

## **Uploadprocedure en voorschriften bestandsformaat**

De gebruiker dient rekening te houden met het volgende:

- Als u de projectovereenkomst gaat ondertekenen, of ondertekend hebt, wijzen wij u op de volgende passage die hierin is opgenomen:  
“Indien Contractant eigen microbestanden aanlevert, verklaart Contractant dat de gegevens rechtmatig zijn verkregen en dat ook de verstrekking van de gegevens aan CBS voldoet aan de eisen van de Algemene verordening gegevensbescherming (AVG) en de uitvoeringswet AVG. Voor zover het gaat om gezondheidsgegevens, moet de verkrijging en verstrekking (mede) aan de eisen van afdeling 7:5 van Boek 7 van het Burgerlijk Wetboek (Wet geneeskundige behandelingsovereenkomst, WGBO) voldoen. CBS kan hierover nadere informatie opvragen.”
- Het CBS zal de bestanden niet gebruiken voor eigen doeleinden noch deze bestanden ter beschikking stellen voor andere Remote Access projecten, tenzij aanleverende instelling hiervoor expliciet schriftelijk toestemming heeft gegeven.
- Mail altijd vooraf naar [microdata@cbs.nl](mailto:microdata@cbs.nl) dat u van plan bent een bestand te gaan uploaden. Vermeld in deze mail welke actie van het CBS wordt verwacht. Geef ook aan hoe deze data past binnen de doelbinding van het project, dat wil zeggen waarom dit bestand noodzakelijk is voor beantwoorden van de onderzoeksvraag van het project.
- Neem het projectnummer op in de naam van het bestand of de zip. Stuur altijd een ingevuld [formulier Upload informatie](#) mee en een recordbeschrijving en zorg dat variabelen voorzien zijn van een omschrijving, anders kan Microdata Services de upload niet in behandeling nemen. Stuur een mail als u de upload hebt uitgevoerd, met als onderwerp uw projectnummer en de naam van het bestand.
- Bij het uploaden van data is het vanwege het beveiligingsbeleid niet toegestaan om executables (.exe) mee te sturen. Ook is het absoluut niet toegestaan om bestanden per mail op te sturen.
- Het CBS wil niet dat in het databestand ***namen van personen, bedrijven of andere instellingen, straatnamen en e-mailadressen in de variabelen voorkomen***; dit om zichtherkenning (het -onbedoeld- herkennen van personen/bedrijven/adressen in de bestanden) te voorkomen.
- Voor de kosten kunt u de [dienstencatalogus](#) raadplegen. De kosten kunnen afwijken indien bestanden niet goed worden aangeleverd. Kosten gelden per te verrinnen/versleutelen (deel)bestand, daarom is het verstandig om zoveel mogelijk uw gegevens in één gecombineerd bestand aan te leveren.

## Inhoudsopgave

|                                                                 |           |
|-----------------------------------------------------------------|-----------|
| <b>1. Inleiding .....</b>                                       | <b>4</b>  |
| <b>2. Uploadprocedure .....</b>                                 | <b>5</b>  |
| Bestandseisen.....                                              | 5         |
| Eigen identificerende variabelen .....                          | 5         |
| Bestandsoverdracht .....                                        | 7         |
| <b>3. Voorschriften .....</b>                                   | <b>11</b> |
| Personen.....                                                   | 11        |
| Adressen.....                                                   | 12        |
| Bedrijven.....                                                  | 12        |
| Overige sleutelvariabelen .....                                 | 12        |
| Open tekstvelden .....                                          | 13        |
| <b>4. Data koppelbaar maken met externe onderzoeksdata.....</b> | <b>14</b> |
| <b>5. Resultaten ná koppeling .....</b>                         | <b>15</b> |

## 1. Inleiding

Het komt regelmatig voor dat onderzoekers een bestand met persoons-, bedrijfs- of adresgegevens willen koppelen aan de beschikbare microdatabestanden van het CBS. Dat is in beginsel mogelijk, mits onderzoekers rechtmatig over die data beschikken en (als het om persoonsgegevens gaat) daar ook toe gerechtigd zijn in het licht van de wetgeving rond de bescherming van persoonsgegevens.

