



Richtlijnen bestanden

Richtlijnen voor het aanleveren van (gegevens)bestanden gebruikt bij WODC onderzoeken aan het WODC (Ministerie van Justitie en Veiligheid).

De organisatie die het onderzoek heeft uitgevoerd zal alle gegevensbestanden die specifiek voor het onderzoek zijn gebruikt, dan wel verkregen aan het WODC/SIBa leveren, ook de gebruikte coderingen (apart bestand, indien van toepassing). De betreffende organisatie zal, indien mogelijk, de bestanden in SPSS ("sav") formaat aanleveren. Indien er hiervoor een specifiek SPSS script is, dan ook dit script (goed gedocumenteerd) aanleveren. Daarnaast zijn op bestandsniveau en op het niveau variabele (veld of kolom) informatie (documentatie) gewenst, zoals onderstaand is weergegeven. Belangrijk is hierbij aan te geven of er sprake is van privacygevoelige elementen. Als er op de variabelen een weegfactor is toegepast worden de oorspronkelijke waarden gevraagd (dus zonder de betreffende weegfactor). Wel is dan de vermelding van het gebruik van een weegfactor en de waarde van die factor(en) vereist. De bestanden met de benodigde documentatie (een MS Excel werkblad in onderstaand formaat) dienen, voorzien van de onderzoeksnaam en het WODC-kenmerk (zoals in het contract vermeld), op cd-rom aan het WODC aangeleverd te worden, zodra het onderzoek is afgerond.

De adressering is:
WODC/SIBa
Postbus 20301
2500 EH Den Haag

Afdeling Extern Wetenschappelijk Beleidsonderzoek (EWB)

Bezoekadres:
Koningskade 4
2596 AA Den Haag

Postadres:
Postbus 20301
2500 EH Den Haag

www.wodc.nl

Ons kenmerk WF-3-2020

Note: Momenteel wordt er gewerkt aan een update van dit document.

De gewenste informatie over de bestanden	
Projectnummer	het WODC projectnummer (kenmerk) zoals vermeld in het contract
Onderzoeksnaam	de (WODC) naam (werktitel) van het onderzoek
Bestandsnaam	de naam van het bestand zoals het wordt aangeleverd
Soort gegevens	de typering van de gegevens; bijv. steekproef, enquête, vragenlijst, gegevensbestand, o.d.
Soort bestand	aanduiding van het soort gegevens in het bestand; bijv. feitelijke data, code/referenties, o.d.
Privacygevoelig	indicatie (ja of nee), of het bestand privacygevoelige gegevens bevat. Nadere informatie, eventueel bij "opmerking"
Detailniveau	(kleinste) eenheid van de variabelen, bijv. persoonsniveau
Gebruikte applicatie	applicatie of toepassing gebruikt bij onderzoek voor het bewerken/bereken van de gegevens; bijv. SPSS, MS Excel
Beschrijving	nadere beschrijving, omschrijving, betekenis van het bestand.

Verzamelmethode	de wijze waarop de gegevens verzameld zijn; bijv. enquête, computergestuurde vragenlijst, of bestand opgeslagen vanuit CBS Statline, of bestand geleverd door organisatie-x
Peildatum/periode	datum waarop de aangeleverde (bestands)gegevens zijn ontstaan, dan wel de periode, waarin zij zijn verzameld
Steekproefmethode	de gevolgde steekproefmethode en opzet van de steekproef; bijv. panel, computergestuurde vragenlijst
Onderzoekseenheid	het niveau waarop het onderzoek (enquête, steekproef) heeft plaatsgevonden; bijv. individueel, persoon
Populatie (doel)	de populatie van de steekproef, bijv. NL-bevolking van 18 jaar en ouder, medewerkers van ..
Steekproefomvang	het aantal uitgezette vragenlijsten (in geval van een steekproef, enquête), bijv. 2546
Respons	het aantal "terugontvangen" vragenlijsten, bijv. 2204 (of in %, bijv. 86,6%)
Aantal variabelen	het aantal variabelen (velden) in het bestand.
Weegfactor J/N	indicatie of er al dan niet een (of meer) weegfactor(en) van toepassing is op de gegevens
Bron	naam van de toepassing of applicatie (of database) waaruit de gegevens afkomstig zijn, bijv. CBS Statline
Bron bestand	naam van het bestand of tabel in de bronapplicatie, bijv. bevolking per postcode
Bron eigenaar	naam van de eigenaar van de bron; bijv. CBS
Aantal rijen	het aantal rijen of records in het aangeleverde bestand (bij steekproef = aantal "terugontvangen" vragenlijsten)
Aantal kolommen	aantal variabelen, velden per rij of record (bij steekproef = aantal vragen, velden per vragenlijst)
Grootte	de omvang van het bestand in Mb
Type/formaat	het formaat, waarin het bestand is vastgelegd, zoals xls (MS Excel), sav (SPSS), doc (MS Word), o.d.
Beperking	bependingen aan het gebruik van de gegevens; bijv. de betrouwbaarheid van de bron of de actualiteit of uitsluitend kwalitatieve gegevens
Opmerking	een nadere toelichting op bijv. het verkrijgen van de gegevens of de inhoud

Afdeling
Extern Wetenschappelijk
Beleidsonderzoek (EWB)

www.wodc.nl

Ons kenmerk
WF-3-2020

Lijst en definities van de variabelen of velden/kolommen per bestand	
Naam variabele	De naam van de variabele(veld of kolom) zoals die in het bestand is opgenomen
Bestand	Naam van het bestand, waartoe het gegeven behoort (zoals het wordt aangeleverd)
Omschrijving	nadere beschrijving, omschrijving, betekenis van de variabele
Type	het type van de variabele, zoals numeriek, alfanumeriek, datum
Lengte	het aantal tekens dat een variabele maximaal heeft
Brongegeven	naam van de variabele (veld of kolom), zoals die in de brontoepassing wordt genoemd of waarvan de variabele is afgeleid, bijv. leeftijd 0 tot 5 jaar in CBS Statline, bevolking, leeftijd
Berekening	de berekening hoe de variabele tot stand is gekomen (algoritme)
Weegfactor	indien van toepassing de voor de variabele gebruikte weegfactor
Beperkingen	beperking aan het gebruik van de variabele; bijv. de actualiteit van het gegeven
Bereik (waarde)	indien van toepassing de minimale en maximale waarde van de variabele
Privacygevoelig	indicatie (ja of nee), of het gegeven privacygevoelig is. Nadere informatie, event. bij "opmerking"
Opmerkingen	een nadere toelichting op bijv. berekening of verkrijgen van het gegeven

**Afdeling
Extern Wetenschappelijk
Beleidsonderzoek (EWB)**

www.wodc.nl

Ons kenmerk
WF-3-2020