

# Uitsluiten onbekende aanvragers en nieuwe aanvragers buitenlands adres Zorg

## Versies van het model

Versie nummer	Datum	Initialen	Belangrijkste wijziging
<b>1.0</b>	30-06-2014	10.2.e	Initieel document
<b>1.1</b>	08-07-2014		Review
<b>1.2</b>	28-10-2014		Inlezen bestand uit Populatie Beheer Tool
<b>1.3</b>	05-01-2015		2 losse programma's samengevoegd tot Uitsluitlijst Zorg.egg
<b>1.4</b>	25-03-2015		Koppelen met conceptbeschikkingen en selectie van de kolom jaarbedrag t.b.v. 10.2.e
<b>1.5</b>	24-12-2015		Aanvullen informatie
<b>1.6</b>	03-05-2016		AWS-aanpassingen

# 1. INHOUDSOPGAVE

1.	Inhoudsopgave.....	3
2.	Beschrijving proces.....	4
2.1.	Bron data.....	4
2.2.	Velden in het model.....	4
2.3.	Parameters.....	5
2.4.	Extra.....	6

## 2. BESCHRIJVING PROCES

Het SAS-script **Uitsluitlijst ZORG.egp** wordt de dag van de risicoselectie-run gedraaid door het risicoselectieteam. Dit resulteert in meerdere exportbestanden, waaronder,

**Uitsluiten\_Zorg\_<Datum>\_PBT.csv** wat als het halfproduct wordt beschouwd. Ook wordt een csv-bestand **MT\_Rapportage\_data\_Zorg** geëxporteerd met kengetallen voor de Managementrapportage.

Zowel het SAS-script als de exportbestanden staan opgeslagen op de AWS:

Script:

*/prj/bi/data/AD010/data/RisicoSelectie/Halfproducten/UITSLUITLIJST\_ZORG*

Exportbestanden:

*/prj/bi/data/AD010/data/RisicoSelectie/Halfproducten/UITSLUITLIJST\_ZORG/Export*

### 2.1. BRONDATA

De input gebruikt in het SAS-script bestaat uit vijf bronbestanden:

#### 1: **civ\_rs\_afgew\_zorgaanvragen\_2.sas7bdat**

Dit bestand bevat een lijst van BSNs met afgewezen aanvragen zorgtoeslag in de afgelopen jaren (geen recht vanwege buitenlands adres). Het bestand wordt door het risicoselectieteam vóór risicoselectie-run aangevraagd bij team BCA/BICC. Uiterlijk om 09.00 uur op de dag van de risicoselectie wordt de brondata opgeleverd in de vorm van sas-bestand op de AWS:

*/prj/bi/data/AD010/data/RisicoSelectie/CIV*

#### 2: **Overzichtresultaat.xlsx**

Dit bestand bevat een lijst van eerder behandelde aanvragen. Voor het bestand zie de map:

*/prj/bi/data/AD010/data/RisicoSelectie/Brondata/EERDERE ACTIES/Source*

Voor de documentatie zie de map:

*Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling 2013\Brondata\Eerdere Acties\Documentatie*

#### 3: **Zorgtoeslag\_niet\_verzekerd.xlsx**

Uitsluiten onbekende aanvragers  
en nieuwe aanvragers buitenlands adres Zorg

Dit bestand bevat een lijst van BSNs van mensen die aanvraag hebben gedaan voor zorgtoeslag, maar niet verzekerd zijn.

Het bronbestand bevindt zich in de volgende map:

*Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Niet\_zorgverzekerd\Controles jaar\Maand jaar*

Het meest actuele bestand (Zorgtoeslag\_niet\_verzekerd\_ddMNDj.xlsx) wordt gekopieerd op de AWS in de volgende map en hernoemd:

*/prj/bi/data/AD010/data/RisicoSelectie/Halfproducten/UITSLUITLIJST\_ZORG*

#### **4: civ\_prof\_nieuwe\_aanvragers.sas7bdat**

Dit bestand bevat een lijst van onbekende aanvragers. Voor het bestand zie de map:

*/prj/bi/data/AD010/data/RisicoSelectie/CIV*

Voor de documentatie zie de map:

*Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling 2013\Brondata\ONBEKENDE AANVR\Documentatie*

#### **5: civ\_prof\_tokens.sas7bdat**

Dit bestand bevat een lijst van toeslagaanvragen HT en Zorg door Vluchtelingen. Voor het bestand zie de map:

*/prj/bi/data/AD010/data/RisicoSelectie/CIV*

Voor de documentatie zie de map:

*Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling2013\Brondata\Vluchtelingenwerk\Documentatie*

## **2.2. VELDEN IN HET MODEL**

De output van dit model bestaat zoals eerder vermeld uit twee export-bestanden:

Uitsluiten\_Zorg\_<Datum>\_PBT.csv

Name	Getalnotatie (csv)
<b>BSN</b>	Standaard
<b>Regeling</b>	Standaard
<b>Jaar</b>	Standaard
<b>Reden</b>	Standaard
<b>Jaarbedrag</b>	Standaard

Uitsluiten onbekende aanvragers  
en nieuwe aanvragers buitenlands adres Zorg

MT\_Rapportage\_data\_Zorg.csv

Name	Getalnotatie (csv)
<b>Categorie</b>	Standaard
<b>Dimensie</b>	Standaard
<b>Aantal</b>	Standaard

Categorie is de categorie waartoe de resultaten behoren. Dit kan zijn 'Data rapportage Zorg', 'Buitenland' of 'Onbekend'. Dimensie betreft de subgroepen waarin de categorieën zijn onderverdeeld. Mogelijkheden zijn bijvoorbeeld 'Aanvrager woont op buitenlands adres' of 'Onbekende aanvrager'. De waarde van 'Aantal' is het kengetal/resultaat behorende bij de dimensie.

## 2.3. FLOWS

### 2.3.1. INIT

#### Libnames

Library's worden gedefinieerd.

#### Macros AWS

In dit script worden de volgende variabelen, parameters en macro's gezet voor de AWS-omgeving: controlejaar, anderjaar, periode, datum-variabelen en de macro's voor het wegschrijven.

Macros lokaal

In dit script worden de volgende variabelen, parameters en macro's gezet voor de lokale server: controlejaar, anderjaar, periode, datum-variabelen en de macro's voor het wegschrijven.

### 2.3.2. ZORG-BESTANDEN VERWERKEN

In het script worden de bronbestanden samengevoegd tot een uitsluitlijst met reden voor Zorg, zodanig dat BSNs die eerder uitgesloten zijn niet nog een keer uitgesloten worden.

**Stap 1:** De drie uitsluitlijsten Zorg CB afgewezen aanvragen-3.txt, Zorgtoeslag\_niet\_verzekerd.xlsx en Overzichtresultaat.xlsx ingelezen.

#### Verwerken

Uitsluiten onbekende aanvragers  
en nieuwe aanvragers buitenlands adres Zorg

**Stap 1:** Ieder van de drie ingelezen tabellen wordt gesorteerd op BSN (en Toeslagjaar) en ontdubbeld.

**Stap 2:** De bestanden worden gekoppeld met een merge statement op BSN. Indien een BSN voorkomt in het bestand van de afgewezen aanvragers zorg en NIET in het bestand waarin eerdere selecties zitten die zijn behandeld en ook NIET in de onverzekerden zorg, wordt het BSN met Toeslagjaar en Regeling meegegeven in de outputtabel **ZORG\_UITSL\_FINAL**.

**Stap 3:** Deze tabel wordt voorzien van een datum (**UITSL\_ZORG\_BTL\_ADR\_&datum**) en naar de volgende locatie geplaatst.:

*/prj/bi/data/AD010/data/RisicoSelectie/Halfproducten/UITSLUITLIJST\_ZORG/Results*

De tabel wordt weggeschreven zodat deze bewaard wordt voor de volgende run.