Als u het voornemen hebt een bestand ter koppeling aan te bieden, neem dan altijd eerst via [microdata@cbs.nl](mailto:microdata@cbs.nl) contact op met uw relatiebeheerder en licht toe wat uw bedoeling is en waarom dit bestand nodig is om de onderzoeksvragen in uw project te beantwoorden (doelbinding).

Voor het fysiek aanleveren van een dergelijk bestand (door u zelf, of direct aan CBS door een derde partij die eigenaar is van het bestand) moet u gebruik maken van [Bestandslevering aan het CBS](#).

U dient daarbij altijd een korte omschrijving van de inhoud van het bestand mee te leveren, een overzicht van alle variabelen (recordbeschrijving), een codelijst en te zorgen dat alle variabelen voorzien zijn van een omschrijving. Tevens stuurt u bij de upload een volledig ingevuld [formulier Upload informatie](#) mee. Stuur na uitvoeren van de upload ook altijd een bericht aan [microdata@cbs.nl](mailto:microdata@cbs.nl) onder vermelding van uw projectnummer en bestandsnaam van de upload.

Na het ontvangen van de data zal het CBS de identificerende variabelen versleutelen zodat deze data met de catalogusbestanden van het CBS kunnen worden gekoppeld. Daarbij worden de (direct) identificerende data verwijderd. Eventueel kunnen andere identificerende variabelen op verzoek versleuteld worden en zo beschikbaar blijven in het bestand.

## 2. Uploadprocedure

### Bestandseisen

De gegevens kunnen in drie formaten worden aangeboden:

- SPSS systemfile (.sav )  
Zorg ervoor dat alle variabelen, en indien van toepassing ook de waarden, voorzien zijn van omschrijvingen. Vermijd het gebruik van DATE field voor datumvelden, maar sla datumvelden op als string variabele of numerieke variabele in de vorm JJJJMMDD.
- Tekstbestand met scheidingstekens tussen de velden (.CSV)  
In de 1<sup>ste</sup> regel van dit bestand dienen de namen van de velden vermeld te staan. Wees er verder alert op dat het scheidingsteken niet als teken in een veld voorkomt. Vermeld ook altijd het decimaalscheidingsteken, hanteer één formaat voor datumvelden, gebruik geen 1000-tal scheidingstekens en valuta teken in de velden. Lever altijd een volledige bestandsbeschrijving aan in het zip-bestand.
- ASCII fixed format  
Zorg voor een goede beschrijving van de data. Het bestand mag bovendien niet te breed worden (maximale recordlengte van 8000 tekens). Gebruik dit bestandsformaat alleen als de dataset slechts weinig (<20) variabelen bevat.

Bij de aanlevering in CSV en/of ASCII fixed format is het van belang om een bestandstoelichting mee te leveren waarin een recordbeschrijving is opgenomen, een omschrijving per variabele (label) en indien van toepassing een codeboek (de mogelijke waarden van de variabele met een omschrijving). Zonder deze bestandstoelichting kunnen we de data niet duiden en wordt de dataset niet geaccepteerd.

***Indien u van deze drie formaten wilt afwijken, dient u dit vooraf te overleggen.***

Geef altijd aan welke de variabelen zijn die persoon, bedrijf of adres identificeren (sleutelvariabelen).

### ***Eigen identificerende variabelen***

Indien het bestand eigen identificerende variabelen bevat (recordnummers, persoonsnummers etc.) dan worden die gewoonlijk door het CBS verwijderd. Indien u deze variabelen in het bestand bewaard wilt houden dan kunt u dit aangeven. Dit kan zinvol zijn om op een later moment via deze identificerende variabelen als koppelsleutel aanvullende variabelen als bestand aan te leveren. Deze identificerende variabelen worden wel encrypt, dit om zichtherkenning te voorkomen.

