

Gegevens van voornamen en familienamen uit de Basisregistratie Personen

Gerrit Bloothoof
Meertens Instituut KNAW, Amsterdam
2022

Deze tekst is gebaseerd op een document dat in 2009 is geschreven voor de Autoriteit Persoonsgegevens (met aanvullingen in 2010) ter onderbouwing van (de beperking van) de persoonsgegevens die sindsdien op de voornamenbank en de familienamenbank online worden getoond. De tekst is geheel herzien en geactualiseerd in december 2022.

Het Meertens Instituut van de Koninklijke Nederlandse Academie voor Wetenschappen heeft als missie het onderzoek naar de taal en cultuur in het dagelijks leven in Nederland, in internationaal, vergelijkend en historisch perspectief. Daar is de naamkunde onder begrepen, in het bijzonder de studie van persoonsnamen. Naast onderzoek is ook publieksvoorlichting over persoonsnamen belangrijk. Daarvoor kan momenteel op het internet de Nederlandse Voornamenbank (NVB, www.meertens.know.nl, initiatief ondersteund door de Nederlandse Vereniging voor Ambtenaren Burgerzaken) en de Nederlandse Familienamenbank (NFB, sinds 2015 beheerd door het Centraal Bureau voor Genealogie, www.cbfgfamilienamen.nl/nfb/) geraadpleegd worden.

In 2006 werd het Meertens Instituut KNAW (samen met de Universiteit Utrecht) door de Minister van Binnenlandse Zaken en Koninkrijksrelaties via de Rijksdienst voor Identiteitsgegevens (RvIG) geautoriseerd tot een selectie uit de Basisregistratie Personen (toen Gemeentelijke Basisadministratie) gekregen met betrekking tot de voornamen van alle Nederlanders (zij met Nederlandse nationaliteit), terwijl in 2007 hetzelfde is gebeurd voor de familienamen. Daarover is het College Bescherming Persoonsgegevens geïnformeerd (document m1393937). Nadien hebben er selectieverstrekingen voor voornamen plaats gevonden in 2010 (update), 2015 (volledig) en 2018 (volledig, ook uit RNI), om de gegevens zoveel mogelijk up-to-date te houden. In 2023 zal weer een volledige selectieverstreking plaatsvinden, nu alleen op autorisatie van het Meertens Instituut KNAW. Het logisch ontwerp van de Basisregistratie Personen met een beschrijving van de daarin opgenomen informatie is te vinden op <https://www.rvig.nl/brp/werken-met-de-basisregistratie-personen-brp/logisch-ontwerp>

In dit document wordt de omvang en de inhoud van de selecties besproken. En omdat bij online presentatie en bij publicatie van onderzoeksresultaten aspecten van privacy aan de orde zijn met betrekking tot namen, naamfrequenties, geboortejaren en geografische verspreiding, wordt daarover een toelichting geven. Hierbij worden de grenzen van gegeven informatie gemotiveerd. Deze grenzen zijn zodanig gekozen dat individuele personen nooit traceerbaar zijn en privacy niet in het geding is.

Dat zowel de online voornamenbank als de familienamenbank in een behoefte voorzien blijkt uit respectievelijk meer dan 10 en 6 miljoen pageviews per jaar. Daarnaast heeft de beschikbaar gestelde informatie geleid tot tientallen wetenschappelijke publicaties. Zeer regelmatig wordt het Meertens Instituut vanuit de media (radio, TV, schrijvende pers) benaderd met vragen over namen, die vanuit de beschikbare gegevens en expertise goed te beantwoorden zijn.

I. De voornamenselectie

I.1 De omvang van de voornamenselectie

Er worden via autorisatie op grond van Wet BRP, Art 3.13, 1^e lid en Art 44 besluit BRP gegevens uit de Basisregistratie Personen (BRP) gevraagd van alle ingeschreven personen met de Nederlandse nationaliteit, en degenen met opgeschorte persoonslijsten (Register Niet-Ingezetenen, RNI). De gedigitaliseerde BRP werd in 1994 geïntroduceerd. Oudere gegevens zijn in de jaren ervoor overgezet van persoonskaarten naar digitale persoonslijsten.