#### Inlezen en verwerken uitsluitlijsten zorg

**Stap 1:** Een selectie wordt gemaakt van de zorgtoeslag aanvragen uit inputbestand **nieuwe\_aanvragers**, i.e. onbekende aanvragers, die aan de volgende eis voldoen: Zorg > 0. Het resultaat wordt gegroepeerd per BSN opgeslagen in een nieuwe tabel **UITSLUIT\_NIEUWE\_AANV.**

**Stap 2:** Uit de tabel **ZORG\_UITSL\_FINAL** (samengevoegde uitsluitlijst) wordt een selectie gedaan op het huidige toeslagjaar. De velden BSN en Toeslagjaar worden opgeslagen in **UITSLUIT\_BTL\_ADR.**

**Stap 3:** De meest actuele uitsluitlijsten aangeleverd door B/CA BICC worden ingelezen. Hierbij wordt een selectie gedaan op de toeslagaanvragen waarbij de regeling (variabele 'regeling') gelijk is aan 'ZORG'.

**Stap 4:** De drie uitsluitlijsten, ingelezen in de vorige stap, worden samengevoegd. Hieraan wordt een veld 'Toeslagjaar' toegevoegd. De waarde van dit veld wordt gelijk aan het huidige toeslagjaar indien de begin van de aanvraag (variabele 'Begin') niet later is dan het toeslagjaar en het einde van de aanvraag (variabele 'Eind') niet eerder is dan het desbetreffende jaar. Deze informatie wordt opgeslagen in de tabel **Uitsluitlijst\_temp**

**Stap 5:** De zelf geproduceerde uitsluitlijsten zorg (**UITSLUITEN\_ZORG\_yyyymmdd**) van de twee vorige runs worden ingelezen.

*(/prj/bi/data/AD010/data/RisicoSelectie/Halfproducten/UITSLUITLIJST\_ZORG/Results)*

Uitsluiten onbekende aanvragers  
en nieuwe aanvragers buitenlands adres Zorg

Pagina 7/10

**Stap 6:** De zelf geproduceerde uitsluitlijsten zorg buitenlands adres (UITSL\_ZORG\_BTL\_ADR\_yyyymmdd) van de twee vorige runs worden ingelezen.  
([/prj/bi/data/AD010/data/RisicoSelectie/Halfproducten/UITSLUITLIJST\_ZORG/Results)

**Stap 7:** De tabel **Uitsluitlijst** wordt aangemaakt. Dit is een samengevoegde lijst van de velden BSN en Toeslagjaar uit het bestand **Uitsluitlijst\_temp** en de outputtabellen van stap 5 en 6. Deze tabel bevat de zorgtoeslagaanvragen die recent zijn uitgesloten.

**Stap 8:** De outputtabel uit de vorige stap (**Uitsluitlijst**), **uitsluit\_nieuwe\_aanv** en **uitsluit\_btl\_adr** worden gesorteerd en ontdubbeld.

**Stap 9:** De tabel **uitsluiten2** wordt aangemaakt. Dit is een lijst van alle onbekende aanvragers die niet eerder voorkwamen op de uitsluitlijst. Deze lijst wordt gecreëerd door middel van een merge statement op de tabellen **uitsluit\_nieuwe\_aanv** en **uitsluitlijst**. De BSNs die wel voorkomen in de eerste tabel en NIET in de tweede tabel worden hieraan toegevoegd.

**Stap 10:** Het bestand **civ\_prof\_tokens** met toeslagaanvragen door vluchtelingen wordt ingelezen.

**Stap 11:** Van de tijdelijke uitsluitlijst **uitsluiten2** worden de BSNs verwijderd die op de lijst Vluchtelingenwerk voorkomen en dus niet mogen worden uitgesloten. Dit resulteert in het bestand **uitsluiten2A**, de definitieve uitsluitlijst met onbekende aanvragers.

**Stap 12:** De uitsluitlijst **uitsluiten2B** van aanvragers wonend op een buitenlands adres en niet eerder voorgekomen op de uitsluitlijst, wordt gecreëerd. Deze tabel bevat alle aanvragen uit het bestand **UITSLUIT\_BTL\_ADR** die niet voorkomen in **uitsluitlijst**.

**Stap 13:** De outputtabellen onbekende aanvragers en buitenlands adres worden samengevoegd. Hieraan wordt het veld 'Regeling' toegevoegd met voor iedere regel de waarde 'Zorg'. Dit resulteert in de tabel **uitsluiten3**.

**Stap 14:** De bestanden **civ\_prof\_conceptbes&controlejaar** en bestanden **civ\_prof\_conceptbes &anderjaar** worden gekoppeld aan uitsluiten3 om het jaarbedrag horende bij de toeslagaanvragen op te kunnen halen. Deze koppeling gebeurt door middel van Left Joins op BSN. Het resultaat wordt opgeslagen in de tabel **UITSLUITEN\_ZORG\_&datum**.

**Stap 15:** De tabel **UITSLUITEN\_ZORG\_&datum** wordt weggeschreven als csv-bestand 'Uitsluiten\_Zorg\_&DateRun.\_PBT.csv' naar de volgende locatie:

/prj/bi/data/AD010/data/RisicoSelectie/Halfproducten/UITSLUITLIJST\_ZORG/Export/

Wegschrijven lokaal

Uitsluiten onbekende aanvragers  
en nieuwe aanvragers buitenlands adres Zorg

Pagina 8/10



In dit script wordt het eindbestand **UITSLUITEN\_ZORG\_&datum.** gedownload en weggeschreven als csv-bestand naar de volgende locatie(huidige run):

*Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling 2013\Profiling\huidig jaar-volnummer over toeslagjaar\Results\*

## EXTRA

Kolom jaarbedrag wordt geselecteerd t.b.v.  Daarvoor moet een left join gemaakt worden met de conceptbeschikkingen huidig toeslagjaar en vorig toeslagjaar (aangeleverd door B/CA BICC). De locatie van de conceptbeschikkingen is:  
*/prj/bi/data/AD010/data/RisicoSelectie/CIV*

### 2.3.3. RAPPORTAGE

In dit script worden de kengetallen gegenereerd die worden gebruikt als input voor het periodieke Management Rapportage betreffende de resultaten van de risicoselectie.

Libname

Vaststellen van library.

#### 1. Onbekend - buitenland

In deze flow wordt het aantal aanvragers zorgtoeslag woonachtig op een buitenlands adres en het aantal onbekende zorgtoeslagaanvragers berekend.

**Stap 1:** De huidige risicoselectierunperiode en huidige datum worden bepaald.

**Stap 2:** Het aantal aanvragers woonachtig op een buitenlands adres wordt geteld. Dit zijn alle aanvragers in het

**UITSLUITEN\_ZORG\_&datum.** met de volgende waarde voor de variabele reden: 'Nieuwe aanvrager op buitenlands adres'.

**Stap 3:** Het aantal onbekende aanvragers wordt geteld. Dit zijn alle aanvragers in het bestand **uitsluiten4** met de volgende waarde voor de variabele reden: 'Onbekende aanvrager'

**Stap 4:** Een output-bestand met het aantal aanvragers woonachtig in het buitenland en het aantal onbekende aanvragers wordt gecreëerd.

#### 3. Export

Het output-bestand gecreëerd in de vorige flow wordt hier als xlsx-bestand 'MT\_Rapportage\_data\_Zorg' geëxporteerd naar de Q-schijf op de locatie (huidige run):

Uitsluiten onbekende aanvragers  
en nieuwe aanvragers buitenlands adres Zorg

*Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling 2013\Profiling\huidig jaar-volgnummer over toeslagjaar*

## 2.4. CHECKS MANAGEMENTRAPPORTAGE

Deze checklist is een lijst welke doorgelopen dient te worden vóór het runnen van de procesflow 'Rapportage'.

Voor het runnen van de managementrapportage wordt aangenomen dat:

1. Er een juist en volledig eindbestand (UITSLUITEN\_ZORG\_&datum.) van het halfproduct gemaakt is. De rapportage flow moet dus altijd als laatste procesflow gedraaid worden.

Om de managementrapportage te runnen moet het volgende gedaan worden:

1. Run de process flow 'Rapportage.' in het SAS Enterprise Guide project 'Uitsluitlijst\_ZORG'.
2. Deze process flow schrijft alle output naar 1 .csv bestand:

*Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling 2013\Profiling\huidig jaar-volgnummer over toeslagjaar\MT\_Rapportage\_data\_Zorg.xlsx*

De gegevens uit deze file moet handmatig gekopieerd worden naar de tab 'Data ZORG' in de file:

*Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling 2013\Profiling\huidig jaar-volgnummer over toeslagjaar\MT\_Rapportage\_Template.xlsx*

Let er hierbij goed op dat de gegevens naar exact dezelfde cellen worden gekopieerd (selecteer cellen, kopieer, en plak op tab 'Data ZORG' in cel A1.). Bij deze actie mogen er geen gegevens uit voorgaande risicoselectie-runs in de tab achtergebleven zijn.

