



Bestandsacceptatie WoON2021

Saskia Janssen-Jansen

Versie: 0.1

CBS Heerlen
CBS-weg 11
6412 EX Heerlen
Postbus 4481
6401 CZ Heerlen
+31 45 570 60 00
www.cbs.nl

projectnummer

305707
SAL
01 juni 2022

Versiebeheer en distributie

Versie	Datum	Beschrijving van de wijziging	Auteur
0.1	01-06-22	Initiële versie	SJNS
	dd-mm-jj	Geplande update.	

Versie	Distributie
0.1	BZK

Inhoudsopgave

Versiebeheer en distributie	3
Inhoudsopgave	4
1. Inleiding	5
1.1 Doel	5
1.2 Leeswijzer	5
1.3 Literatuur	5
2. Samenvatting: Bestandsacceptatie	6
2.1 Kwaliteitscontrole	6
2.2 Responsbestand WoON 2021	6
3. Controle en Consistentie checks	7
3.1 Controle op volledigheid respons	7
3.2 Controle op interne consistentie respons	8
3.3 Controle met behulp van registerinformatie	10
3.4 Conclusie	11
4. Respons WoON 2021	12
4.1 Respons in WoON 2021 bestand	12
4.2 Respons reguliere Woningmarktmodule	13
4.3 Respons naar oversamplingdeelnemers	13
4.4 Conclusie respons	14

1. Inleiding

1.1 Doel

Dit document beschrijft de controle- en kwaliteitsprocessen die starten bij het binnenhalen van de veldwerkgegevens van de module Woningmarkt en eindigen met de samenstelling van een responsbestand. De informatie heeft zowel betrekking op het reguliere deel van de Woningmarktmodule als op de oversampling.

1.2 Leeswijzer

In hoofdstuk 2 worden de belangrijkste conclusies gepresenteerd. De veldwerkgegevens afkomstig uit twee verschillende bronnen (CBS: CAPI/CATI en CAWI) worden samengevoegd tot één uniform responsbestand. Het complete uniforme responsbestand wordt vervolgens beoordeeld op kwaliteit. Deze kwaliteitscontrole staat beschreven in hoofdstuk 3.

In het laatste hoofdstuk (4) wordt het uiteindelijke responsbestand beschreven aan de hand van een aantal kenmerken.

1.3 Literatuur

- Janssen, B (2022). WoON 2021 – Woningmarkt module. Onderzoeksverantwoording. CBS, juni 2022.
- Projectovereenkomst WoON 2021, Module Woningmarkt, Maart 2021.
- W. Gielen, 'Dataprocesing WoON2021-controle en consistentie (C&C)', december 2021.

2. Samenvatting: Bestandsacceptatie

2.1 Kwaliteitscontrole

Na samenvoegen van de gegevensbestanden tot één uniform bestand bestaat het bestand uit 48.798 responsen. Door de controles op volledigheid zijn er 2.140 responsen uit het bestand verwijderd. Het uniforme responsbestand met volledige vragenlijsten bestaat dan nog uit 46.658 records. In het gecorrigeerde bestand zijn vervolgens nog een aantal inconsistenties geconstateerd. Deze inconsistenties zijn gecorrigeerd, zoals beschreven is in de Controle en Consistenties van Willem Gielen (2021).

Het uniforme responsbestand met 46.658 cases wordt als kwalitatief goed beoordeeld en acceptabel bevonden voor verdere analyses.

2.2 Responsbestand WoON 2021

In het responsbestand met 46.658 cases zijn er 39.736 afkomstig van de reguliere Woningmarkt module en 6.922 zijn afkomstig van de oversampling. De responspercentages komen uit op 52% voor het totale bestand. Als gevolg van de coronapandemie zijn zowel binnen de reguliere Woningmarkt module als bij 7 deelgebieden van de oversampling de afgesproken responseisen niet behaald.