Het is voor naamkundig onderzoek essentieel om een volledig overzicht te hebben van de voornamen van de gehele bevolking, in plaats en tijd. Daarbij wordt beperkt tot hen die de Nederlandse nationaliteit bezitten. Om optimale inzichten te krijgen in de (historische) naamgeving zijn niet alleen gegevens van de huidige bevolking van belang maar ook die van overledenen. Deze kunnen in de Basisregistratie Personen zijn opgenomen met een eigen persoonslijst, maar ook met een aantal basisgegevens als ouders bij kinderen worden genoemd. Daarnaast zijn degenen die in Nederland zijn geboren, maar er niet meer woonachtig zijn (en in de RNI zijn opgenomen) voor een volledig beeld van de naamgeving relevant.

De selectie omvat gegevens uit circa 21 miljoen persoonslijsten. Na reconstructie van personen zonder persoonslijst, die vaak meervoudig vermeld zijn als ouders, is naamkundige informatie over circa 25 miljoen personen beschikbaar.

I.2 De gegevens in de voornamenselectie

Voor het naamkundig onderzoek naar voornamen worden gegevens gevraagd van de categorieën <01:persoon>, <02:ouder1>, <03:ouder2>, <09:kind>, <51:persoon historisch>, <07:inschrijving>, <08:verblijfplaats> en <58:verblijfplaats historisch>.

Naast de categorie persoon zijn de categorieën ouder1, ouder2 en kind nodig (1) voor de reconstructie van personen zonder a-record waardoor gegevens (populariteit) verder terug in de tijd kunnen worden bepaald en (2) voor generationeel onderzoek, bijvoorbeeld naar overdacht van naamgeving van ouders op (klein)kinderen en naar naamgeving in verschillende generaties migranten, of naar onderzoek naar relaties tussen de voornamen van kinderen in een gezin. Met de gegevens voor persoon historisch kan naamsverandering worden bestudeerd.

Van deze categorieën wordt gevraagd <01.10 A-nummer> voor reconstructie en generationeel onderzoek, <02.10 voornamen> als kerngegeven, <02.20 adellijke titel/predicaat> voor naamonderzoek in deze speciale maatschappelijke groep, <03.10 geboortedatum>, <03.20 geboorteplaats>, <03.30 geboorteland>, <04.10 geslachtsaanduiding> als achtergrondgegevens voor een naam. Voor categorie <51:persoon historisch> wordt alleen de historische voornaam en geslacht gevraagd.

Daarnaast wordt gevraagd uit de categorie inschrijving de <67.10 datum opschorting

bijhouding> (overlijdens- of emigratiedatum) en <67.20 omschrijving reden opschorting bijhouding> (overlijden of emigratie), zodat de voornamen van de in Nederland levende bevolking kan worden onderzocht.

Uit categorieën verblijfplaats en verblijfplaats(historisch) wordt <11.60 postcode> gevraagd om de huidige verspreiding van voornamen te kennen, en om de geboortegemeente in 2007 te bepalen. Het jaar 2007 is voor de geografische presentatie gekozen als jaar met nog redelijk gedetailleerde gemeentegrenzen (ten tijde van de eerste selectieverstrekking), die door gemeentelijke herindelingen steeds minder gedetailleerd wordt. Uit de historische postcodes geeft de oudste postcode vanaf 1977 helderheid over de woonplaats bij geboorte ingeval van (inmiddels talrijke) ziekenhuisgeboorten. Daarnaast wordt gevraagd <14.20 datum vestiging in Nederland>, wat van belang is voor generationeel onderzoek naar naamgeving in migranten: of men zeer jong of als volwassene naar Nederland is gekomen en welke invloed dat heeft op de naamgeving aan eigen kinderen.