3. Dit zorgt ervoor dat automatisch de juiste getallen getoond worden in de tab 'Resultaten ZORG' in het rapport. Doe hier echter voor de zekerheid nog een sanity check op!

Wanneer ook de informatie voor HT en KOT juist is opgenomen in het rapport (zie hiervoor de documentatie van 'HT\_Model' en het 'KOT\_Model') wordt dit document opgeslagen als **MT\_Rapportage\_jjjjmmdd.xlsx**. Als de volledige controlematrix voor de risicoselectie is afgerond, kan het rapport toegestuurd worden aan de regisseur van de risicoselectie, die het rapport uiteindelijk deelt met het management.

Uitsluiten onbekende aanvragers  
en nieuwe aanvragers buitenlands adres Zorg

Pagina 10/10



# Signaallijst adressen

## Versies van het model

Versie nummer	Datum	Initialen	Belangrijkste wijziging
<b>1.0</b>	15-05-2013	10.2.e	Initieel document
<b>1.1</b>	15-05-2013		Aanpassing Key aan Bag lijst
<b>1.2</b>	30-05-2013		Aanpassing brondata
<b>1.3</b>	21-12-2015		Informatie aanvullen
<b>1.4</b>	29-04-2016		AWS-aanpassingen

# 1. INHOUDSOPGAVE

1.	Inhoudsopgave.....	3
2.	Beschrijving proces.....	4
2.1.	Bron data.....	4
2.2.	Velden in het model.....	4
2.3.	Flows.....	4

## 2. BESCHRIJVING PROCES<sup>1</sup>

Het SAS-script **Signaallijst Adressen.egp** wordt één dag voor de risicoselectie-run gedraaid door het risicoselectieteam. Dit resulteert in het halfproduct **signaallijstadressen.sas7bdat**, welke wordt gebruikt tijdens de risicoselectie-run.

Zowel het SAS-script als het halfproduct staan opgeslagen op de AWS:

Script:

```
/prj/bi/data/AD010/data/RisicoSelectie/Halfproducten/UITGESLOTEN_ADRES/SAS  
Enterpriseguide
```

Halfproduct:

```
/prj/bi/data/AD010/data/RisicoSelectie/Halfproducten/UITGESLOTEN_ADRES/Results
```

### 2.1. BRON DATA

De input gebruikt in het SAS-script is het bronbestand **Niet Rechthebbende objecten.xls**. Dit bestand bevat een signaallijst met verboden adressen voor het aanvragen van toeslagen. Voor het bestand zie de map:

```
/prj/bi/data/AD010/data/RisicoSelectie/Brondata/UITGESLOTEN_ADRES/Source
```

### 2.2. VELDEN IN HET MODEL

Onderstaande velden zijn in de signaallijst opgenomen.

Name	Type	Length	Format	Informat
<b>KeyAdres</b>	Character	27		
<b>Reden</b>	Character	99	\$CHAR99.	\$CHAR99.

### 2.2. FLOWS

In dit script wordt het bronbestand ingelezen en worden enkele bewerkingen op de data gedaan:

**Stap 1:** Het bronbestand wordt ingelezen.

---

<sup>1</sup> Na overgang tot AWS zal het groen gemarkeerde (locatie en naam brondata/halfproduct/SAS-script) moeten worden aangepast.

**Stap 2:** Een nieuwe tabel wordt aangemaakt met de veld 'Reden' en het nieuw gecreëerde veld 'KeyAdres'. Hier worden meerdere velden uit het ingelezen bronbestand als volgt samengevoegd: Postcode || '&' || Huisnummer || '&' || Huisletter || Toevoeging.

**Stap 3:** De tabel wordt ontdebeld en gesorteerd op 'KeyAdres'. Hierbij mag het adres niet leeg mag zijn, i.e. de waarde van 'KeyAdres' mag niet gelijk zijn aan de waarde '&&'.

**Stap 4:** Dubbele sleutels (KeyAdres) worden verwijderd met behulp van de optie NODUPKEY.



**6PP**  
Postcodes

6PP1

Versie nummer	Datum	Initialen	Belangrijkste wijziging
<b>0.01</b>	10-07-2013		Initieel document
<b>1.0</b>	13-10-2014	10.2.e	Aanpassing n.a.v. wijziging in levering

# 1. INHOUDSOPGAVE

1. Inhoudsopgave.....	3
2. Beschrijving.....	4
2.1. Bron data.....	4
2.2. Velden.....	4
2.3. Indicatoren.....	5
3. Beslissingen.....	5

## 2. BESCHRIJVING

Voor het bepalen van de afstand tussen woonplaats en plaats van de kinderopvanginstelling hebben we de lengte- en breedtegraden nodig van de postcodes.

Deze zijn te koop bij d-centralize ([www.d-centralize.nl](http://www.d-centralize.nl)). we hebben een jaarabonnement op deze gegevens dat ingaat op 10 oktober 2014. Dit abonnement staat op naam van

10.2.e

De kosten van het abonnement zijn €75, per jaar (excl. BTW).

De inloggegevens zijn:

Het CMS: <http://cms.pro6pp.nl/cms>

E-mailadres: 10.2.e@belastingdienst.nl

Wachtwoord: 10.2.e

In het CMS vindt u:

- \* een overzicht van uw abonnementen en facturen.
- \* gebruiksstatistieken.
- \* downloaden van postcode lijsten.

Deze site is overigens niet via de huidige browser te benaderen.

Door in te loggen kunnen verschillende postcodelijsten gedownload worden. De lijst die voor ons van belang is, is: download\_nl\_sixpp.zip.

De data wordt opgeslagen op de netwerkschijf:

*Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling 2013\Brondata\6PP\Source\6pp.csv*

Vervolgens wordt de bron data ingelezen in SAS.

### 2.1. BRON DATA

De bron data bestaat uit 12 csv-bestanden (inclusief aantal records):

Bronbestand	Omschrijving	Aantal records
<b>6pp.csv</b>	Lijst met postcodes van Nederland met lengte- en breedtegraad	467.615 (1-10-2014)

### 2.1. VELDEN

Onderstaande velden zijn opgenomen:

PROVINCE

MUNICIPALITY

6PP4

WOONPLAATS

AREACODE

'6PP'N

STRAAT

STREETNUMBERS

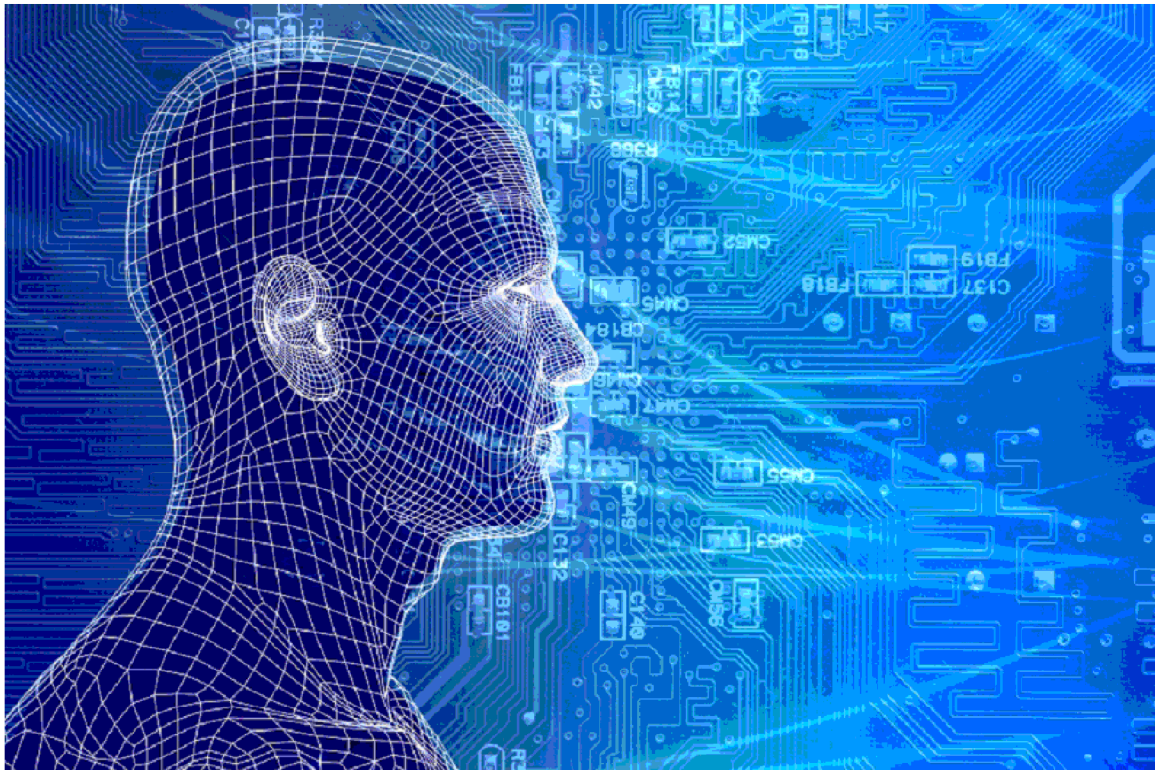
LAT

LNG

## 2.2. INDICATOREN

Veld	Benodigd voor Indicator	Omschrijving	Ontbrekende waarde indicator mits gekoppeld aan beschikking
N.v.t.			

## 3. BESLISSINGEN



# Terugkoppeling

## Vorige acties

#### Versies van het model

Versie nummer	Datum	Initialen	Belangrijkste wijziging
<b>0.00</b>	09-10-2013	<div>10.2.e</div>	Initieel document
<b>1.0</b>	21-12-2015		Aanvullen informatie

# 1. INHOUDSOPGAVE

1.	Inhoudsopgave.....	3
2.	Beschrijving proces.....	4
2.1.	Brondata.....	4
2.2.	Velden in het model.....	4
2.3.	Flows.....	5



## 2. BESCHRIJVING PROCES<sup>1</sup>

Het SAS-script **Maak lijst vorige-acties.sas** wordt één dag voor de risicoselectierun gedraaid door het risicoselectieteam. Dit resulteert in de halfproducten **vorige\_acties\_huur.sas7bdat** en **vorige\_acties\_kot.sas7bdat**, die wordt gebruikt tijdens de risicoselectierun HT en KOT respectievelijk.

Zowel het SAS-script als de halfproducten staan opgeslagen **op de netwerkschijf**. Let op! Het SAS-script staat opgeslagen in de map 'control\_klantbehandeling' en de halfproducten staan opgeslagen op twee locaties:

SAS-script:

*Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling 2013\Halfproducten\control\_klantbehandeling*

Halfproducten:

*Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling 2013\Halfproducten\control\_klantbehandeling*

*Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling 2013\Halfproducten\Terugkoppeling\Results*

### 2.1. BRONDATA

De input gebruikt in het SAS-script is het bestand **rsr\_teruglevering.sas7bdat**. Dit bestand bevat alle acties voortgekomen uit eerdere risicoselectieruns. Het bestand bevat onder andere de BSN van de toeslagaanvrager, de periode, hoofdbatch, status en resultaat van de aanvraag.

Het inputbestand wordt maandelijks, uiterlijk één dag vóór de risicoselectierun opgeleverd door

10.2.e
--------

 van het team Bedrijfsvoering. Dit gebeurt **per mail**. Vervolgens wordt het SAS-bestand door het risicoselectieteam opgeslagen op de netwerkschijf. De locatie is **dezelfde map** als de locatie van het script en de halfproducten:

*Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling 2013\Halfproducten\control\_klantbehandeling*

### 2.2. VELDEN IN HET MODEL

De SAS-bestanden **vorige\_acties\_huur** en **vorige\_acties\_kot** bestaan uit dezelfde velden:

<sup>1</sup> Na overgang tot AWS zal het groen gemarkeerde (locatie en naam brondata/halfproduct/SAS-script) moeten worden aangepast.

Name	Type	Length	Format	Informat
<b>BSN</b>	Numeric	8	BEST12.	BEST12.
<b>Actie_&lt;datum&gt;</b>	Character	50	\$CHAR23.	

Actie\_<datum> representeert hier meerdere velden.

## 2.3. FLOWS

In het script worden de vorige acties uit het inputbestand uitgesplitst naar KOT en HT. Daarnaast wordt het inputbestand getransponeerd zodat elke kolom een specifieke runperiode betreft.

**Stap 1:** Aan *rsr\_teruglevering* wordt een veld toegevoegd. Dit bevat voor elke aanvraag de waarde *Actie\_periode*. Periode is hier de runperiode. Daarnaast, indien het veld Batch leeg is en de runperiode na 201308 is, wordt de waarde van het veld Batch gelijk aan 'werkbundel'.

**Stap 2:** De acties behorende bij Huur worden geselecteerd.

**Stap 3:** Velden worden geselecteerd en het bestand met alle acties HT wordt getransponeerd. Dit resulteert in een nieuw bestand waarin alle waarden *Actie\_periode* uit het veld Run de kolommen betreffen. Daar wordt het kolom BSN aan toegevoegd. Zo bestaat elke regel uit een BSN waarvoor eerdere actie(s) is/zijn uitgevoerd. Het veld *Actie\_periode* is gevuld indien de desbetreffende actie in de runperiode is uitgevoerd. De waarde is dan gelijk aan de batchwaarde uit *rsr\_teruglevering*.

**Stap 4:** Stap 2 en 3 worden hetzelfde uitgevoerd voor KOT.

**Stap 5:** De twee resulterende tabellen worden opgeslagen op twee locaties: in de map 'control\_teruglevering' en in 'terugkoppeling'. De bestanden opgeslagen in 'terugkoppeling' worden gebruikt tijdens de risicoselectierun.

## DEELJAAR BESCHIKKING HT

Versie nummer	Datum	Initialen	Belangrijkste wijziging
<b>0.01</b>	17-07-2013	10.2.e	Initieel document
<b>0.02</b>	24-02-2014		Het programma voor het genereren van deeljaar beschikking heeft een andere naam
<b>1.0</b>	30-11-2015		Informatie aanvullen

## 1. INHOUDSOPGAVE

1.	Inhoudsopgave.....	3
2.	Beschrijving.....	4
2.1.	Bron data.....	4
2.2.	Velden.....	4
2.3.	Indicatoren.....	5
3.	Beslissingen.....	5

## 2. BESCHRIJVING<sup>1</sup>

Deze brondata bevat een lijst met deeljaar beschikkingen voor huurtoeslag en wordt gebruikt in het halfproduct DTM\_HT.

Het risicoselectieteam vraagt de brondata maandelijks **per mail op bij** 10.2.e Na formeel beschikken wordt de data **door BCA/BICC** opgeleverd als SAS-bestand en **geplaatst op de netwerkschijf**:

*Q:\VEPROW52\TSL\_UT\_Datawisseling\Noodoplossingen voor EDW 2014.1\Profiling Huur of Profiling Kot*

Dit bestand heet over het algemeen 'dljr\_beschikking\_huur\_<datum>' of 'deeljaar\_beschikking\_huur\_<datum>'.

Vervolgens wordt de data door het risicoselectieteam **gekopieerd naar een andere locatie**:

*Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling 2013\Brondata\ DEELJAAR BESCHIKKING HT \Source*

Op deze locatie is de naam van het bronbestand aangepast tot **'deeljaar\_beschikking\_huur'**. Dit bestand wordt gebruikt bij het creëren van het halfproduct.

### 2.1. BRONDATA

De brondata van DEELJAAR BESCHIKKING HT (inclusief aantal records):

Bronbestand	Omschrijving	Aantal records
<b>deeljaar_beschikking_huur.sas7bdat</b>	Lijst met deeljaar beschikkingen huur	2.098.623 (24-02-2014)

### 2.1. VELDEN

Onderstaande velden zijn opgenomen:

Name	Type	Length	Format
<b>Bsn</b>	Numeric	8	
<b>Toeslagjaar</b>	Numeric	8	20.
<b>Beschikkingnummer</b>	Character	17	\$17.
<b>Ingangsdatum</b>	Date	8	YYMMDDD10.