3. Controle en Consistentie checks

Het document 'Dataprocesing WoON2021-controle en consistentie (C&C)' van Willem Gielen (2021) beschrijft de controles die uit gevoerd zijn. In de bestandsacceptatie is een onderscheid gemaakt tussen controles op volledigheid (zie paragraaf 3.1), controles die checken op interne inconsistenties binnen de vragenlijst (3.2) en controles die checken met behulp van registerinformatie (3.3).

3.1 Controle op volledigheid respons

Indien een respondent scoort op een van de controles die checken of een vragenlijst volledig is afgenomen, wordt deze respons afgekeurd en alsnog uit het uniforme responsbestand verwijderd. In totaal zijn er 3 controles op volledigheid. Het gaat om de volgende Controles:

- SF01b: > 15% 'weet niet' en 'weigert';
- SF03_1: Steekproefpersoon is niet gelijk aan de respondent;
- SF04b: Geboortedatum respondent;

In totaal hebben 2.140 respondenten één of meer keer gescoord op de controles op volledigheid (zie tabel 1). Veruit de meest responsen (99%) vallen af doordat meer dan 15% van de vragen is beantwoord met 'weet niet' en of 'weigert'. Dit gaat vooral ten koste van de CAWI responsen. In WoON2021 zijn er 1.000 responsen minder afgevallen door dan in WoON2018, maar dit komt vooral doordat het bestand minder responsen bevat ten behoeve van de oversampling. Na de verwijdering van de 2.140 onvolledige vragenlijsten uit het uniforme responsbestand bevat het responsbestand nog 46.658 responsen.

Tabel 1. Aantal responsen dat niet volledig is na de controle.

	Regulier	Oversampling	Totaal
Totaal volledige veldwerk responsen	41.558	7.240	48.798
Totaal onvolledige veldwerkresponsen door C&C	1.822	318	2.140
Totaal volledige responsen na C&C	39.736	6.922	46.658
Totaal onvolledige respons	1.822	318	2.140
• CAPI	7	7	14
• CATI	11	4	15
• CAWI	1.894	1.091	2.985
SF01b: > 15% 'weet niet' en 'weigert'			
• CAPI	2	0	2
• CATI	11	1	12
• CAWI	1.799	315	2.114
SF03_1: Steekproefpersoon is niet gelijk aan de respondent			
• CAPI	5	0	5
• CATI	3	1	4
• CAWI	3	0	3
SF04a Geboortedatum respondent			
• CAPI	2	1	3
• CATI	1	0	1
• CAWI	0	0	0

3.2 Controle op interne consistentie respons

In totaal zijn er 15 controles uitgevoerd om inconsistenties van gegeven antwoorden op te sporen. Als antwoorden niet consistent zijn wordt een consistente waarde geïmputeerd. De controles op inconsistentie zijn uitgevoerd op het uniforme responsbestand, met daarin alleen de volledige responsen (n=46.658). De volgende controles zijn uitgevoerd:

- SF06 : Leeftijdsverschil respondent en partner
- SF07 : Leeftijdsverschil oudste ouder en oudste kind
- SF11b : Jaar op huidig adres komen wonen (2)
- SF13a : Jaar koopwoning gekocht (1)
- SF13b : Jaar koopwoning gekocht (2)
- SF18 : Jaar op vorig adres komen wonen
- SF19 : Bouwjaar vorige woning
- SF24: Verdiepingen woning/rookmelders
- SF25: Bouwlagen woongebouw
- SF32: Benedenburen zijn aanwezig in een eengezinswoning
- SF35: Belangrijkste reden opgevuld
- SF36 : Verhuizing naar het buitenland
- SF37 : Controle op antwoordpatronen
- SF39: Leeftijd respondent en partner is plausibel
- SF40: Controle op kosten stads-/blokverwarming

In tabel 2 is weergegeven hoe vaak een inconsistentie is voorgekomen en indien mogelijk is aangepast.