Voorbeeld van een recordbeschrijving (te leveren in txt formaat):

| Variable    | Beginpositie | Eindpositie | Lengte | Type | Decimalen |
|-------------|--------------|-------------|--------|------|-----------|
| Variabele1  | 1            | 6           | 6      | A    | 0         |
| Variabele2  | 7            | 11          | 5      | N    | 3         |
| Variabele3  | 12           | 12          | 1      | A    | 0         |
| Variabele4  | 13           | 16          | 4      | A    | 0         |
| Soort adres | 17           | 18          | 2      | A    | 0         |

Toelichting bij type A=alfanumeriek en N is numeriek

Voorbeeld van een codeboek:

Variabelen omschrijvingen:

Postc: Postcode van het adres.

Huisnr: Huisnummer

Huisltr: Huisnummerletter

Huisnrtoe: Huisnummertoevoeging

Soortadres: Typering van een adres naar woningtype:

Voorkomende waarden bij een variabele

Variabele: Soortadres:

| Code | Omschrijving       |
|------|--------------------|
| 00   | Hoekwoning         |
| 01   | Rijtjeswoning      |
| 02   | Twee onder één kap |
| 03   | Vrijstaand         |
|      | ....               |

### **Bestandsoverdracht**

Voor het opsturen van bestanden maakt u gebruik van de beveiligde uploadfaciliteit van het CBS.

De procedure is als volgt:

Zip het bestand, het ingevulde formulier Upload informatie, het codeboek en de bestand- en variabelen beschrijving, neem het nummer van uw project op in de naam van bestand. . (voorbeeld: 1051LeerlingenAmsterdam.zip)

In een ZIP kunnen meerdere bestanden worden opgenomen, maar deze worden wel afzonderlijk behandeld en in rekening gebracht.

Desgewenst kunt u de ZIP beveiligen met een wachtwoord, dat u doorstuurt naar [microdata@cbs.nl](mailto:microdata@cbs.nl), onder vermelding van uw projectnummer en bestandsnaam.

Een bestand dat via de upload wordt verstuurd mag **niet groter zijn dan 1 Gb**, én het mag **geen executables (.exe bestanden)** bevatten. Indien het zip bestand groter is dan 1 Gb neem dan contact ons op ([microdata@cbs.nl](mailto:microdata@cbs.nl)).

Formulier Upload informatie

De AVG verplicht ons om informatie op te vragen over upload van uw data naar de beveiligde omgeving van de Microdata CBS voor uw project. Hiervoor dient dit [formulier Upload informatie](#).

Dit ingevulde formulier dient u op te slaan als uploadXXXXJJJMMDD.xlsx (XXXX=projectnummer, JJJMMDD is bijvoorbeeld 20180703) en in het zip bestand van de upload te plaatsen, of separaat aan [microdata@cbs.nl](mailto:microdata@cbs.nl) te zenden onder vermelding van uw projectnummer en de naam van het bestand dat via de upload is aangeboden.

Nadat u deze voorbereidingen hebt getroffen kunt U het ZIP bestand naar het CBS uploaden door te klikken op: <http://www.cbs.nl/bestandslevering>

**Bij het invullen van de internet pagina dient u de volgende codes te gebruiken:**

**Enquêtecode: CVB**

**Correspondentienummer: 2008-111-401**

**Controlenummer: 817808**

U komt dan op het scherm dat u op de volgende pagina ziet:



## Inzenden van een bestand met opgaven

Bij vragen kunt u contact opnemen met het Contact Center CBS:

Telefoon: (045) 570 64 00 voor inhoudelijke vragen, op werkdagen tussen 9-17 uur.  
(045) 570 66 27 voor technische vragen, op werkdagen tussen 10-12 en 14-16 uur.