<sup>1</sup> Na overgang tot AWS zal het groen gemarkeerde (aanvraag / oplevering / locatie en naam brondata) moeten worden aangepast. Daarnaast zal te zijner tijd de query, gebruikt om de data op te halen uit de database, toegevoegd worden.

<b>Einddatum</b>	Date	8	YYMMDDD10.
<b>Maandbedrag_toeslag</b>	Numeric	8	14.2
<b>Postcode</b>	Character	6	\$6.
<b>Huisnummer</b>	Character	6	\$6.
<b>Huisnummer_toevoeging</b>	Character	12	\$12.
<b>Huisletter</b>	Character	5	14.2
<b>Nadere_aanduiding</b>	Character	5	14.2
<b>Rekenhuur_per_maand</b>	Numeric	8	14.2
<b>Rekenhuur_regulier</b>	Numeric	8	14.2
<b>Kale_Huur</b>	Numeric	8	14.2
<b>Standplaatshuur</b>	Numeric	8	14.2
<b>DienstEnRecruitmenkosten</b>	Numeric	8	14.2
<b>Kosten_energie_coll_ruimten</b>	Numeric	8	14.2
<b>Kosten_schoonmaak_coll_ruimten</b>	Numeric	8	14.2
<b>Kosten_huismeester</b>	Numeric	8	14.2
<b>Draagkracht</b>	Numeric	8	20.
<b>Huishoudgrootte</b>	Numeric	8	
<b>Aantal_kinderen_ouder_18</b>	Numeric	8	
<b>Maximale_huurgrens</b>	Numeric	8	14.2

## 2.1. INDICATOREN

Veld	Benodigd voor Indicator	Omschrijving	Ontbrekende waarde indicator mits gekoppeld aan beschikking
N.v.t.			

## 3. BESLISSINGEN

## **Eerdere Acties**



## Versies van het model

Versie nummer	Datum	Initialen	Belangrijkste wijziging
1.0	30-11-2015	10.2.e	Initieel document
1.1	19-01-2016		Zaak-bestand toegevoegd
1.2	25-04-2016		AWS-aanpassingen
1.3	04-08-2016		OverzichtResultaat geleverd op de AWS

## 1. INHOUDSOPGAVE

1.	Inhoudsopgave.....	3
2.	Beschrijving.....	4
2.1.	Brondata.....	4
2.2.	Velden.....	4
2.3.	Indicatoren.....	5
3.	Beslissingen.....	5

## 2. BESCHRIJVING

Deze brondata bevat een lijst van eerder behandelde aanvragen Huur, KOT en Zorg met reden hoog risico, harde stopper, HOT-HOR post, onbekende aanvrager of aanvrager buitenlands adres. De brondata wordt gebruikt in het halfproduct Eerdere Acties.

De brondata wordt door het risicoselectieteam **per mail** opgevraagd bij **Queryloket Toeslagen**.

In principe gaat de aanvraag twee keer per maand uit, één keer voor de proefrun en één keer voordat de echte risicoselectie wordt gedraaid. Van december tot en met mei gebeurt dit vier keer wegens het draaien van de runs voor twee verschillende toeslagjaren.

De brondata wordt opgeleverd als xlsx-bestand en opgeslagen **op het netwerk**:

*Q:\VEPROW52\TSL\_UT\_Datawisseling\intelligence\Profiling 2015\jjjmmdd*

en **op de AWS**:

*/prj/bi/data/AD010/data/RisicoSelectie/Brondata/EERDERE ACTIES/Source*

Het bestand op deze locatie wordt ingelezen in SAS.

Ook het bestand 'Zaak' is nodig. Dit wordt elke maandag en/of donderdag verversd door 10.2.e

10.2.e en geplaatst op de AWS:

*/prj/bi/data/AD010/data/RisicoSelectie/Brondata/IGS/Source*

### 2.1. BRONDATA

De brondata van Eerdere Acties:

Bronbestand	Omschrijving
<b>OverzichtResultaat.xlsx</b>	Lijst met eerder behandelde aanvragen
<b>Zaak.sas7bdat</b>	De resultaten van eerdere behandelingen

### 2.1. VELDEN

Onderstaande velden zijn opgenomen:

Naam	Getalnotatie (csv)
<b>Bsn</b>	Standaard
<b>Bundelnaam</b>	Standaard
<b>Reden</b>	Standaard
<b>Regeling</b>	Standaard
<b>Toeslagjaar</b>	Standaard
<b>UitgeslotenVanaf</b>	Datum
<b>UitgeslotenTm</b>	Datum
<b>Mailing</b>	Standaard

<b>ResultaatVerplicht</b>	Standaard
<b>Behandelwijze</b>	Standaard
<b>Resultaat</b>	Standaard
<b>NieuwToeslagbedrag</b>	Standaard
<b>UserId</b>	Standaard
<b>Wijzigingsdatum</b>	Datum
<b>Voortgang</b>	Standaard

## 2.1. INDICATOREN

Veld	Benodigd voor Indicator	Omschrijving	Ontbrekende waarde indicator mits gekoppeld aan beschikking
N.v.t.			

## 3. BESLISSINGEN

**Ouder Kind**

Ouder\_Kind

30-11-20151

Versie nummer	Datum	Initialen	Belangrijkste wijziging
<b>0.01</b>	07-10-2013		Initieel document
<b>1.0</b>	30-11-2015	10.2.e	Aanvullen informatie

# 1. INHOUDSOPGAVE

1. Inhoudsopgave.....	3
2. Beschrijving.....	4
2.1. Brondata.....	4
2.2. Velden.....	4
2.3. Indicatoren.....	4
3. Beslissingen.....	4

## 2. BESCHRIJVING

Deze bron data bevat een lijst met ouder-kind relaties en wordt gebruikt in halfproduct OUDER\_KIND.

Maandelijks vraagt het risicoselectieteam de brondata **per mail op bij BCA / UPG**. Deze aanvraag verloopt **via** . Vóór het draaien van de risicoselectierun wordt de brondata opgeleverd als csv-bestand en **op de netwerkschijf opgeslagen**:

*Q:\VEPROW52\TSL\_UT\_Datawisseling\Queries Formeel Beschikken*

Vervolgens wordt het bestand door het risicoselectieteam gekopieerd naar **de volgende locatie**:

*Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling 2013\Brondata\ OUDER\_KIND\Source*

Vanaf deze locatie wordt het bestand ingelezen in SAS.

### 2.1. BRONDATA

De brondata van Ouder Kind:

Bronbestand	Omschrijving
<b>Klantbeeld ouderkind.csv</b>	Lijst met ouder kind relaties

### 2.1. VELDEN

Onderstaande velden zijn opgenomen:

BSN kind  
BSN ouder1  
BSN ouder2

### 2.2. INDICATOREN

Veld	Benodigd voor Indicator	Omschrijving	Ontbrekende waarde indicator mits gekoppeld aan beschikking
N.v.t.			

## 3. BESLISSINGEN

Er komen een flink aantal kinderen voor met meer dan 2 ouders, dit is nog niet opgelost.



## Onbekende aanvragers

Versie nummer	Datum	Initialen	Belangrijkste wijziging
<b>0.01</b>	30-06-2014	10.2.e	Initieel document
<b>0.02</b>	08-07-2014		Review
<b>1.0</b>	18-12-2015		Informatie aanvullen
<b>1.1</b>	04-08-2016		AWS-aanpassingen

# 1. INHOUDSOPGAVE

1. Inhoudsopgave.....	3
2. Beschrijving.....	4
2.1. Brondata.....	4
2.2. Velden.....	4
2.3. Indicatoren.....	4
3. Beslissingen.....	5

## 2. BESCHRIJVING

Deze brondata bevat een lijst met BSNs van onbekende aanvragers, d.w.z. aanvragers die een BSN hebben dat jonger is dan 12 maanden en die nog nooit een formele beschikking > 0 hebben gehad. De brondata wordt gebruikt in het halfproduct 'Uitsluitlijst onbekende aanvragers Zorg'.