Tabel 2. Aantal geconstateerde inconsistenties.

	CAPI	CATI	CAWI	Totaal
SF06 Leeftijdsverschil respondent en partner				
Fout geen actie	0	0	0	0
Fout + actie	3	2	12	17
SF07 Leeftijdsverschil kind ouder				
Fout geen actie	7	4	0	11
Fout + actie	7	16	7	30
SF11b Geboortjaar ligt na het jaar dan men op het huidig adres is komen wonen				
Fout + actie	0	0	0	0
SF13a Jaar koopwoning meer dan 3 jaar voor jaar komen wonen gekocht				
Fout geen actie	49	19	213	281
Fout + actie	0	0	0	0
SF13b Koopwoning gekocht voor de geboortedatum				
Fout + actie	0	0	0	0
SF18 Jaar op vorig adres komen wonen				
Fout + actie	0	0	0	0
SF19 Bouwjaar vorige woning				
Fout + actie	0	0	0	0
SF24 Verdiepingen vs verdiepingen met rookmelder				
Fout + actie	74	197	592	863
SF25 Verdiepingen woning vs gehele woongebouw				
Fout + actie	53	41	767	861
SF32 Benedenburen zijn aanwezig terwijl het om een eengezinswoning gaat				
Fout + actie	89	48	213	350
SF35a Belangrd aangepast				
Aangepast	793	209	2.454	3.456
SF35b Voorredn aangepast				
Aangepast	1.257	876	7.778	9.911
SF35c Belnksgv aangepast				
Aangepast	1.691	1.264	9.476	12.431
SF36 Buitenland verhuiscwensen consistent gemaakt				
Fout + actie	23	15	108	146
SF37 Controle op antwoordpatronen				
Antwoordpatroon op vraag 12.12 t/m 12.17	1	0	46	47
Antwoordpatroon op vraag 13.2 t/m 13.8	0	0	0	0
Antwoordpatroon op vraag 13.9 t/m 13.16	0	0	0	0
SF 40 Opgegeven servicekosten of VVE Bijdrage en verwarmingskosten kloppen niet				
Stads of blokverwarming is gelijk aan VVE bijdrage	1	1	9	11
Stads of blokverwarming is gelijk aan service kosten	35	20	131	186

3.3 Controle met behulp van registerinformatie

In totaal zijn er 14 controles beschreven om inconsistenties van gegeven antwoorden op te sporen met behulp van registerinformatie. De controles van registerbestand zijn uitgevoerd op een registratiegegevens die doorlopen tot 1 oktober 2021 (=meest recent). De volgende controles zijn uitgevoerd:

- SF08 : controle huishouden
- SF10a : Jaar in NL komen wonen
- SF11a : Jaar op huidig adres komen wonen (1)
- SF13c : Jaar koopwoning gekocht (3)
- SF13d : Jaar koopwoning gekocht (4)
- SF14 : Lift bij hooggelegen woonkamer
- SF16a : Oppervlakte woonkamer en woonoppervlak (1)
- SF16c : Oppervlakte woonkamer en woonoppervlak (3)
- SF17a : Oppervlakte woonkamer en woonoppervlak (1)
- SF17c : Oppervlakte woonkamer en woonoppervlak (3)
- SF26a : Bouwjaar gekochte, niet-bestaande woonruimte (1)
- SF29 : Woontype huidige woning met BAG
- SF30 : Jaar komen wonen in GBA
- SF33: Corporatie/particuliere huurwoning

De controles met behulp van registerinformatie zijn uitgevoerd op het uniforme responsbestand, met daarin alleen de volledige responsen (n=46.658). De resultaten staan in tabel 3 gepresenteerd. Opvallend is nog steeds dat 14% van de respondenten aangeeft dat er minder kinderen thuis wonen dan in BRP is geregistreerd. Dit percentage komt overeen met dan in WoON2018. Verschillen zijn net als bij de interne consistentiecontroles het grootst bij de CAWI waarneming.