E-mail: [contactcenter@cbs.nl](mailto:contactcenter@cbs.nl)

### Vul eerst deze gegevens in:

De enquêtecode, uw correspondentienummer en uw controlnummer staan vermeld in de brief of e-mail die u is toegezonden door het CBS samen met de link naar deze pagina.

|                           |                                                                    |
|---------------------------|--------------------------------------------------------------------|
| De enquêtecode:           | <input type="text"/>                                               |
| Uw correspondentienummer: | <input type="text"/> - <input type="text"/> - <input type="text"/> |
| Uw controlnummer:         | <input type="text"/>                                               |

### Contactpersoon:

Indien de contactinformatie in de brief of e-mail niet meer correct is, kunt u hier de juiste gegevens invullen

|                 |                                                        |                      |
|-----------------|--------------------------------------------------------|----------------------|
| Naam:           | <input type="radio"/> Dhr. <input type="radio"/> Mevr. | <input type="text"/> |
| Telefoonnummer: |                                                        | <input type="text"/> |
| E-mailadres:    |                                                        | <input type="text"/> |

Afdrukken

Verder

Bij het invullen van de internet pagina dient u de volgende codes te gebruiken:

Enquêtecode: CVB

Correspondentienummer: 2008-111-401

Controlnummer: 817808

Uw gegevens bij contactpersoon invullen en klik dan op verder.




Centraal Bureau voor de Statistiek

### Inzenden van een bestand met opgaven

Bij vragen kunt u contact opnemen met het Contact Center CBS:

Telefoon: (045) 570 64 00 voor inhoudelijke vragen, op werkdagen tussen 9-17 uur.  
(045) 570 66 27 voor technische vragen, op werkdagen tussen 10-12 en 14-16 uur.

E-mail: [contactcenter@cbs.nl](mailto:contactcenter@cbs.nl)

---

|                               |                     |
|-------------------------------|---------------------|
| <b>Enquêtecode:</b>           | <b>CVB</b>          |
| <b>Correspondentienummer:</b> | <b>2008-111-401</b> |

**Verzenden van het bestand:**  
Druk na het invullen van de gevraagde gegevens op de knop [ **Verzenden** ] onderaan de pagina. Het verzenden van het bestand start direct. Bij grote bestanden kan dit enige tijd in beslag nemen, afhankelijk van de verbinding.  
Zodra het bestand is ontvangen krijgt u een ontvangstbevestiging van het CBS.

**Te verzenden bestand**

**LET OP:**  
Indien u tussentijds de pagina afsluit of opnieuw op de verzendknop drukt wordt slechts een deel van uw bestand bij het CBS afgeleverd. Wacht dus op de ontvangstbevestiging, hierin staat onder meer de grootte van het ontvangen bestand aangegeven.  
Wij adviseren geen bestanden groter dan 600 MB te verzenden.

Hierna dient u het zip-bestand te selecteren (Bladeren..) wat u gaat uploaden en klikt u vervolgens op verzenden. Hierna wacht u af (dit kan even duren) totdat u de volgende melding (ontvangstbevestiging) op uw scherm krijgt dat de upload is gelukt (zie volgende pagina).



**Centraal Bureau voor de Statistiek**

### Inzenden van een bestand met opgaven

Bij vragen kunt u contact opnemen met het Contact Center CBS:

Telefoon: (045) 570 64 00 voor inhoudelijke vragen, op werkdagen tussen 9-17 uur.  
(045) 570 66 27 voor technische vragen, op werkdagen tussen 10-12 en 14-16 uur.

E-mail: [contactcenter@cbs.nl](mailto:contactcenter@cbs.nl)

---

|                               |                     |
|-------------------------------|---------------------|
| <b>Enquêtecode:</b>           | <b>CVB</b>          |
| <b>Correspondentienummer:</b> | <b>2008-111-401</b> |

### Bedankt voor het inzenden van uw opgave

Bestandsnaam: Upload informatie.xlsx  
Bestandsgrootte: 15.937 bytes  
Datum / tijd ontvangst: 03-07-2018 09:03:15

**Afdrukken**

U kunt deze ontvangstbevestiging afdrukken voor uw eigen administratie.

**Stuur altijd een e-mail** naar [microdata@cbs.nl](mailto:microdata@cbs.nl) als u een upload heeft gedaan, met als onderwerp het projectnummer en de naam van het zip bestand.

Het geüploade bestand komt doorgaans de volgende dag bij ons binnen.

### Doorlooptijd

Wij streven ernaar om de verwerking van uw bestanden binnen 2 weken af te hebben. Vaak kan het iets sneller en soms duurt het iets langer.