Het risicoselectieteam vraagt de data maandelijks per mail op bij **Toeslagen IM**

**Portfoliomanagement\_Postbus**. Deze aanvraag stuurt IM middels een queryverzoek door naar BICC.

Vóór de risicoselectierun levert **BCA\BICC** de brondata op als SAS-bestand en wordt de data opgeslagen op het netwerk:

*Q:\VEPROW52\TSL\_UT\_Datawisseling\Profiling van CIV\jjjjmdd*

en op de AWS:

*/prj/bi/data/AD010/data/RisicoSelectie/CIV/*

Vanaf deze locatie wordt de data ingelezen in SAS om het halfproduct te creëren.

### 2.1. BRONDATA

De brondata van Onbekende aanvragers:

Bronbestand	Omschrijving
<b>civ_prof_nieuwe_aanvragers.sas7bdat</b>	BSNs van nieuwe aanvragers

### 2.1. VELDEN

Onderstaande velden zijn opgenomen:

Name	Type	Length	Format
<b>Aanvrager_BSN</b>	Numeric	8	11.
<b>Toeslagjaar</b>	Numeric	8	20.
<b>Huur</b>	Numeric	8	14.2
<b>Zorg</b>	Numeric	8	14.2
<b>Kot</b>	Numeric	8	14.2
<b>Kgb</b>	Numeric	8	14.2

### 2.2. INDICATOREN

Veld	Benodigd voor Indicator	Omschrijving	Ontbrekende waarde indicator mits gekoppeld aan beschikking
N.v.t.			

### 3. BESLISSINGEN

n.v.t.

## Nihilstellers

Versie nummer	Datum	Initialen	Belangrijkste wijziging
<b>0.01</b>	23-12-2013	10.2.e	Initieel document

## 1. INHOUDSOPGAVE

1. Inhoudsopgave.....	3
2. Beschrijving.....	4
2.1. Bron data.....	4
2.2. Velden.....	4
2.3. Indicatoren.....	4
3. Beslissingen.....	4



## 2. BESCHRIJVING

Deze bron data wordt gebruikt in het halfproduct Nihilstellers.

De bron data is opgevraagd bij het Queryloket 10.2.e. De huidige levering betreft de jaren 2007 t/m 2011, deze zijn allen afgesloten. Het bestand is daarom gefixeerd en wordt niet periodiek opgevraagd. Bij afsluiting van het jaar 2012 dient om een update gevraagd te worden.

De bron data worden opgeleverd als .xlsx-bestand en wordt opgeslagen op het netwerk: Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling 2013\Brondata\Nihilstellers\Source. Om het inlezen in SAS te vergemakkelijken wordt het bestand handmatig opgeslagen als .csv.

### 2.1. BRON DATA

De bron data van Nihilstellers:

Bronbestand	Aantal records
IC 8880 nihilstellers_2007_2011.xlsx	87.856
Nihilstellers.csv	87.856

### 2.1. VELDEN

Onderstaande velden zijn opgenomen:

BSN  
Jaar

### 2.2. INDICATOREN

Veld	Benodigd voor Indicator	Omschrijving	Ontbrekende waarde indicator mits gekoppeld aan beschikking
BSN	Nihilstelling_KOT	BSN is in het verleden nihil gesteld (0=nee; 1=ja)	0
BSN	Aant_nihilstelling_KOT	Aantal keer dat BSN nihil gesteld is	0

### 3. BESLISSINGEN

## **LRK**

Landelijk Register Kinderopvang

Versie nummer	Datum	Initialen	Belangrijkste wijziging
<b>1.0</b>	18-12-2015	10.2.e	Initieel document
<b>1.1</b>	25-04-2016		AWS-aanpassingen

## 1. INHOUDSOPGAVE

1.	Inhoudsopgave.....	3
2.	Beschrijving.....	4
2.1.	Brondata.....	4
2.2.	Velden.....	4
2.3.	Indicatoren.....	5
3.	beslissingen.....	5

## 2. BESCHRIJVING

Deze brondata bevat een lijst met kinderopvanginstellingen uit het Landelijk Register Kinderopvang ([www.landelijkregisterkinderopvang.nl](http://www.landelijkregisterkinderopvang.nl)). Deze lijst wordt gebruikt in het halfproduct LRK.

De brondata wordt via DUO (Dienst Uitvoering Onderwijs) aan team BCA\BICC geleverd, die het bestand eens per twee weken updatet en als txt-bestand op het netwerk plaatst:

*Q:\VEPROW52\TSL\_UT\_Datawisseling\LRK en WKO\Export LRK\LRK<datum>*

Vóór het draaien van de risicoselectie-run wordt de meest actuele bron data door het risicoselectieteam gekopieerd naar de volgende locatie op de AWS:

*/prj/bi/data/AD010/data/RisicoSelectie/Brondata/LRK/Source*

Het meest recente bestand wordt gebruikt om het halfproduct LRK te vernieuwen voordat de risicoselectie wordt gedraaid.

### 2.1. BRONDATA

De brondata van LRK:

Bronbestand	Omschrijving
<b>Export-bd-new.txt</b>	Lijst met kinderopvanginstellingen

### 2.1. VELDEN

Onderstaande velden zijn opgenomen:

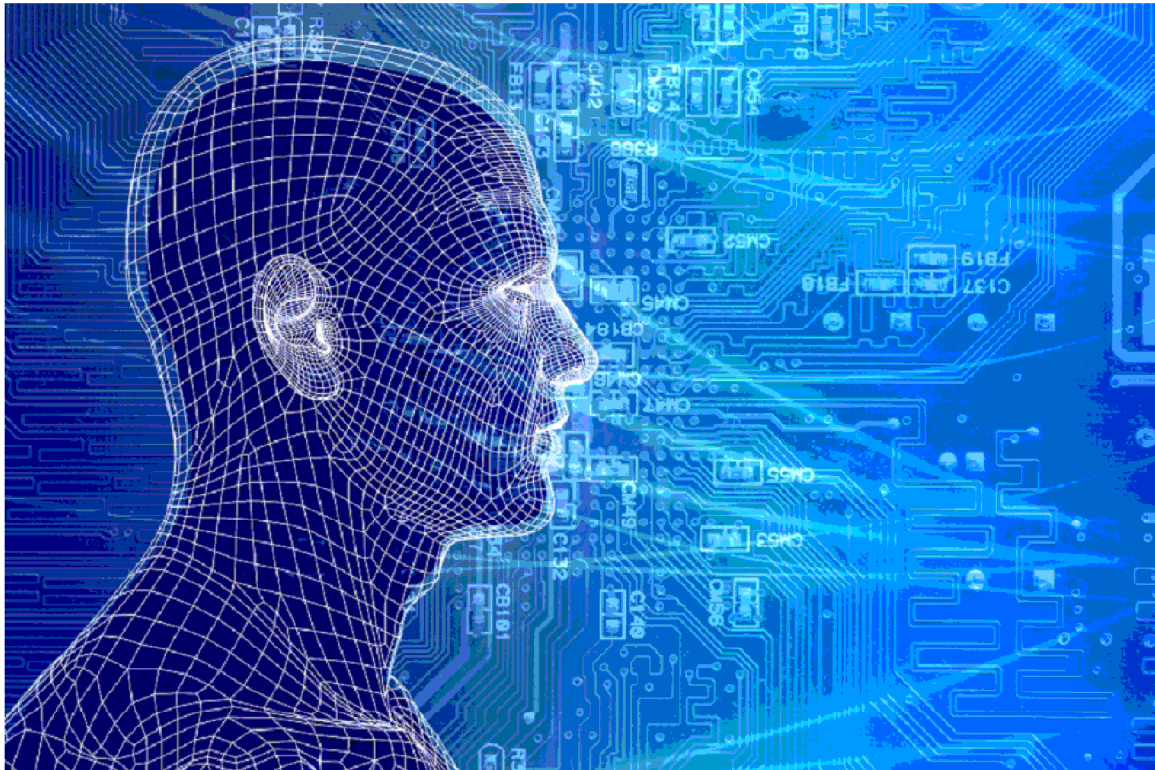
Name	Length
<b>LRK_ID</b>	8
<b>Soort_opvang</b>	3
<b>Datum_1</b>	8
<b>Naam</b>	80
<b>BSN</b>	8
<b>KvK</b>	8
<b>Huisnr</b>	8
<b>Toevoeging</b>	5
<b>Postcode</b>	6
<b>Plaats</b>	80
<b>Straat</b>	80
<b>Datum2</b>	8
<b>Status</b>	25

<b>Datum3</b>	8
<b>Datum4</b>	8

## 2.2. INDICATOREN

Veld	Benodigd voor Indicator	Omschrijving	Ontbrekende waarde indicator mits gekoppeld aan beschikking
N.v.t.			