II. De familienamenselectie

Voor naamkundig onderzoek en voor de presentatie van de familienamen op de Nederlandse Familienamenbank zijn in 2007 de familienaam en aanvullende kerngegevens gevraagd van 16 miljoen personen met Nederlandse nationaliteit die in Nederland wonen (autorisatie op grond van Wet BRP, Art 3.13, 1^e lid en Art 44 besluit BRP).

De familienaamgegevens komen uit de categorieën <01:persoon>, <08:verblijfplaats> en <58:verblijfplaats historisch>. Uit de categorie persoon zijn verkregen <02.30 voorvoegsel geslachtsnaam>, <02.40 geslachtsnaam>, <03.20 geboorteplaats> en <03.30 geboorteland>, en uit de categorie verblijfplaats <09.10 gemeente van inschrijving> en <10.20 gemeentedeel> (om de woonplaats preciezer te kennen) en uit de categorie verblijfplaats historisch eveneens <09.10 gemeente van inschrijving> als een indicatie van de oorspronkelijke herkomstplaats. In deze selectieverstrekking is de postcode niet opgenomen.

III. Voornamen, familienamen en privacy

De voornaam wordt meestal als minder gevoelige informatie beschouwd dan de familienaam omdat koppeling van gegevens op basis van alleen de voornaam zelden eenduidig is. Voor de familienaam ligt dat anders.

We moeten in het algemeen de vraag stellen of privacy in het geding is wanneer:

1. Het bestaan van een voor- of familienaam in Nederland wordt gegeven.
2. De frequentie van een voor- of familienaam voor de levende bevolking van Nederland wordt gegeven.
3. De frequentie van een voornaam per geboortejaar wordt gegeven.
4. De frequentie van een voor- of familienaam per gemeente wordt gegeven.

Bij hoogfrequente gegevens op enig aggregatieniveau bieden de bovengenoemde vier presentaties geen informatie die op persoonlijk niveau als bezwaarlijk kan worden ervaren. Bij kleine aantallen, of zelfs unieke namen, kan dat anders zijn.

III.1 Het bestaan van een voornaam of familienaam in Nederland

Kennis over het bestaan van een voornaam of familienaam in Nederland is van belang voor ambtenaren burgerzaken en ouders bij het beoordelen van de toelaatbaarheid van een gewenste voornaam, of het veranderen van de voor- of familienaam. Daarbij mag er vanuit worden gegaan dat een vermelde naam onder Nederlands recht gegeven is (bij voornamen van buitenlandse herkomst wordt dat in algemene termen ook vermeld, zie het onderstaande voorbeeld). We gaan er vanuit dat met het presenteren van een naam (zonder nadere gegevens over aantal, tijd en plaats) een algemeen maatschappelijk belang is gediend. Het is aannemelijk dat de Nederlandse Voornamenbank onenigheid over de toelaatbaarheid van een voornaam, en zelfs juridische procedures daarover sterk heeft verminderd.

Voorbeeld

“Naast geboorten in Nederland is de voornaam 'Mohamed' in ons land vooral bekend via naamdragers die geboren zijn in Marokko, Suriname, Somalië, Egypte, Tunesië, Sierra Leone, Algerije, Irak, Voormalig Frans West Afrika, Soedan (inclusief Zuid-Soedan), Afghanistan.”

III.2 De frequentie van een voor- of familienaam voor de levende bevolking van Nederland

Kennis over het aantal personen met een weinig voorkomende naam kan soms ongewenst zijn. Wij hanteren daarom een ondergrens van 5 bij de presentie van frequenties in Nederland (op alle aggregatieniveaus). Wanneer vergelijkbare voornaamgegevens in andere landen worden gepresenteerd is dit de hoogste ondergrens (per jaar):

land	bron	minimum
Noorwegen	www.ssb.no/en/befolkning/navn/statistikk/navn	4
Frankrijk	www.insee.fr/fr/statistiques/2540004	3
USA	www.ssa.gov/oact/babynames/	5
Vlaanderen	statbel.fgov.be/nl/themas/bevolking/namen-en-voornamen/voornamen-van-meisjes-en-jongens#figures	5
Schotland	https://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/vital-events/names/babies-first-names	1
Engeland en Wales	www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/datasets/babynamesinenglandandwalesfrom1996	3

Voorbeeld

De familienaam *Schotanus* kwam in 2007 1.382 keer in Nederland voor.