### 3. Voorschriften

#### Personen

De meeste onderzoekers willen een bestand koppelen op persoonsniveau. Hiervoor moeten personen uit het bestand dezelfde koppelsleutel krijgen als die binnen het CBS wordt gebruikt. Dit is de RINPERSOON. Het proces van het omzetten van persoon identificerende kenmerken in het een bestand naar het RINPERSOON noemen we het 'verrinnen' van een bestand.

Dat gebeurt veelal op BSN, op A- nummer, of op onderwijsnummer. Het is ook mogelijk om met een combinatie van geslacht, geboortedatum, postcode en/of adres of overlijdensdatum personen binnen de GBA te identificeren. Bij adresgegevens is het essentieel dat een peildatum van het moment dat de desbetreffende persoon daar heeft gewoond opgegeven is.

Bij het verrinnen van een bestand verwijdert het CBS de identificerende variabelen uit het bestand en alleen RINPERSOON wordt teruggegeven. Ook unieke nummers die mogelijk naar personen zijn te herleiden, bijvoorbeeld een intern nummer, worden verwijderd. De reden hiervan is om te voorkomen dat gebruikers met deze variabelen een link met het originele bestand kunnen leggen en zo vertrouwelijke persoonsgegevens uit gekoppelde CBS-bestanden kunnen halen.

Voorschrift waaraan de essentiële koppelvariabelen moeten voldoen van bestanden zonder BSN, A-nummers of onderwijsnummers:

| Veld                                      | Formaat | Voorbeeld     |
|-------------------------------------------|---------|---------------|
| Geslacht                                  | A1      | M of V        |
| Geboortedatum                             | N8      | 19590202      |
| Overlijdensdatum                          | N8      | 20191126      |
| Postcode6                                 | A6      | 2498CM        |
| Postcode4                                 | A4      | 2498          |
| Huisnummer                                | N5      | 1             |
| Huisletter                                | A1      | A             |
| Huisnummertoevoeging                      | A4      | IIH of fl01   |
| Jaar van geldigheid adres<br>of Peildatum | N4/N8   | 2010/20101231 |

Zie ook *Eigen identificerende variabelen* op pagina 5.

## Adressen

Het komt ook voor dat u op basis van adresgegevens een bestand koppelbaar wilt maken met CBS data. De koppelsleutel is RINOBJECTNUMMER (verrinobjecten).

Dit kan op basis van de BAG\_ID of VBO\_ID of via postcode, huisnummer, huisletter, huisnummertoevoeging en jaar ( of peildatum) van geldigheid van de adresgegevens.

Bij het verrinnen van een bestand verwijdert het CBS de identificerende variabelen die gebruikt zijn voor het RINOBJECTNUMMER.

Let op dat het jaar van geldigheid of peildatum van het adres altijd is opgenomen in het bestand anders kan er geen versleuteling plaatsvinden naar RINOBJECTNUMMER.

Voorschrift waaraan de essentiële koppelvariabelen op adres moeten voldoen:

| Veld                                | Form  | Voorbeeld             |
|-------------------------------------|-------|-----------------------|
| Postcode                            | A6    | 2498CM                |
| Huisnummer                          | N5    | 1; 12345              |
| Huisletter                          | A1    | A of a                |
| Huisnummertoevoeging                | A4    | IIH of fl01           |
| Combinatie huisnr+letter+toevoeging | A10   | 1A; 1AFL10;19999AFL10 |
| Jaar geldigheid adres of peildatum  | N4/N8 | 2010/20140131         |

Zie ook *Eigen identificerende variabelen* op pagina 5.

## Bedrijven

Bij bedrijven kan de koppeling met CBS bestanden gelegd worden via het Kamer van Koophandel nummer (KvKnr) of het fiscaal identificatienummer (FI-nr). Het KvKnr en/of FI-nr wordt daartoe versleuteld zodat deze via het microdatabestand 'Algemeen Bedrijven Register' (ABR) gekoppeld kan worden aan een Bedrijfseenheid (BE), de statistische eenheid voor veel bedrijfseconomische statistieken. Doel van die versleuteling is tevens om te voorkomen dat direct vertrouwelijke bedrijfseconomische gegevens uit gekoppelde CBS-bestanden gehaald kunnen worden.