## 2. BESLISSINGEN



## KOT\_Tijdreizen



## Versies van de documentatie

Versie nummer	Datum	Initialen	Belangrijkste wijziging
<b>0.01</b>	28-11-2013	10.2.e	Initieel document
<b>0.02</b>	10-02-2014		Nieuwe maand toegevoegd: 2014-01
<b>1.0</b>	20-12-2015		Wijziging nav overzetten naar AWS en bevroren toeslagjaar 2013.
<b>1.1</b>	10-02-2016		Stappenplan toegevoegd
<b>1.2</b>	30-11-2018		Nieuwe opzet, 2013 apart

1. INHOUDSOPGAVE	
1. Inhoudsopgave.....	3
2. Doel.....	4
3. Beschrijving project.....	5
3.1. Bron data.....	5
3.2. Flows.....	5
3.3. Beschrijving Halfproduct.....	6
3.3.1. Autoexec.....	6
3.3.2. A. Initialize.....	6
3.3.3. B. Steekproef cases 2013.....	6
3.3.4. C. Create Testcases mm-YY.....	6
3.3.5. Z. Append cases.....	7
4. Update KOT_Tijdreizen.....	7
5. Stappenplan.....	8
5.1.1. Programma tijdreizen:.....	8
5.1.2. Input data toevoegen:.....	9
6. Nieuwe opzet.....	9
7. Resultaten per tijdreizen.....	10
7.1. November 2018.....	10

## 2. DOEL

Doel van tijdreizen is testcases voorzien van de informatie zoals deze op het moment van beoordeling (als goed of fout) actueel was.  
Dit heeft tevens tot gevolg dat als er nieuwe indicatoren worden toegevoegd aan het model, deze ook met terugwerkende kracht voor alle trainingscases toegevoegd moeten kunnen worden.

## 3. BESCHRIJVING PROJECT

### 3.1. BRON DATA

Er is generieke brondata en brondata per risicoselectierun:

- Generiek is het SAS-bestand 'Trainingscases\_KOT' dat volgt uit het SAS project 'Trainingscases' (Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling 2013\Trainingscases).
- Daarnaast is er per risicoselectierun die in het verleden is uitgevoerd een tabel met alle indicatoren vereist. Dit is de tabel KOT\_Score. Let op: de tabel Export\_Temp3 kan ook, die bevat dezelfde indicatoren, en daar boven op een aantal andere variabelen die uit de risicoselectierun volgen, maar niet gebruikt worden tijdens het tijdreizen.
- Voor KOT is zijn na 2015 2 nieuwe indicatoren toegevoegd, namelijk of iemand bijstand heeft, en zo ja, het bedrag.

De bijstandsgegevens zijn afkomstig uit het Halfproduct FLG, het resultaat hiervan wordt gekopieerd naar ..... dit bestand wordt per jaar opnieuw gedraaid.

### 3.2. FLOWS

Het SAS project wordt met ingang van december 2015 uitgevoerd op de AWS omgeving: AD010\data\RisicoSelectie\Tijdreizen\KOT\SAS\KOT\_Tijdreizen.

Het project bestaat uit meerdere process flows:

**Autoexec:** toewijzen van libnames

**A. Initialize:** Initialisatie van scripts en input Trainingscases\_KOT

**B. Steekproef cases 2013:** trekken van steekproef uit bestand 2013

**C.-H. Create Testcases 2014-2019:** het toevoegen van details vanuit een historische risicoselectierun die uitgevoerd is vlak voor de datum waarop de case als trainingscase beoordeeld is.

**Z. Append:** voeg alle testcases samen in 1 tabel

In december 2015 is besloten om de trainingscases voor toeslagjaar 2013 te bevriezen. De process flows die dit regelden waren omvangrijk en er werden nauwelijks meer nieuwe trainingscases voor dit toeslagjaar toegevoegd.

De bevvriezing houdt in dat het tijdreizen voor deze toeslagjaren eenmalig is gerund (zie aparte documentatie KOT\_Tijdreizen\_2013 voor een toelichting op deze process flows), waarna het bestand met cases inclusief alle indicatoren opgeslagen is. Dit bestand wordt aan de gehele set met trainingscases toegevoegd in de process flow Z. Append. De implicatie is dat geen nieuwe trainingscases voor 2013 meer toegevoegd kunnen worden. Omdat het een omvangrijk en niet representatief bestand is, wordt een steekproef van 2013 aan het totaal toegevoegd (zie par. 3.3.3).

Met ingang van maart 2019 zijn de cases van 2014 en 2015 bevroren, ook hieraan kunnen geen nieuwe cases toegevoegd worden. (wat te doen met fraudecases?)



### 3.3. BESCHRIJVING HALFPRODUCT

#### 3.3.1. *AUTOEXEC*

Deze process flow wordt gebruikt voor het aanmaken om:

- Verschillende libraries aan te maken, 1 generieke, en 1 per maand;
- De generieke brondata (BSN's van alle testcases in te lezen).

#### 3.3.2. *A. INITIALIZE*

In het script A02\_SetMonth vindt de bepaling van de meest historische data per testcase plaats. Hiervoor wordt eerst per testcase bepaald wat de datum van beoordeling was, op basis van de diverse beschikbare datumkolommen.

Vervolgens worden aan de testcases alle beschikkingen uit de risicoselectieruns gelijk of voor de beoordeeldatum gekoppeld. Principe hierachter is dat de beoordeling van de BSN waarschijnlijk plaatsvindt naar aanleiding van een conceptbeschikking, en dat die conceptbeschikking terug te vinden zou moeten zijn in de run die op of voor die datum plaats heeft gevonden. Als de BSN in meerdere risicoselectieruns terugkomt, wordt alleen de laatste mutatie behouden.

Dus bijvoorbeeld:

- BSN is beoordeeld op 2 augustus 2016.
- Er zijn gescoorde mutaties beschikbaar voor 31 mei, 2 augustus en 30 september 2016.
- De testcase wordt toegewezen aan maand augustus 2016.

En een tweede voorbeeld:

- BSN is beoordeeld op 20 december 2016.
- Er zijn gescoorde mutaties beschikbaar voor 2 augustus, 30 september en 13 december 2016.
- De testcase wordt toegewezen aan maand december 2016.

#### 3.3.3. *B. STEEKPROEF CASES 2013*

In deze process flow wordt het steekproefbestand voor 2013 aangemaakt. Het programma is in deze process flow opgenomen zodat het mogelijk is om een andere steekproef te trekken indien dat wenselijk is. Het resultaat van de steekproef staat in de map Results/KOT\_training\_2013\_selectie.

#### 3.3.4. *C. CREATE TESTCASES 2014*

Per jaar is er momenteel 1 process flow.

Brondata specifiek voor deze flow:

KOT_score_yyyymm	Output van het model zoals gedraaid tijdens een risicoselectierun (bijvoorbeeld januari 2014 over toeslagjaar 2014).
------------------	--

Deze flows kennen 2 stappen:

1. selecteer in tabel met alle trainingscases de cases die beoordeeld zijn op het moment van de betreffende risicoselectierun, en gebruik hiervoor alle indicatoren beschikbaar in de KOT\_score tabel voor die risicoselectierun.
2. Voeg toe of aanvragers een bijstandsuitkering hebben en koppel dat aan het bestand. Deze indicator is later toegevoegd, vandaar dat het toevoegen van deze gegevens in een apart programma gaat.

### 3.3.5. Z. APPEND CASES

**Z01:** In deze flow worden alle testcases zoals samengesteld in de voorgaande flows samengevoegd tot één tabel: KOT\_TRAININGSCASES\_yyyymmdd. Alleen velden die nodig zijn om te modelleren in Enterprise Miner worden meegenomen, de rest wordt gedropt.