De voornaam *Wietske* kwam in 2017 3.141 keer in Nederland voor.

De familienaam *Bleier* kwam in 2007 minder dan 5 keer in Nederland voor.

De voornaam *Trinetty* kwam in 2017 minder dan 5 keer in Nederland voor.

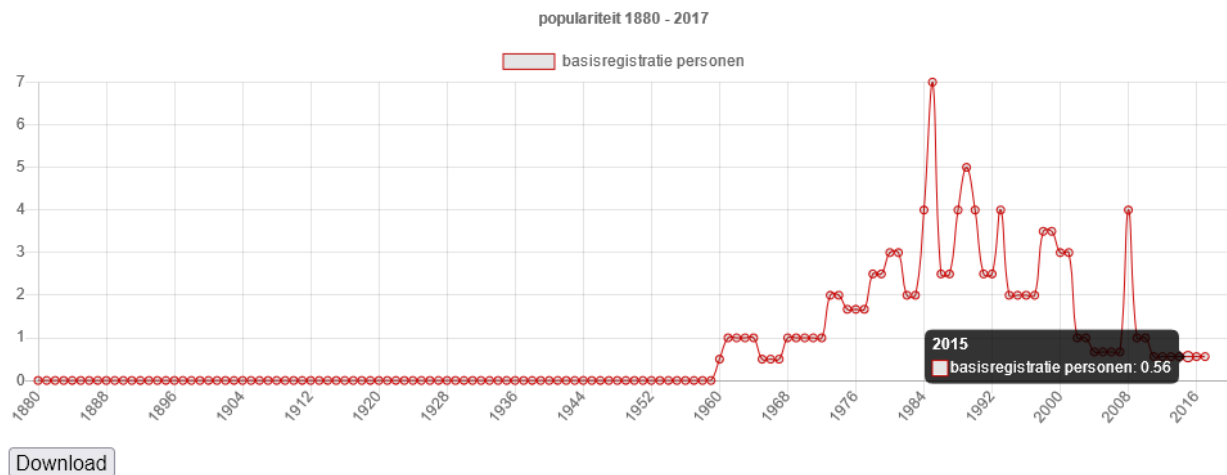
III.3 De frequentie van een voornaam per geboortjaar

In de voornamenbank geeft de uitsplitsing van voornaamfrequentie naar geboortjaar een beeld van de historische populariteit van een voornaam, die in 2022 voor traditionele

voornamen zelfs kan worden gegevens vanaf 1790. Voor gegevens van na 1880 wordt de grens van 5 naamdragers per geboortjaar aangehouden (zie tabel in III.2). Wanneer dat het geval is wordt de frequentie van naburige jaren bijeen wordt genomen tot de grens van 5 is gehaald (waarbij wordt gemiddeld over de betrokken jaren). Naast absolute aantallen kan ook de relatieve frequentie worden getoond, onder dezelfde presentatieregels.

Voorbeeld

Populariteit van 'Annieke' als eerste naam voor vrouwen in de periode 1880-2017



De populariteit van *Annieke* sinds 1880. Na 2011 is de informatie gemiddeld.

