

Documentatie
Personen en Beroepen gekoppeld aan Bedrijven
(PBB)

Bronvermelding

Publicatie van uitkomsten geschiedt door de onderzoeksinstituting of de opdrachtgever op eigen titel. Verwijzing naar het Centraal Bureau voor de Statistiek (CBS) betreft uitsluitend het gebruik van de niet-openbare microdata. Deze microdata zijn onder [bepaalde voorwaarden](#) voor statistisch en wetenschappelijk onderzoek toegankelijk. Voor nadere informatie microdata@cbs.nl. Hier een voorbeeld voor de formulering van de bronvermelding:

“Resultaten [gedeeltelijk] gebaseerd op eigen berekeningen [*naam onderzoeksinstituting, c.q. opdrachtgever*] op basis van niet-openbare microdata van het Centraal Bureau voor de Statistiek betreffende Personen en Beroepen gekoppeld aan Bedrijven.”

Engelse versie

“Results based on calculations by [*name of research institution or commissioning party*] using non-public microdata from Statistics Netherlands.”

Under [certain conditions](#), these microdata are accessible for statistical and scientific research. For further information: microdata@cbs.nl.

Beschikbare bestand(en):

De bestanden zijn per jaar beschikbaar over de periode 2010 t/m 2022.

In de [Versiegeschiedenis](#) ziet u een overzicht van de beschikbare bestanden.



De gebruiker dient rekening te houden met het volgende:

- Zie [Onderzoeksomschrijvingen](#) voor meer informatie over alle onderzoeken.





Inhoudsopgave

Bronvermelding	2
Beschikbare bestand(en):	3
1. Inleiding	5
2. Toelichting op de inhoud van de bestanden	6
Beschrijving	6
Beschrijving van de populatie	6
Bijlagen	8
3. Bestandsopbouw en toelichting	9
Bestandsopbouw	9
Toelichting op de variabelen	10
Versiegeschiedenis.....	11

1. Inleiding

Het team Microdata Services van het CBS stelt, onder [bepaalde voorwaarden](#), niet-openbare microdata (geanonimiseerde data op persoons-, bedrijfs- en adresniveau) toegankelijk voor statistisch en wetenschappelijk onderzoek. Dit rapport beschrijft de inhoud en structuur van *“Personen en Beroepen gekoppeld aan Bedrijven”*.

Hoofdstuk 2 beschrijft in het algemeen de microdatabestanden zoals populatieafbakening, methodologische bijzonderheden, kwaliteit en herkomst van de gegevens.

In hoofdstuk 3 wordt de bestandsindeling gegeven en worden voor de categoriale variabelen alle mogelijke scores en hun betekenis opgesomd.

2. Toelichting op de inhoud van de bestanden

Beschrijving

Het uitgangspunt van dit bestand bestaat uit de polisadministratie. Deze bron bevat alle loonaangiften in Nederland, inclusief identificerende sleutels van de personen die het loon hebben ontvangen (RINPERSOONS en RINPERSOON) alsook de bedrijfseenheden die het loon heeft uitgekeerd (SBEID). In het hier besproken bestand worden deze gegevens verrijkt met informatie over het beroep (EBBTW1ISCO2008V) van deze personen, voor zover deze personen in de periode 2010-2022 zijn waargenomen in de Enquête Beroepsbevolking (EBB). Indien een persoon in één van deze jaren is waargenomen wordt de informatie over het beroep aan alle observaties van deze persoon in de polisadministratie toegekend. De variabele EBB_waarneming in het bestand geeft aan of het beroep op de desbetreffende regel is waargenomen in de EBB (1) of is geïmputeerd (0).

De gegevens zijn beschikbaar voor de periode 2010-2022. Er zijn vooralsnog geen updates van dit bestand voorzien.