Voorschrift waaraan de essentiële koppelvariabelen moeten voldoen:

| Veld  | Formaat | Voorbeeld | Bijzonderheden      |
|-------|---------|-----------|---------------------|
| Kvknr | A8      | 00123456  | Met voorlooppnullen |
| FI-nr | A9      | 000123456 | Met voorlooppnullen |

## Overige sleutelvariabelen

Ook kunnen andere identificerende variabelen uit geüploade bestanden worden gebruikt om koppelingen te maken met CBS bestanden, zoals aan schoolgegevens met BRIN als koppelsleutel. Voor eventuele overige mogelijkheden dient u contact op te nemen met [Microdata@cbs.nl](mailto:Microdata@cbs.nl).

### **Open tekstvelden**

De bestanden mogen in principe geen open tekstvelden (dus antwoorden op open vragen) bevatten. Mochten zulke velden noodzakelijk zijn en geen identificerende informatie bevatten, dan dient dit expliciet aangegeven te worden en moet een overzicht gegeven worden van de inhoud van de variabele ter controle op identificerende informatie. Voor deze controle kunnen kosten in rekening gebracht worden.

#### **4. Data koppelbaar maken met externe onderzoeksdata**

Microdata Services faciliteert ook de mogelijkheid om door onderzoekers geselecteerde externe data koppelbaar te maken met CBS databestanden over personen. Het CBS heeft daartoe een koppelbestanden tussen de populatie van de diverse onderzoeken en de persoonsgegevens van het CBS beschikbaar in de Remote access omgeving. Via deze koppeling kunnen data uit die diverse onderzoeken gebruikt worden in een microdataonderzoek bij het CBS. Het betreft de volgende externe onderzoeksdata:

- DNB Household survey (DHS);
- LIFELINES Cohort;
- LISSPANEL;

Zie hiervoor [externe onderzoeksdata](#) waar u de procedures kunt vinden.

## **5. Resultaten ná koppeling**

### **Resultaten van de verrining van personen:**

De kwaliteit van de koppeling hangt af van de kwaliteit van de aangeleverde sleutelvariabelen. Er vanuit gaande dat dit in orde is, kunnen we onderstaande percentages verwachten van het verrinnen van persoonsbestanden:

- Op BSN (of bijv. onderwijsnummer) is in principe 100% te verrinnen. In de praktijk kan dit echter een paar procent minder zijn, bijvoorbeeld door verschillen in peilmoment en/of populatie (niet in de GBA ingeschrevenen)
- Op geboortedatum, geslacht, adres en peildatum van het adres is rond de 90% te verrinnen.
- Op geboortedatum, geslacht, en postcode4 en jaar van geldigheid van de postcode is rond de 80% te verrinnen.

Van elke koppeling wordt een rapport opgesteld dat ter beschikking wordt gesteld.

### **Resultaten van het verrinobjecten van adressen:**

Het resultaat van het omzetten van VBO\_ID of BAG\_ID naar RINOBJECTNUMMER is 100%.

Op basis van de adresgegevens is dit rond de 90%.

### **Resultaten van versleuteling van bedrijfs indentificatoren:**

KvK-nummers en FI-nummers versleutelen lukt altijd, maar of er een corresponderende bedrijfseenheid (BE) in het ABR bestaat kan van verschillende factoren afhangen.

In tegenstelling tot persoonsstatistieken, waar voor het verrinnen altijd sprake is van een 1-op-1 relatie, ligt de situatie met bedrijven een stuk ingewikkelder. Een ‘bedrijfseenheid’ (BE) in het ‘Algemeen Bedrijven Register’ (ABR) kan meerdere Kamer van Koophandel (KvK)-nummers en/of Fiscale (FI)-nummers omvatten. Ook kan het voorkomen dat een KvK- of FI-nummer niet in een BE komt, bijvoorbeeld omdat deze feitelijk niet economisch actief is. Voor verdere details wordt verwezen naar de documentatie, en met name naar de bijlagen van deze documentatie, van het [ABR](#).