

Aanleverspecificatie PVM CBS

1. Inleiding

Met de PVM CBS is het mogelijk om:

1. Persoonsgegevens ter pseudonimisatie aan te bieden (paragraaf 4);
2. Pseudoniemen ter conversie aan te bieden (paragraaf 5).

2. Bestandsindeling

De Privacy Verzend Module (PVM) CBS leest de te pseudonimiseren gegevens uit een bestand in CSV formaat met de volgende kenmerken:

- Het scheidingsteken is de puntkomma (;)
- Het quote-teken is het dubbele aanhalingsteken ("")
- Elke regel van het bestand bevat evenveel kolommen
- De tekenset van het bestand is ISO-8859-1
- De bestandsnaam dient te eindigen op de extensie .csv

3. Bestandsnaamconventie

Er wordt door de PVM CBS geen controle op de bestandsnaam uitgevoerd. ZorgTTP raadt aan een conventie te hanteren om eventuele afstemming over een aanlevering te vergemakkelijken. Maak daarbij onderscheid tussen een test- of productie aanlevering, de databron, de aanleverdatum en eventueel volgnummer.

Naam	Betekenis
cbs_prod_\$afzender\$_20180801_1.csv	Productie aanlevering van \$afzender\$ op 1 augustus 2018, 1 ^e bericht
cbs_test_\$afzender\$_20180201_2.csv	Test bericht van \$afzender\$ op 1 februari 2018, 2 ^e poging

4. Pseudonimisatie

Voor de verwerking van bestanden maakt de PVM gebruik van kolomlabels, de kolomlabels moeten op de eerste rij staan. De volgende kolomlabels worden als input gebruikt voor pseudonimisatie: **BSN**, **GEBDAT**, **GESLACHT** en **PC6**. In onderstaande tabel zijn de criteria voor de variabelen opgenomen.

Kolom	Criteria
VOLGNR	Betekenisloos nummer dat gebruikt kan worden voor de terugkoppeling door het CBS.
BSN	Een geldig burger service nummer met een lengte van 8 of 9 karakters, alléén cijfers.
GEBDAT	Een datum in 'EEJJMMDD' formaat van 8 karakters ná 1 januari 1850 en niet in de toekomst, 27 april 2011 wordt onder deze conventie bijvoorbeeld 20110427.
GESLACHT	1 karakter waarbij 1, M of m voor mannelijk en 2, V of v voor vrouwelijk staat, 0 voor onbekend en 9 voor niet gespecificeerd.
PC6	6 karakters voor een Nederlandse postcode van 4 cijfers en 2 letters A-Z, NNNNAA.

Genoemde persoonsgegevens worden gebruikt om drie pseudoniemen aan te maken:

1. Pseudoniem op basis van het BSN;
2. Pseudoniem op basis van geboortedatum, geslacht en numerieke deel van de postcode;
3. Pseudoniem op basis van geboortedatum, geslacht en volledige postcode.

De genoemde persoonsgegevens worden gebruikt voor het genereren van pseudoniemen en daarna verwijderd uit de dataset. Alle overige variabelen worden ongewijzigd aangeboden aan het CBS.

Voorbeeld input

VOLGNR;GESLACHT;GEBDAT;PC6;BSN

1;2;19750301;3448KX;123456782

2;1;19671020;2215PR;

3;2;19620624;2321HG;

5. Conversie

Opdrachtgevers van ZorgTTP kunnen met de PVM CBS pseudoniemen aan ZorgTTP aanbieden ten behoeve van het CBS. ZorgTTP zal de pseudoniemen van het domein van de databron converteren naar het domein van het CBS.

Voor de verwerking van bestanden maakt de PVM CBS gebruik van kolomlabels. De volgende kolomlabels worden als input gebruikt voor domeinconversie: **PSEUDONIEM BSN**, **PSEUDONIEM CBS** en **PSEUDONIEM NIVEL**.

Alle overige variabelen worden ongewijzigd aangeboden aan het CBS.

Kolom	Criteria
PSEUDONIEM BSN , PSEUDONIEM CBS of PSEUDONIEM NIVEL	Een ZorgTTP pseudoniem, bestaande uit een prefix van 4 karakters en een waarde van 28 karakters. Alle pseudoniemen in deze kolom moeten afkomstig zijn uit hetzelfde domein.

De informatie uit het inputbestand in andere dan de genoemde, voor pseudonimisatie relevante kolom wordt versleuteld door de keten getransporteerd en *onveranderd* bij het CBS afgeleverd.

Voorbeeld input

```
VOLGNR;PSEUDONIEM BSN;PSEUDONIEM CBS;PC4;DATONTVANGST;LIS_BIOPT
1; TP2B0001g9davihecoh429lot182t6gb;TP2C0027kd9sfil1t89mp6vjdn79gj71c;3995;12-10-2018;LIS
```

Waarbij de PVM een bestand alleen verwerkt wanneer de eerste regel kolomnamen bevat en de kolom 'PSEUDONIEM BSN', 'PSEUDONIEM CBS' of 'PSEUDONIEM NIVEL' daarin voorkomt

Voor het aanleveren van (te pseudonimiseren) persoonsgegevens worden de volgende labels gebruikt. Deze labels mogen **NIET** in combinatie met het aanleveren van (te converteren) pseudoniemen worden gebruikt:

- BSN Burgerservicenummer
- GEBDAT Geboortedatum
- GESLACHT Geslacht
- PC6 Volledige postcode