Beschrijving van de populatie

Polisadministratie [(S)POLISBUS]

Het uitgangspunt van het PBB bestand zijn alle loonaangiften in de polisadministratie. De informatie over het beroep uit de EBB wordt in principe alleen uitgevraagd voor de eerste werkring. We nemen vanuit de polisadministratie daarom ook maar één RINPERSOONS-RINPERSOON-SBEID regel per jaar mee. De volgende beslisregels die hierbij worden gehanteerd cq. toegepast zijn de volgende (in deze volgorde):

- a) Uit alle loonaangiften worden uitsluitend gegevens meegenomen van bedrijfseenheden met een niet-fictieve SBEID. Bedrijfseenheden met loonaangiften die bijvoorbeeld niet in het Algemeen Bedrijven Register (ABR) worden teruggevonden worden dus verwijderd uit het bestand. Deze gegevens zouden in een later stadium hoe dan ook niet gekoppeld kunnen worden aan andere bronnen met bedrijvengegevens en laten we daarom buiten beschouwing.
- b) Na stap a) worden loonaangiften binnen dezelfde persoon-bedrijfsrelatie (op basis van de sleutels RINPERSOON-RINPERSOONS-SBEID) op jaarbasis samengeteld. Dus, alle lonen en andere aggregerbare data wordt opgeteld zodat er per RINPERSOON-RINPERSOONS-SBEID-JAAR combinatie één regel overblijft. Hierna worden alle RINPERSOON-RINPERSOONS-SBEID-JAAR combinaties met een SBASISLOON ≤ 0 verwijderd uit de data.
- c) We behouden alleen de RINPERSOON-RINPERSOONS-SBEID-JAAR combinatie behouden met het hoogste SBASISLOON.
- d) Het kan nu voorkomen dat er nog mensen in de data zitten met een identiek SBASISLOON bij meer dan één SBEID. Ter illustratie: In de 2022 data zien we dat dit zo is voor 0,03% van de personen. In stap d) wordt er daarom voor gekozen de gegevens van de laagste SBEID (de langst-bestaande bedrijfseenheid) te behouden.

Stappen a)-d) zorgen voor een uniek bestand op de combinatie gegevens uit de RINPERSOONS-RINPERSOON-SBEID. Met andere woorden: We eindigen met jaarbestanden (2010-2022) waarin elke persoon (RINPERSOONS-RINPERSOON) wordt gekoppeld aan één bedrijfseenheid (SBEID).

Enquête Beroepsbevolking (EBB)

In de EBB selecteren we uitsluitend respondenten in de eerste peiling van de Enquête Beroepsbevolking (EBB) in de leeftijd van 15 tot 75 jaar die behoren tot de werkzame beroepsbevolking (volgens de internationale definitie). In deze eerste peiling wordt het beroep in de eerste werkkring vastgesteld.¹ De eerste werkkring is die werkkring waarin het grootste aantal uren per week gewerkt wordt. Op basis van de ontvangen informatie over het beroep van de respondent wordt de International Standardized Classification of Occupations (ISCO) code afgeleid.

Bovenstaande exercitie leidt tot één bestand per jaar (2010-2022) met EBB gegevens, met daarin de variabelen RINPERSOONS, RINPERSOON, EBBTW1ISCO2008V, EBBGEWJAARGEWICHTP1A, JAAR en EBB-waarneming. Dit bestand bevat maximaal één regel per persoon; m.a.w. het bestand is uniek op de sleutels RINPERSOON-RINPERSOONS.

Koppeling EBB-bestand en Polisadministratie

De voorlaatste stap in het proces betreft het verrijken van de bestanden met informatie uit de polisadministratie met de EBB-data. Omdat beide bestanden op jaarbasis informatie over één persoon bevatten, is de koppeling eenvoudig; elke regel in de bestanden van de polisadministratie zal (op jaarbasis) gekoppeld worden met hoogstens één regel in de EBB-data.

In deze koppeling worden in eerste instantie alle regels afkomstig uit de polisadministratie behouden; ook regels die niet op jaarbasis koppelen met een regel uit de EBB zijn nodig in de volgende stap.

Vooruit- en terugleggen informatie over beroep

Na het uitvoeren van bovenstaande stappen ligt er een bestand met daarin per jaar (2010-2022) informatie over personen, hun beroepen en hun werkgever. In een laatste stap wordt de informatie over beroep afkomstig uit één van de jaargangen van de EBB ook toegekend aan andere jaren, waarvoor we geen EBB-data beschikbaar hebben. Stel: een persoon komt in 2014 als respondent voor in de EBB en we hebben van deze persoon informatie over zijn of haar beroep. En, stel: deze persoon komt in 2012-2022 voor in de polisadministratie, omdat zij in die periode loon ontvangt in Nederland. In deze laatste stap kennen we de beroepcode zoals

¹ Het kan in een enkel geval voorkomen dat een persoon meerdere keren in één kalenderjaar in de eerste peiling van de EBB voorkomt. In deze gevallen wordt uitsluitend de observatie van deze persoon met het hoogste gewicht (EBBGEWJAARGEWICHTP1A) meegenomen in de analyse.

geobserveerd in 2014 toe aan alle jaren waarin we deze persoon aantreffen in de polisadministratie.

