepaalde betekenis zoals in paragraaf 1.2.1.

Voorbeelden uit de enquête:

2. Welke bedrag besteed je gemiddeld per week in het restaurant?  
..... euro per week
7. Hoeveel jaar ben je als student ingeschreven bij Pandion?  
..... jaar

### 1.2.3 Open vragen met tekstantwoorden

Ook bij dit type enquêtevraag staan geen vooraf aangegeven antwoordmogelijkheden op het enquêteformulier. Nu bestaat het antwoord van de respondent echter uit tekst. In de Pandion-enquête komt dit type vraag niet voor.

Een voorbeeld van een open vraag met tekstantwoord is:

Ik zou het assortiment graag uitgebreid zien met de volgende producten:  
.....

Het verwerken van de enquêtevraag en de antwoorden van de respondenten kan op twee manieren:

1. We willen het door de respondent gegeven antwoord letterlijk overnemen en als tekst invoeren. We definiëren hiervoor een tekstvariabele in SPSS. Er is een extra module bij SPSS verkrijgbaar, SPSS Text Analysis for Surveys, waarmee je tekstanalyses kunt maken, zoals het tellen van steekwoorden. Dit valt echter buiten het bestek van dit boek.
2. We kunnen ook een numerieke variabele in SPSS definiëren, waarbij we na het afnemen van de enquête de gegeven antwoorden coderen. Vooral als het aantal antwoorden dat respondenten (kunnen) geven beperkt is, is dit een aan te bevelen methode.

### 1.2.4 Multiple response vragen

Bij dit type enquêtevraag heeft de respondent de mogelijkheid om meer dan één antwoord te geven.

Voorbeeld uit de Pandion-enquête:

4. Welke producten koop je wel eens in het restaurant?  
(meerdere antwoorden mogelijk)
- Broodje ham/kaas
  - Een ander broodje
  - Zuivelproducten
  - Koffie of thee
  - Soep
  - Salades

In feite worden in bovenstaand voorbeeld zes vragen aan de respondent gesteld:

Koopt u wel eens een broodje ham/kaas?

Koopt u wel eens een ander broodje?

Enzovoort

Bij elk van de zes vragen moet de respondent een keus maken tussen de mogelijke antwoorden 'Ja' of 'Nee'. Er is natuurlijk een vrijheid bij het kiezen van codes, maar het verwerken in SPSS van multiple response vragen (zie hoofdstuk 5) is het eenvoudigst als de antwoorden gecodeerd worden met 1=Ja en 0=Nee.

Bovenstaande enquêtevraag moet dus in SPSS ingevoerd worden met behulp van zes variabelen met telkens een dichotome antwoordstructuur: 1=Ja, 0=Nee.

## 1.3 Meetniveaus

In de analysefase moeten met behulp van statistische analyses uiteindelijk antwoorden geformuleerd worden op de onderzoeksdeelvragen. Welke statistische methode bruikbaar is, hangt af van het soort enquêtevraag. Beter gezegd, het meetniveau van de betreffende variabele is bepalend voor de analysemethode die gebruikt kan worden.

Met een klein voorbeeld proberen we dit te verduidelijken. Eén van de elementaire analysemethoden is het uitrekenen van gemiddelden. Opvallend is dat bij sommige variabelen het gemiddelde niet en bij andere variabelen het gemiddelde wel uitgerekend kan worden. Praten over het gemiddelde *Geslacht* is onzinnig, terwijl bij bijvoorbeeld de variabele *Bedrag* wel een gemiddelde uitgerekend kan worden. Dit verschil heeft dus te maken met het niveau waarop variabelen gemeten worden. Deze zogenaamde meetniveaus van variabelen bepalen in hoge mate welke statistische bewerkingen mogelijk zijn. Er zijn voor ons vakgebied drie meetniveaus te onderscheiden.

### 1.3.1 Nominaal meetniveau

Een variabele heeft een **nominaal meetniveau** als bij de mogelijke scores slechts keuze is uit een aantal categorieën, waarbij tussen deze categorieën geen volgorde aanwezig is. In de bovenstaande enquête van Pandion heeft bijvoorbeeld de variabele *Studierichting* een nominaal meetniveau en de antwoordcategorieën zijn te beschouwen als namen.

Andere voorbeelden zijn:

- Woonsituatie (Thuiswonend, Uitwonend);
- Kopen van het product 'Broodje ham/kaas' (Ja, Nee).

Bij variabelen op nominaal niveau valt er weinig te rekenen. Een gemiddelde uitrekenen is niet mogelijk. Ook het bepalen van de mediaan kan niet. Turven (het tellen hoe vaak elke score voorkomt) is een statistische bewerking die wel mogelijk is. Merk op dat SPSS wel in staat is bij de variabele *Studierichting* een gemiddelde te berekenen. De antwoorden op de vraag naar de studierichting worden immers met *numerieke codes* ingevoerd. Het door SPSS gegenereerde gemiddelde voor de variabele *Studierichting* heeft echter geen betekenis! Interpretatie van door SPSS gemaakte berekeningen blijft mensenwerk.