Bovendien wordt een laatste filter gezet op de cases die uiteindelijk als trainingscase naar Enterprise Miner gaan.

Het bestand met trainingscases uit 2013 (zie hierboven) wordt vervolgens afzonderlijk toegevoegd. Vervolgens wordt een steekproef getrokken op het hele bestand omdat het bestand onevenwichtig is opgebouwd voor wat betreft herkomst (voor een beschrijving van 'herkomst' zie de documentatie van Trainingscases). We hebben geëxperimenteerd met de SAS Enterprise Miner en op basis daarvan gekozen voor een steekproef van:

- 4.000 posten op de DTCheck van 2017 (dit betreft alleen goede posten)
- 1.000 posten uit GreenLane 2015 (dit betreft alleen goede posten)
- 1.000 posten uit GreenLane 2016 (dit betreft alleen goede posten)
- 4.000 goede posten uit het bestand van eerder gecontroleerde posten
- 4.000 foute posten uit het bestand van eerder gecontroleerde posten

Het programma is zodanig opgezet dat de steekproef eenvoudig aangepast kan worden.

**Z02:** check op alle numerieke velden of missende waarden voorkomen. Deze output wordt verder niet gebruikt, is bedoeld voor controledoeleinden.

## 4. UPDATE KOT\_TIJDREIZEN

Het project is per risicoselectierun opgezet. Dat betekent dat:

1. Er diverse scripts zijn waarin hard-coded de betreffende maanden aangeroepen worden;
2. Er 1 process flow per risicoselectierun is, waarin de relevante testcases geselecteerd worden en de indicatoren vanuit die risicoselectierun toegevoegd worden aan de cases die vlak na die risicoselectierun zijn beoordeeld.

Bij het toevoegen van een nieuwe risicoselectierun moet dus:

1. Nieuwe (bron)data die relevant is voor deze risicoselectierun ontsloten worden. De KOT\_Score tabel uit de betreffende risicoselectierun wordt tijdens het maken van de back-up van de desbetreffende risicoselectierun in de KOT\_Tijdreizen folder op de AWS omgeving gezet.
2. Een tab ingevoegd worden waarin testcases voor de betreffende risicoselectierun geselecteerd en indien nodig aangevuld worden.
3. Diverse algemene scripts aangepast worden zodat ook de nieuwe risicoselectierun meeloopt. Het gaat om:  
Autoexec: A01: nieuwe library  
A02: nieuwe risicoselectierun hardcoded toevoegen aan stap 3 (2x).  
Z01: finaal testcases bestand voor de nieuwe risicoselectierun toevoegen aan set statement in stap 1.
4. Bepaald worden welke indicatoren nog niet beschikbaar zijn, en (indien van toepassing) hoe deze op basis van de tijdens de run voor die risicoselectierun gebruikte gegevens toegevoegd kunnen worden. Dit komt momenteel niet voor.

## 5. STAPPENPLAN

In dit deel staan kort de verschillende stappen die ondernomen moeten worden bij het aanvullen van het tijdreizen. In het stappenplan moet de input data worden aangevuld, en het programma moet worden aangepast.

### 5.1.1. PROGRAMMA TIJDREIZEN:

[AWS/files/AD10/data/RisicoSelectie/tijdreizen/kot\_training/SAS/KOT\_tijdreizen.epg]

- 1.) In process flow 'A0.Initialize' in programma 'A01\_SetParameters' onderaan een libname toevoegen met de nieuwe run. Voorbeeld: LIBNAME LIB0316 "&PATH./Input/2016";
- 2.) Zelfde process flow in programma 'A02\_Set month' de libnames toevoegen en onderaan de nieuwe risicoselectie run met betreffende datum en het jaar waarop de run betrekking heeft toevoegen. Voorbeeld: LIB0316.KOT\_score\_201603 (IN = IN032016 KEEP = BSN)

```
En: IF IN032016 THEN DO;
      runnr = 201603;
      Toeslagjaar = 2016;
      Rundatum = INPUT('29/01/2016', ddmmyy10.);
END;
```

- 3.) Vervolgens wordt er in de process flow van het huidige jaar een nieuw stukje aan het programma geplakt waarin nieuwe runs worden toegevoegd. Voorbeeld van een programma:

```
PROC SQL;
CREATE TABLE LIB0316.KOT_training_201603_Final AS
SELECT A.*,
       B.Type,
       B.Datum_beeoordeeld,
       B.Rundatum,
       B.Herkomst
FROM LIB0316.KOT_score_201603 AS A
INNER JOIN KOTTrain.Trainingscases_KOT_FINAL AS B
WHERE B.runnr = 201603;
QUIT;
```

- 4.) In de Process Flow 'ZZ.Append' wordt tot slot in het programma 'Z01\_append' de libname ook toegevoegd, voorbeeld: LIB0316.KOT\_TRAINING\_201603\_FINAL

#### 5.1.2. INPUT DATA TOEVOEGEN:

Vervolgens moet de data van elke risicoselectie run worden toegevoegd in: AWS/files/AD10/data/RisicoSelectie/tijdreizen/kot\_training/Input/"betreffende jaar"/kot\_score\_'jaar'runnr'.

Met ingang van 2017 gebeurt dit automatisch tijdens het maken van de back-up van de risicoselectie.

## 6. NIEUWE OPZET

Met ingang van april 2019 zijn ook de jaren 2014 en 2015 bevroren. Deze bestanden zijn opgeslagen in xxxx en worden rechtstreeks in de append cases toegevoegd.

Als het proces is afgerond het programma en de datasets opslaan op de q-schijf:  
Q:\VEPROW63\TSL\_DM\_Handhavingsregie\Profiling 2013\Trainingscases\Backup  
Het project in de map sas\_epg en de bestanden in de map: bestanden.  
Telkens opslaan met bestandsnaam\_jjjjmmdd.

## 7. RESULTATEN PER TIJDREIZEN

Vanaf maart 2019 wordt per keer dat we tijdreizen draaien een overzicht toegevoegd van het totale bestand en de steekproef naar type, herkomst en toeslagjaar.

### 7.1. MAART 2019

		Totaal			Steekproef		
		goed	fout	totaal	goed	fout	totaal
Herkomst	Toeslagjaar						
DT Check	2017	12.567	0	12.567	4.000	0	4.000
Excel	2013	1.500	1.938	3.438	1.500	1.938	3.438
	2014	622	859	1.481	622	859	1.481
	2015	15	61	76	15	61	76
	<b>totaal</b>	<b>2.137</b>	<b>2.858</b>	<b>4.995</b>	<b>2.137</b>	<b>2.858</b>	<b>4.995</b>
Fraudeteams	2013	0	562	562	0	562	562
	2014	0	657	657	0	657	657
	2015	0	262	262	0	262	262
	2016	0	263	263	0	263	263
	2017	0	167	167	0	167	167
	2018	0	99	99	0	99	99
	2019	0	4	4	0	4	4
	<b>totaal</b>	<b>0</b>	<b>2.014</b>	<b>2.014</b>	<b>0</b>	<b>2.014</b>	<b>2.014</b>
GreenLane	2015	3.382	0	3.382	1.000	0	1.000
	2016	4.740	0	4.740	1.000	0	1.000
	2017	410	30	440	410	30	440
	<b>totaal</b>	<b>8.532</b>	<b>30</b>	<b>8.562</b>	<b>2.410</b>	<b>30</b>	<b>2.440</b>
Zaak	2013	131	123	254	39	58	97
	2014	1.324	1.538	2.862	361	758	1.119
	2015	3.774	2.673	6.447	987	1.329	2.316
	2016	2.956	1.196	4.152	778	623	1.401
	2017	3.575	1.491	5.066	981	734	1.715
	2018	2.975	978	3.953	839	485	1.324
	2019	47	21	68	15	13	28
	<b>totaal</b>	<b>14.782</b>	<b>8.020</b>	<b>22.802</b>	<b>4.000</b>	<b>4.000</b>	<b>8.000</b>
<b>Totaal</b>		<b>38.018</b>	<b>12.922</b>	<b>50.940</b>	<b>12.547</b>	<b>8.902</b>	<b>21.449</b>