Aandachtspunten

De gebruiker dient rekening te houden met het volgende:

- Gegevens afkomstig van de polisadministratie worden door Team Verwerking Polisadministratie op meso- en macro-niveau op plausibiliteit gecontroleerd. Dit betekent dat er op micro-niveau (op het niveau van een aangifte) fouten kunnen staan in de polisadministratie. Dit is niet tot in detail te controleren.
- Uit de polisdata zijn alleen loonaangiftes meegenomen met een SBEID die teruggevonden kunnen worden in het ABR.
- Uit de polisdata zijn alleen gegevens meegenomen van unieke RINPERSOON-RINPERSOONS-SBEID combinaties die op jaarbasis een positief geaggregeerd SBASISLOON hebben.
- Het aandeel personen in de EBB waaraan we geen beroep kunnen toekennen (EBBTW1ISCO2008V == 9999) varieert in de jaren 2010-2022 tussen de 1,7 en 0,5 procent.
- Omdat het uitgangspunt de polisadministratie is, zullen veel zelfstandig ondernemers die zichzelf geen loon uitkeren niet voorkomen in het bestand.
- Het op jaarbasis optellen van alle gewichten (EBBGEWJAARGEWICHTP1A) zal niet leiden tot het aantal personen in de werkzame beroepsbevolking zoals gepubliceerd op [StatLine](#). Dit komt enerzijds doordat we de polisadministratie als uitgangspunt nemen (zie voorgaande punt), alsook het maken van enkele selecties in de data (zie eerder in dit document).

Bijlagen

variabelenwaardenPBB.xlsx

3. Bestandsopbouw en toelichting

Bestandsopbouw

Hieronder volgt een overzicht van alle variabelen in het bestand.

Nr	Variabele en omschrijving	Form.
1	RINPERSOONS Indicator oorsprong persoon identificerend nummer	A1
2	RINPERSOON Identificatiesleutel van een natuurlijk persoon	A9
3	JAAR Jaar waarop de gegevens betrekking hebben	F4
4	SBEID Identificatiesleutel van een bedrijfseenheid	F8
5	EBBTW1ISCO2008V Beroep volgens de <i>International Standardized Classification of Occupations</i>	A4
6	EBBGEWJAARGEWICHTP1A Weegfactor persoon	F4.2
6	EBB_waarneming Indicatie of EBBTW1ISCO2008V uit de desbetreffende regel afkomstig is van een EBB-waarneming (1) of van een imputatie (0).	A1

Toelichting op de variabelen

In deze paragraaf wordt informatie over de variabelen gegeven.

RINPERSOONS

Deze variabele geeft weer vanuit welke bron het persoon identificerend nummer (RINPERSOON) is afgeleid.

RINPERSOON

Dit nummer identificeert een natuurlijk persoon. Het is een betekenis- en dimensie-loos nummer.

JAAR

Jaar waarop de gegevens betrekking hebben.

SBEID

Nummer van de bedrijfseenheid, afkomstig uit de polisadministratie. De bedrijfseenheid (SBEID) is een door het CBS gemaakte volgnummer van bedrijven en wordt als statistische eenheid gebruikt.

Een bedrijfseenheid wordt gekenmerkt door autonomie en externe gerichtheid. De autonomie heeft betrekking op beslissingen over het productieproces van een productengroep en het aanbieden van producten of diensten aan derden. Daarnaast moet de bedrijfseenheid te enquêteren zijn. Een bedrijfseenheid kan uit meer dan één juridische eenheid bestaan.

EBBTW1ISCO2008V

Een viercijferige code die een beroep weergeeft, gebaseerd op de *International Standardized Classification of Occupations* (ISCO).

EBBGEWJAARGEWICHTP1A

Een gewicht dat wordt toegekend aan een persoon, waarna schattingen van populatieparameters verkregen worden door de gegevens met deze weegfactoren te vermenigvuldigen en vervolgens op te tellen.

EBB_waarneming

Indicatie of EBBTW1ISCO2008V uit de desbetreffende regel afkomstig is van een EBB-waarneming (1) of van een imputatie (0)



Versiegeschiedenis

<i>Bestandsnaam</i>	<i>Reden</i>
PBB2010-2022V1	Eerste plaatsing