Tabel 3. Aantal geconstateerde inconsistenties vergeleken met registratie.

	CAPI	CATI	CAWI	Totaal
SF08 Controle huishouden				
Minder personen in vragenlijst opgegeven	269	312	969	1.550
Meer personen opgegeven in vragenlijst	313	213	502	1.028
Minder kinderen in vragenlijst opgegeven	1.046	1.457	4.160	6.663
Meer kinderen opgegeven in vragenlijst	160	131	262	553
Plaats in huishouden wijkt af	377	339	978	1.694
Type huishouden wijkt af	389	323	1.046	1.758
SF10a Jaar in NL komen wonen				
Fout geen actie	0	0	0	0
SF11a Jaar op huidige adres komen wonen				
Fout geen actie	50	14	88	152
Fout + actie	100	76	230	406
SF13c Jaar koopwoning gekocht (3)				
Fout geen actie	29	10	116	155
Fout + actie	18	37	136	191
SF13d Jaar koopwoning gekocht (4)				
Fout geen actie	-	-	7	7
Fout + actie	3	40	105	148
SF14 Lift bij hoge gebouwen				
Fout + actie	2	-	8	10
SF16a Oppervlakte woonkamer (1)				
Fout + actie	25	8	91	124
SF16c Oppervlakte woonkamer (3)				
Fout + actie	19	5	63	87
SF17a Oppervlakte woonkamer (5)				
Fout + actie	46	79	1.358	1.483
SF17c Oppervlakte woonkamer (6)				
Fout + actie	16	8	108	132
SF26a Bouwjaar nieuwbouw woning				
Fout geen actie	-	3	18	21
Fout + actie	6	34	79	119
SF29 Woningtype versus BAG				
Fout geen actie	255	177	770	1.202
SF30 Jaarkomen wonen in huidige woning				
Fout geen actie	343	723	2.258	3.324
SF33 Corporatie/particuliere huurwoning				
Verhuurder omgezet naar corporatie	87	31	119	237
Verhuurder omgezet naar overig verhuurder	285	211	947	1.443

3.4 Conclusie

Na samenvoegen van de gegevensbestanden tot één uniform bestand bestaat het bestand uit 48.798 responsen. Door de controles op volledigheid zijn er 2.140 responsen uit het bestand verwijderd. Het uniforme responsbestand met volledige vragenlijsten bestaat dan nog uit 46.658 records. In het gecorrigeerde bestand zijn vervolgens nog een aantal inconsistenties geconstateerd. Deze inconsistenties zijn gecorrigeerd, zoals beschreven is in de Controle en Consistenties van Willem Gielen (2021).

Het uniforme responsbestand met 46.658 cases wordt als kwalitatief goed beoordeeld en acceptabel bevonden voor verdere analyses.

4. Respons WoON 2021

4.1 Respons in WoON 2021 bestand

In de projectovereenkomst met BZK staat dat de reguliere Woningmarktmodule minimaal 40.000 volledige responsen moeten zijn opgenomen. Met een aantal van 39.736 gerealiseerde volledige responsen is hieraan niet voldaan. In het oversampling deel zijn eveneens de afgesproken responsaantallen niet overal gerealiseerd. In totaal zijn er 6.922 cases opgenomen die voor de oversampling zijn waargenomen (zie tabel 5). Door de corona pandemie is het CAPI veldwerk ruim 4 maanden geheel uitgevallen en in een aantal maanden beperkt mogelijk geweest. Om dit verlies te compenseren is besloten om de veldwerkperiode te verlengen tot eind september 2021. In overleg met BZK is niet besloten om dit verlies aan CAPI respondenten te compenseren met een nieuwe steekproef en waarneming via een andere mode. Dit omdat het niet ten goede zou komen van de kwaliteit van het responsbestand. Helaas heeft het verlengen van de veldwerkperiode er niet voor kunnen zorgen dat alle steekproefpersonen aan huis benaderd konden worden. Hierdoor zijn de afgesproken aantallen bij WoON Regulier en bij Oversampling niet gerealiseerd.