Voor presentatie van het aantal naamdragers per jaar moet het totaal aantal naamdragers in een periode een grens overschrijden. Ook dit beperkt de presentatie van gegevens van zeldzamere namen. Die perioden zijn 1790-1879 en 1880-2017, waarbij de gegevens voor 1880 zijn afgeleid uit de namen van bruid en bruidegom in huwelijksakten die in wiewaswie.nl online beschikbaar zijn (en geen privacy beperkingen hebben, maar voor een duidelijk beeld wordt toch een ondergrens van 20 gebruikt). De onderstaande tabel geeft de grenswaarden aan, en het aantal namen dat er aan voldoet (ongeacht geslacht of als eerste voornaam of volnaam). Veruit de meeste voornamen komen maar weinig voor, en hun jaarlijkse populariteit wordt niet getoond.

populariteit tonen in periode	aantal naamdragers 1790-1879	aantal naamdragers 1880-2017	aantal voornamen
niet tonen	<20	<50	427.524
1790-1879	>=20	<50	2.277
1790-2017	>=20*	>=50	4.723
1880-2017	<20	>= 50	10.639

* bij volnaam als eerste voornamen ook nog >=2 om patroniemen uit te sluiten

III.4 De frequentie van een voor- of familienaam per gemeente

In de voornamenbank wordt in de geografische presentatie de naamfrequentie per gemeente gegeven, en via een kleurschaal in een kaart gepresenteerd, om een indruk te geven van de verspreiding van een voornaam of familienaam over Nederland. Deze

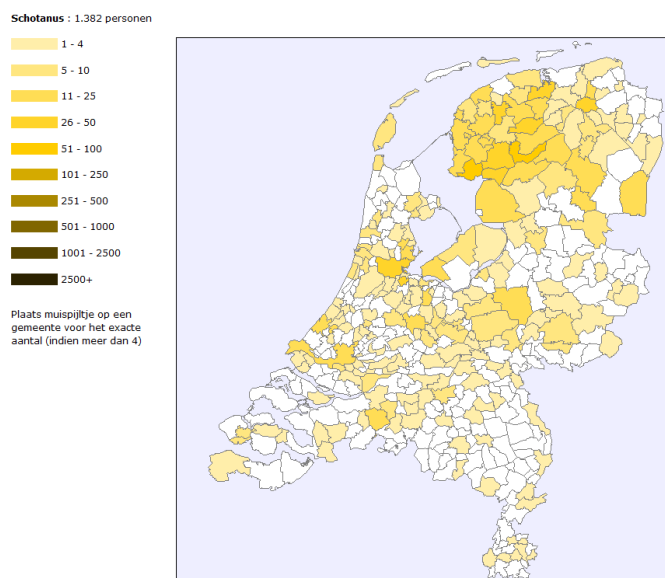
geografische presentatie wordt niet getoond wanneer de totale frequentie in Nederland lager is dan 5.

Voor de presentatie van het aantal naamdragers per gemeente wordt een procedure gehanteerd (zie bijlage A) die op grond van het totaal aantal naamdragers in Nederland een minimum inwonersaantal van een gemeente vereist. Pas als een gemeente daaraan voldoet wordt informatie getoond.

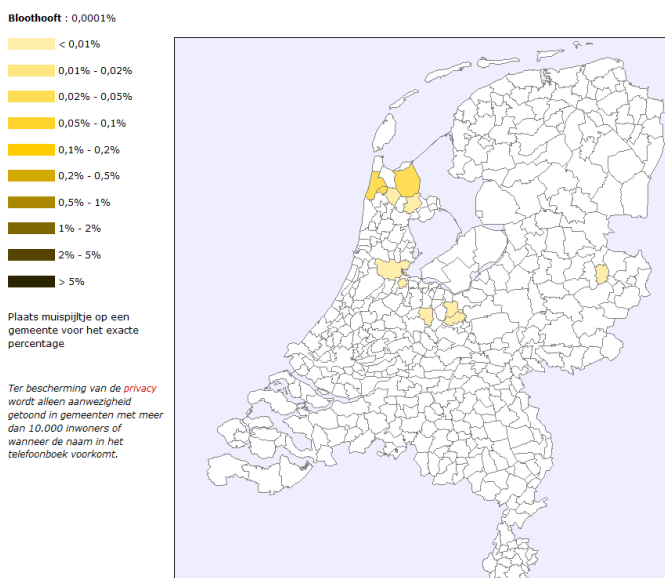
Wanneer een verzameling namen met behulp van een reguliere expressie wordt gezocht, wordt wel de kleurpresentatie van de geografische verspreiding gegeven maar geen getalsmatige informatie.

