

Hoofdstuk 1. Elementen voor de basis

Onderzoekseenheden zijn objecten over wie of wat het onderzoek een uitspraak doet, bijvoorbeeld tweedejaarsstudenten. Variabelen zijn de kenmerken van de onderzoekseenheden, bijvoorbeeld leeftijd, geslacht etc. Variabelen kunnen waarden hebben. Bij sommige kenmerken zijn de waarden of values al een getal zoals bij leeftijd, bij andere kenmerken zou je voor de waarden een getal kunnen verzinnen bijvoorbeeld man is 0 en vrouw is 1. In dat geval wordt 'man' bij de waarde 0 een valuelabel genoemd. In een frequentietabel kan op een overzichtelijke wijze een variabele met bijbehorende categorieën worden gepresenteerd. Absolute frequentie staat voor het aantal variabelen. Soms zijn deze lastig te begrijpen vandaar dat je ook vaak naar percentages kijkt. Stel dat in een frequentietabel staat dat 8 personen van in totaal 22 personen cola drinken, dan drinkt dus 36,4 procent cola. Hierbij is het getal 8 de absolute frequentie en het percentage 36,4 is de relatieve frequentie. Een datamatrix is een spreadsheet waarin per onderzoekseenheid alle kenmerken in afzonderlijke variabelen worden beschreven.

Percentages

Valid Percent is het percentage dat berekend is over de respondenten die hebben deelgenomen én hebben geantwoord. Het wordt gebruikt wanneer iemand niet gereageerd heeft op een vraag en er daardoor een antwoord ofwel waarde ontbreekt (**missing value**). Door te kijken naar Valid Percent kun je toch een goede uitslag krijgen. **Cumulative Percent** betekent dat de percentages van elke volgende waarde bij de eerste worden opgeteld, bijvoorbeeld 53,8 procent is 16 jaar of jonger.

Frequentietabel

Een frequentietabel geef je in SPSS weer via Analyse → Descriptive Statistics → Frequencies. Een grafiek geef je in SPSS weer via Analyse → Descriptive Statistics → Frequencies en dan Charts. Je kunt hierbij kiezen uit een taartdiagram, een staafdiagram en een histogram.

Kruistabel

Bij een kruistabel wordt gebruik gemaakt van twee variabelen in één analyse, bijvoorbeeld zowel sekse als leeftijd. Om overzicht te krijgen, is het goed om de percentages erbij te geven. Dit kan aan de hand van:

- totaalpercentages (totaal aantal onderzoekseenheden op 100% stellen);
- kolompercentages (onderzoekseenheden in de kolommen op 100% stellen);
- rijpercentages (onderzoekseenheden in rijen op 100% stellen).

Het is afhankelijk van welke uitspraken je wilt doen op basis van je onderzoek, welke van de drie je kiest.

Een kruistabel maak je in SPSS via Analyse → Descriptive Statistics → Crosstabs. Je kunt hierbij aangeven welke variabelen in de rijen en welke in de kolommen moeten komen. Om aan te geven dat de kolommen gepercenteerd moeten worden, klik je op Cells. In SPSS wordt vervolgens achter Count de absolute frequenties weergegeven. Onder Count staat % within variabele.

JoHo Samenvatting – Beschrijvende Statistiek

Onafhankelijke en afhankelijke variabelen

Een **onafhankelijke variabele** is de variabele die invloed uitoefent en meestal wordt aangegeven met een x . Bijvoorbeeld geslacht is onafhankelijk, geen enkele variabele kan hierop invloed uitoefenen. De **afhankelijke variabele** is de variabele die beïnvloed wordt en meestal wordt aangegeven met een y . Bijvoorbeeld: de onafhankelijke variabele geslacht heeft invloed op het televisieprogramma dat iemand graag kijkt (afhankelijk).

Meetniveaus

Er worden verschillende meetniveaus gehanteerd binnen de statistiek. Ten eerste is er het **nominaal** meetniveau. De numerieke waarde heeft hierbij als getal geen betekenis. Er zou net zo goed voor andere getallen of symbolen kunnen worden gekozen. Voorbeelden van nominale meetniveaus zijn: geslacht, geloof en voorkeur smaak ijs. Ten tweede is er het meetniveau **ordinaal**. Hierbij is sprake van een rangordening. Een voorbeeld hiervan is boodschappbestedingen per week (minder dan 50, 50 tot 100, 100 tot 200, meer dan 200 euro per maand). Als derde meetniveau hebben we **interval**. Hierbij is sprake van een vaste meeteenheid waarbij de waarden voor de stappen betekenis toekennen aan de afstanden tussen de stappen. Een voorbeeld hiervan is temperatuur of geboortejaar. Er is geen vast nulpunt (arbitrair). Het laatste meetniveau is **ratio**. Bij dit niveau is sprake van een rangordening, waarbij de intervallen betekenis hebben en is er sprake van een natuurlijk nulpunt. Bij dit meetniveau hebben niet alleen de verschillen tussen de afzonderlijke waarden betekenis, maar ook het quotiënt, het resultaat van een deling. Voorbeelden van dit meetniveau zijn lengte, leeftijd en afstand in meters. Er zijn dus vier criteria waar je rekening mee moet houden om het meetniveau te bepalen: de classificatie, de rangordening, de intervalbetekenis en het natuurlijke nulpunt.

Waarden van variabelen

Variabelen kunnen verschillende waarden aannemen. Een **discrete meetschaal** houdt in dat alleen gebruik wordt gemaakt van gehele getallen, tussenliggende waarden hebben geen betekenis. Je kan bijvoorbeeld niet spreken van 20,6 proefpersonen, wel van 20 of van 21 proefpersonen. Bij een **continue meetschaal** kunnen alle mogelijke waarden de uitkomst zijn van de meetprocedure, bijvoorbeeld niet alleen 3 of 4 maar ook 3,3 en 3,825. De laatste mogelijk waarde die een variabele zich kan aannemen, is de **dichotome variabele**. Deze meetschaal kan slechts twee mogelijke waarden aannemen, bijvoorbeeld ja (0) of nee (1).

Analyseniveaus

Er kunnen in een analyse meerdere variabelen worden gebruikt. Dit is mede afhankelijk van wat je te weten wilt komen. De **univariate analyse** is een frequentietabel waarbij maar één variabele wordt gebruikt. Bij de **bivariate analyse** worden twee variabelen met elkaar vergeleken. Dit is het geval bij kruistabellen. Bij de **multivariabele analyse** worden meer dan twee variabelen gebruikt.