Binnen de reguliere Woningmarktmodule is 12 procent waargenomen via CAPI, 17 procent via CATI en 71 procent via CAWI. Voor de oversampling is het aandeel van de waarneemmethode CAPI, CATI en CAWI respectievelijk 14, 14 en 72 procent. In de projectovereenkomst is de mode verdeling geschat op 68% CAWI, 18% CATI en 14% CAPI. Zoals in tabel 5 is aangegeven is het aandeel respondenten dat via CAWI deel heeft genomen hoger dan verwacht.

Tabel 5. Verdeling van de respons naar waarneemmethode

Responsbestand	Regulier		Oversampling		Totaal	
	N	%	N	%	N	%
Totaal	39.736		6.922		46.658	
Methode						
• Capi	4.953	12	964	14	5.917	13
• Cati	6.713	17	972	14	7.685	16
• Cawi	28.070	71	4.986	72	33.056	71

In het steekproefontwerp voor WoON-regulier is Nederland verdeeld in 19 deelgebieden. Per gebied is het uitgangspunt om 1.200 responsen te realiseren. Voor de oversamplingdeelnemers is voor elke contractpartner een apart steekproefontwerp vastgesteld, met aparte responsafspraken. Gedurende de dataverzameling is gecontroleerd of de respons overeenkwam met de verwachting. Tijdens het veldwerk zijn nog een paar extra steekproeven getrokken om tegenvallende CAWI responsen te compenseren.

Tabel 6 Ongewogen responscijfers WoON 2018

	Regulier	Oversampling	Totaal
	%	%	%
Uitgezette steekproef	100	100	100
Kaderfouten	8	8	8
Uitgezette steekproef min de kaderfouten	100	100	100
Non respons	44	48	45
Respons	56	52	55

De totale respons van WoON-regulier bedraagt 51,2% op de uitgezette steekproef. Dit responspercentage is een stuk lager als in WoON2018 en ook ruim onder het afgesproken responspercentage van 56%. In de onderzoeksverantwoording (Janssen, B, 2021) zijn de responscijfers in detail beschreven.

4.2 Respons reguliere Woningmarktmodule

In de projectovereenkomst van WoON2021 zijn responseisen naar 19 regionale gebieden vastgesteld. In tabellen 7 staan de resultaten voor de reguliere woningmarkt module gepresenteerd. De target per gebied is 1.200 responsen en dit is in alle gebieden gerealiseerd.

Tabel 7 Aantal gerealiseerde responsen voor Woon-regulier naar 19 gebied.

	Target	Responsen
Corop Agglomeratie Den Haag	1.200	1.855
Corop Arnhem/Nijmegen	1.200	1.826
Corop Groot - Amsterdam en Corop Zaanstreek	1.200	3.077
Corop Groot - Rijnmond	1.200	2.970
Corop Twente	1.200	1.414
Corop Zuidoost Noord-Brabant	1.200	1.847
Gemeente Utrecht en omgeving	1.200	1.642
Provincie Drenthe	1.200	1.344
Provincie Flevoland	1.200	1.202
Provincie Friesland	1.200	1.485
Provincie Gelderland excl. Corop Arnhem/Nijmegen	1.200	2.974
Provincie Groningen	1.200	1.352
Provincie Limburg	1.200	2.701
Provincie Noord - Holland excl. Corop Groot - Amsterdam en Corop Zaanstreek	1.200	2.879
Provincie Noord-Brabant excl. Corop Zuidoost Noord-Brabant	1.200	4.090
Provincie Overijssel excl. Corop Twente	1.200	1.220
Provincie Utrecht excl. Gemeente Utrecht en omgeving	1.200	1.359
Provincie Zeeland	1.200	1.348
Provincie Zuid-Holland excl. Corop Groot-Rijnmond en Agglomeratie Den Haag	1.200	3.151

In het veldwerverslag wordt verder ingegaan op de behaalde responsen naar gebied en achtergrondkenmerken (Janssen, B, 2021).