Voorbeeld

De familienaam Schotanus, volledig (aantal per gemeente in 2007)



De familienaam Bloothoofd, met beperking (percentage per gemeente in 2007)



Kanttekeningen

- In de kaartpresentatie ziet de gebruiker door met de muis op een gemeente te gaan staan, zowel de naam van de gemeente als de (relatieve) frequentie
- De verspreidingskaart van familienamen is gebaseerd op de *woonplaats* in 2007, die van de voornamen op de *geboorteplaats*. Dat maakt het traceren van een persoon die zowel een zeldzame voornaam als een zeldzame familienaam heeft, door het combineren van de beide kaarten minder kansrijk.
- In het Repertorium Nederlandse Familienamen (Van Gorcum 1962-1980), die gebaseerd is op de volkstelling van 1947, worden per gemeente alle familienamen genoemd, met frequentie tot en met 1.

IV. Motivatie van de gegevenspresentatie op de voornamen- en de familienamenbank

IV.1 Voornamen

Bij de presentatie van informatie over voornamen is er een groot maatschappelijk belang in verband met de keuze van voornamen door ouders. Daarbij moet een ambtenaar van de burgerlijke stand beslissen of een door ouders gekozen voornaam voor een in Nederland geboren kind in overeenstemming is met de Nederlandse naamwet (dwz onder Nederlands recht). Het is bijzonder nuttig dat het voor zowel ouders als voor de ambtenaar van de burgerlijke stand mogelijk is om te weten welke voornamen in Nederland al eerder geaccepteerd zijn (wat als openbaarheid van bestuur kan worden beschouwd). Dat kan de consistentie in besluitvorming daarover sterk bevorderen.

De daarbij te tonen gegevens zijn: de enkelvoudige naam, frequentie als eerste naam voor mannen, frequentie als volgnaam voor mannen, frequentie als eerste naam voor vrouwen, frequentie als volgnaam voor vrouwen. Met daarnaast een verspreidingskaart van de geboorteplaatsen voor een voornaam, en zo mogelijk een uitleg over de betekenis van een naam. Zo'n kaart geeft inzicht in regionale voorkeuren en kan daarbij ouders behulpzaam zijn bij het maken van een keuze. Dat geldt ook voor de betekenis van de voornaam.

Toelichting

(1) Met een enkelvoudige voornaam wordt bedoeld een naam zonder spaties, eventueel wel met koppelteken of apostrof. Goedkeuring van een samengestelde voornaam zal vrijwel altijd neerkomen op goedkeuring van de afzonderlijke enkelvoudige voornamen. Voor beslissingen over toelaatbaar aantal voornamen of over het toelaatbaar aantal namen die met een koppelteken aan elkaar zijn verbonden zou een aparte leidraad wenselijk zijn.

(2) Veel voornamen zijn indicatief voor het geslacht van de drager (in Duitsland is dit zelfs verplicht). Maar er zijn ook veel namen die zowel aan jongens als meisjes worden gegeven, zij het vaak met een voorkeur voor een geslacht. Bij een volgnaam, waarbij vaak van vernoeming sprake is, komt dat nog vaker voor. Het is informatief om te weten hoe de verhoudingen per naam liggen.

Voorbeeld (2017):

Jan	man	vrouw
eerste naam	177.095	15
volgnaam	199.217	670

Jip	man	vrouw
eerste naam	2912	1132
volgnaam	194	71

IV.2 Familienamen

Een bestaande familie naam mag niet als voornaam worden gekozen, tenzij de naam als voornaam gebruikelijk is. Ook mag bij naamsverandering niet voor een bestaande familienaam worden gekozen. In deze gevallen is het voor betrokkenen, ambtenaren van de burgerlijke stand en rechters van belang om de huidige familienamenvoorraad van Nederland te kennen.

