

DATAPROCESSING



Na de uitgave van de ABF kernpublicatie over het WoonOnderzoek Nederland 2009 (WoON 2009) wordt het onderzoeksbestand vrijgegeven. Het WoON-bestand bevat van ruim 70.000 respondenten een verscheidenheid aan informatie op het terrein van demografie, de woningvoorraad, woontevredenheid en woonwensen. Het spreekt voor zich dat een onderzoeksbestand van deze omvang, dat beleidsbepalend is op het gebied van wonen, van hoge kwaliteit moet zijn. ABF Research is een van de betrokken partijen geweest om dat voor elkaar te krijgen.

Dataprocessing: expertise ABF

In de loop van de jaren heeft ABF Research grote expertise opgebouwd op het gebied van dataprocessing. Sinds begin jaren negentig van de vorige eeuw is ABF Research altijd betrokken geweest bij de dataprocessing van het WoON (voorheen WBO). Andere aansprekende onderzoeksbestanden waarvoor ABF Research de dataprocessing uitvoert, zijn de Veiligheidsmonitor en Bewoners Nieuwe Woningen.

Dataprocessing: van veldwerk naar onderzoek

De kwaliteit van een onderzoeksbestand is in eerste instantie afhankelijk van de kwaliteit van het verrichte veldwerk. Maar de wijze waarop de onderzoeksgegevens verwerkt worden (dataprocessing), is ook van groot belang. In deze fase wordt vanuit het ruwe veldwerkbestand – met daarin de primaire veldwerkresultaten – toegewerkt naar een onderzoeksbestand. Het onderzoeksbestand wordt ontdaan van bestandsvervuiling en allerlei andere oneffenheden, zodat de gebruikers van het bestand hier in een later stadium geen hinder meer van ondervinden.

Dataprocesing in vogelvlucht

Dataprocesing omvat een aantal stappen. De belangrijkste zijn: routingcontrole, consistentie, bijschatting, weging en plausibiliteit.

- **Routingcontrole:** Routing is van belang voor het juist doorlopen van een vragenlijst, omdat niet altijd alle vragen hoeven te worden beantwoord. Als een respondent bijvoorbeeld aangeeft een huurder te zijn, is het vragenblok over de koopwoning niet meer van toepassing en moet juist het vragenblok over de huurwoning ingevuld worden. Wordt de routing van de vragenlijst niet goed doorlopen, dan levert de enquête weinig bruikbare uitkomsten op. Daarom zijn tussentijdse routingcontroles van belang om het veldwerk tijdig bij te sturen.
- **Consistentie:** Er moet een 'logisch' verband zijn tussen de antwoorden. Zo kan een respondent die in 1967 geboren is vrijwel zeker geen ouder zijn van een kind dat in 1978 geboren is. Het is in dat geval waarschijnlijk dat er sprake is van een tikfout en dat het geboortjaar 1987 bedoeld wordt. Omwille van de consistentie wordt daarom het geboortjaar van het kind aangepast.
- **Bijschatting:** Soms geeft een respondent geen antwoord op een vraag, terwijl hij dat wel had moeten doen. Dit wordt ook wel (item) non-respons genoemd. Wanneer hier niets mee gedaan wordt, wordt impliciet verondersteld dat non-respondenten zich hetzelfde gedragen als respondenten. Menig onderzoek heeft in het verleden al uitgewezen dat deze veronderstelling meestal onjuist is, omdat non-respons vaak samenhangt met bepaalde kenmerken van de respondent. ABF Research maakt voor problemen rond non-respons al jaren gebruik van de zogenaamde imputatietechniek, waarmee ontbrekende waarden worden bijgeschat. Bijvoorbeeld voor de bijschatting van de opleiding van een respondent worden onder andere de leeftijd, de opleiding van een eventuele partner en het inkomen gebruikt.
- **Weging/ophoging:** Het komt geregeld voor dat steekproefbestanden een selectieve respons kennen. Dit houdt in dat de steekproef niet representatief is voor de populatie, omdat groepen van respondenten over- dan wel ondervertegenwoordigd zijn. Met behulp van weging kan voor deze selectiviteit een correctie plaatsvinden, zodat de gewogen steekproef wel representatief is. Na de weging wordt ook een zodanige ophoging toegepast dat de gewogen aantallen zo veel mogelijk overeenkomen met andere bronnen, waaronder die van het CBS. Zo moet de gewogen som van alle respondenten in de steekproef overeenkomen met het feitelijke populatieaantal.
- **Plausibiliteit:** Tot slot wordt de plausibiliteit van de gewogen uitkomsten tegen het licht gehouden. De uitkomsten worden vergeleken met informatie die afkomstig is uit andere bronnen (registratiebestanden) of uit het verleden (tijdreeksen).

Meer informatie

Voor meer informatie kunt u contact opnemen met ABF Research:

Kenneth Gopal	tel. 015 - 27 99 332	kenneth@abf.nl
Eric Verhoog	tel. 015 - 27 99 322	eric@abf.nl

ABF RESEARCH
VERWERSDIJK 8
2611 NH DELFT
T [015] 27 99 300
F [015] 27 99 333
E INFO@ABFRESEARCH.NL
S WWW.ABFRESEARCH.NL
ING 1769518
ABN-AMRO 44.21.66.338
ABF RESEARCH BV
K.V.K. TE DELFT 27.31.05.58
BTW NR. NL 8188.62.312.B01