4.3 Respons naar oversamplingdeelnemers

In totaal hebben 11 oversamplingdeelnemers meegedaan. Met deze deelnemers zijn afspraken gemaakt over de te behalen respons. In tabel 8 staan de gerealiseerde responsen naar oversamplingdeelnemers. Voor zeven deelgebieden is de target niet gehaald: Amersfoort, Den Haag Particuliere huur en corporatie huur, Utrecht wijk 06 en 3 regio's van Gelderland (Achterhoek, FoodValley en Arnhem).

Tabel 8. Gerealiseerde responsen naar oversamplinggemeenten.

	Target	Regulier	Oversampling	Totaal
1,00 Gemeente Enschede	600	365	255	620
2,00 Gemeente Den Haag Koop	600	592	150	742
3,00 Gemeente Den Haag Corporatie huur	600	277	320	597
4,00 Gemeente Den Haag Particuliere huur	600	200	390	590
5,00 Gemeente Breda	600	430	218	648
6,00 Gemeente Maastricht	600	272	391	663
7,00 Gemeente Amersfoort	600	356	235	591
8,00 Gemeente Altena	600	126	484	610
9,00 Gemeente Zwolle Midden	200	60	151	211
10,00 Gemeente Zwolle Oost	250	53	210	263
11,00 Gemeente Zwolle Noord	250	59	197	256
12,00 Gemeente Zwolle West	250	66	232	298
13,00 Gemeente Zwolle Zuid	250	75	223	298
14,00 Gemeente Utrecht Wijk 01 West	250	64	189	253
15,00 Gemeente Utrecht Wijk 02 Noordwest	250	91	173	264
16,00 Gemeente Utrecht Wijk 03 Overvecht	250	52	205	257
17,00 Gemeente Utrecht Wijk 04 Noordoost	250	96	187	283
18,00 Gemeente Utrecht Wijk 05 Oost	250	86	178	264
19,00 Gemeente Utrecht Wijk 06 Binnenstad	250	48	195	243
20,00 Gemeente Utrecht Wijk 07 Zuid	250	73	204	277
21,00 Gemeente Utrecht Wijk 08 Zuidwest	250	88	171	259
22,00 Gemeente Utrecht Wijk 09 Leidsche Rijn	250	69	194	263
23,00 Gemeente Utrecht Wijk 10 Vleuten-De Meern	250	86	191	277
27,00 Gemeente Oldenzaal	600	70	556	626
28,00 Gemeente Kampen	600	122	481	603
29,00 Provincie Gelderland Regio Achterhoek	600	596	-	596
30,00 Provincie Gelderland Regio Noord-Veluwe	600	391	191	582
31,00 Provincie Gelderland Rivierenland	600	512	91	603
32,00 Provincie Gelderland Regio Food Valley	600	779	-	779
33,00 Provincie Gelderland Regio Stedendriehoek	600	1.012	-	1.012
34,00 Provincie Gelderland Regio Arnhem eo	600	763	-	763
35,00 Provincie Gelderland Regio Nijmegen eo	600	834	-	834
36,00 Provincie Gelderland Regio Liemers	600	333	260	593

4.4 Conclusie respons

In het responsbestand met 46.658 cases zijn er 39.736 afkomstig van de reguliere Woningmarkt module en 6.922 zijn afkomstig van de oversampling. De responspercentages komen uit op 52% voor het totale bestand. Als gevolg van de coronapandemie zijn zowel binnen de reguliere Woningmarkt module als bij 7 deelgebieden van de oversampling de afgesproken responseisen niet behaald.