Kaarten met de verspreiding van de familienamen dienen vooral een cultureel doel (en kunnen ook als cultureel erfgoed worden beschouwd): velen zijn geïnteresseerd in deze informatie die indicatief kan zijn voor de herkomst van een familie.

Bijlage A. Regels voor het tonen van het aantal naamdragers in een gemeente

Bij een beslissing om informatie over het aantal naamdragers in een gemeente al dan niet te tonen moet met de onderstaande factoren en aannamen rekening worden gehouden:

(1) de kans om iemand te vinden is omgekeerd evenredig met het *totaal aantal naamdragers* van een naam (N). Hoe groter het aantal naamdragers hoe moeilijker het is om een bepaalde persoon te traceren.

(2) de kans om iemand in een gemeente te vinden is omgekeerd evenredig met het *inwoneraantal* van die gemeente (i_g). Hoe meer inwoners een gemeente heeft, hoe moeilijk het zal zijn daar iemand te vinden

(3) de kans om iemand in een gemeente te vinden is evenredig met het *aantal naamdragers* in die gemeente (n_g). Hoe meer naamdragers in een gemeente wonen, hoe groter de kans is dat een gezochte persoon daarbij zit.

(4) de kans dat iemand geheim wenst te blijven is omgekeerd evenredig met het aantal naamdragers in een gemeente (n_g). Hier gaat het om een aanname over het begrip 'onbekend willen blijven'. De gedachte is dat hoe minder naamdragers in een gemeente wonen, hoe groter de kans is dat daarbij iemand zit die dat geheim wil houden (want gemakkelijker vindbaar).

Bovenstaande kansen kunnen vermenigvuldigd worden (waarbij er nog een constante factor aan wordt toegevoegd):

P = de vindkans in een gemeente voor iemand die geheim wenst te blijven

$$P = (1/N) * (1/i_g) * (n_g) * (1/n_g) = 1/(N*i_g)$$

P moet onder een bepaalde waarde blijven. Daar kiezen we $1/K$ voor, wat voorgesteld kan worden als de kans om iemand met een bepaalde naam te vinden, die dat niet wil, in een stad met K inwoners.

$$P = 1/(N*i_g) < 1/K$$

waaruit volgt $i_g > K/N$

Dit betekent dat we voor een bepaalde naam het aantal naamdragers kunnen geven in de gemeenten die aan deze eis voldoen.

Onder deze condities blijkt $K = 200.000$ een redelijke keuze.

$N < 5$ wordt nooit in kaart getoond

$N = 5$ $i_g > 40.000$ (naamdragers worden alleen getoond in gemeenten met meer dan 40.000 inwoners)

$N = 10$ $i_g > 20.000$

$N = 20$ $i_g > 10.000$ (veel gemeenten)

$N > 100$ $i_g > 2.000$ (vrijwel overal)

In enkele gevallen kan het voorkomen dat relatief veel naamdragers in een kleine gemeente niet getoond worden. We gaan er echter van uit dat wanneer er meer dan 10 naamdragers in een gemeente zijn, er van een wens tot geheim blijven geen sprake zal zijn. Het aantal van deze naamdragers in een gemeente wordt daarom altijd getoond.

Indien naamdragers niet in alle gemeenten worden getoond komt bij de kaart te staan: "Ter bescherming van de privacy wordt alleen aanwezigheid in gemeenten met meer dan inwoners getoond." (i_g wordt naar beneden afgerond op hele duizendtallen)

Voor de familienamen is gecontroleerd of een naam in het telefoonboek van 2007 (selectiejaar) in een bepaalde gemeente voorkwam. Indien dat het geval is wordt de naam altijd in die gemeente getoond, ongeacht de uitkomst van voorgaande berekeningen (met nog steeds de beperking dat een aantal naamdragers kleiner dan 5 niet wordt getoond).