### 1.3.2 Ordinaal meetniveau

Bij een **ordinaal meetniveau** gaat het ook om een variabele die is ingedeeld in categorieën. Nu kunnen alle categorieën echter in een logische volgorde gezet worden. Er is dus tussen de categorieën een *ordering* aanwezig. Voorbeelden in de Pandion-enquête zijn de zes variabelen van enquêtevraag 3, waarin vier categorieën met een logische ordening (goed, voldoende, onvoldoende, slecht) als antwoordmogelijkheden zijn vermeld.

Ook een variabele met een klassenindeling rekenen we tot het ordinale meetniveau. Bijvoorbeeld *Leeftijd* (met vier categorieën met een logische ordening: 15-<25, 25-<40, 40-<60, 60-<90).

Bij variabelen op ordinaal niveau zijn meer berekeningen mogelijk dan bij variabelen op nominaal niveau. Bij ordinale variabelen kan bijvoorbeeld de middelste score (mediaan) berekend worden. In het geval van een klassenindeling kan soms zelfs een gemiddelde berekend (benaderd) worden.

### 1.3.3 Interval- en ratiomeetniveau

Het hoogste meetniveau is het **ratiomeetniveau**. Het betreft hier variabelen die niet in categorieën gemeten worden en waarbij met de mogelijke scores berekeningen uitgevoerd kunnen worden (zoals het bepalen van een gemiddelde). Een voorbeeld in de Pandion-enquête is de variabele *Bedrag*: deze wordt gemeten met losse scores (dus niet in categorieën!) en met de scores kan gerekend worden. Ook de variabele *Aantal inschrijffaren* is op ratiomeetniveau gemeten.



Het **intervalmeetniveau** komen we tegen bij variabelen als *Temperatuur* (gemeten in graden Celsius) of *Datum* (zoals we die kennen van de kalender). Het verschil met het **ratiomeetniveau** is dat deze variabelen geen absoluut nulpunt hebben. We gaan hier niet nader op in, want SPSS maakt geen onderscheid en geeft deze beide meetniveaus aan met *Scale*.

---

**Opmerking** Een variabele kan, afhankelijk van de wijze waarop hij gemeten wordt, verschillende meetniveaus hebben. De variabele leeftijd uit de vorige subparagraaf wordt vanwege de klassenindeling gemeten op ordinaal niveau. Stel dat de vraagstelling in de enquête als volgt was geweest:

Hoe oud bent u? ... jaar.

Op deze wijze krijgt de variabele *Leeftijd* een **ratio**meetniveau!

Als men gevraagd zou hebben:

Tot welke categorie rekent u zichzelf? Jongeren/Ouderen

dan zou men slechts op **nominaal** niveau meten.

---

### **Samenvattend**

Het meetniveau van een variabele wordt bepaald door te kijken hoe er wordt gemeten, dus door te kijken naar de antwoorden die er op een bepaalde vraag in een enquête mogelijk zijn. Wij onderscheiden drie meetniveaus:

- Nominaal
- Ordinaal
- Ratio/interval (in SPSS: scale)

## **1.4 Het codeboek**

Wanneer de verzamelde gegevens met SPSS worden verwerkt, dan is het noodzakelijk dat iedere vraag wordt omgezet in een variabele. Bij de vertaling van de enquêtevragen naar variabelen maak je een **codeboek**. Het codeboek is dus een lijst met afspraken waarin wordt vastgelegd hoe de enquêtevragen worden vertaald in variabelen en hoe de antwoorden worden gecodeerd. Ook wordt aangegeven welk meetniveau de variabele heeft.

Zoals in paragraaf 1.2 is toegelicht, kan voor één enquêtevraag in SPSS één variabele worden gebruikt, behalve als het een **multiple response** vraag betreft. In dat geval definiëren we net zo veel variabelen als er antwoordmogelijkheden zijn bij de betreffende **multiple response** vraag.

Enige aanwijzingen bij het kiezen van namen voor variabelen en coderingen:

- De variabelen krijgen een beknopte naam in SPSS. Het eerste teken van deze naam moet een letter zijn. De rest van de naam mag cijfers en letters bevatten. Ook het gebruik van de symbolen @, #, \$ en \_ is toegestaan. Spaties en punten zijn echter verboden.
- De naam van de variabele die in SPSS wordt ingevoerd is vaak een beknopte weergave van het gemeten kenmerk in de enquêtevraag. Aan deze variabelenaam wordt een **label** (=omschrijving van de variabelenaam) toegekend. In de tabellen en grafieken die in de analysefase worden gemaakt, wordt dit label opgenomen. Zorg dus voor duidelijke omschrijvingen.
- De (eventueel) te gebruiken codes (*values* in SPSS) bij het invoeren van de gegevens dienen eveneens van een label te worden voorzien. Deze **Value Labels** worden eveneens in de tabellen en grafieken opgenomen. Hier moeten dus ook weer heldere en kernachtige omschrijvingen gebruikt worden.
- Het meetniveau van een variabele (*measure* in SPSS) is vooral van belang voor de analysefase. Een statistische analyse mag alleen uitgevoerd worden als de variabelen het vereiste meetniveau hebben. Bij sommige analyses laat SPSS alleen maar die variabelen zien waarmee de betreffende analyse kan worden gedaan.