

HANDLEIDING STATISTIEKPRODUCTIE

15.01.2025



Vlaanderen
is sterk in cijfers

INLEIDING

De hoofdtak van het netwerk Statistiek Vlaanderen en de Vlaamse Statistische Autoriteit is de productie en publicatie van statistieken, ook wel Vlaamse Officiële Statistieken (VOS'en) genoemd. Maar wat verstaan we eigenlijk onder een statistiek? En welke informatie is essentieel om tot een kwaliteitsvolle statistiek te komen? Om het netwerk goed te laten functioneren, is het cruciaal om met een gedeelde woordenschat te werken. Momenteel bestaan er echter grote verschillen in hoe termen en concepten binnen het netwerk worden gebruikt. Deze nota heeft daarom als doel enkele basisconcepten in de Vlaamse statistiekproductie te verduidelijken.

In deze handleiding wordt terminologie uit de GSBPM- en de SDMX-standaarden gebruikt, al verwijzen we voorlopig slechts sporadisch in enkele kanttekeningen naar deze standaarden. De integratie van beide standaarden kan in een volgende versie opgenomen worden.

1 STATISTIEKEN

1.1 WAT IS EEN STATISTIEK?

Veel mensen definiëren een statistiek op de eerste plaats als een cijfer. Maar zo'n definitie is onvolledig. Statistiek gaat namelijk niet alleen over het berekenen van cijfers; het draait ook om de interpretatie ervan.

Neem bijvoorbeeld het cijfer "525 935". Dit is een bestaande statistiek, maar op zichzelf zegt het weinig. Waarvan zijn er 525 935? Of wat stelt dit getal voor? Zonder aanvullende context blijft het betekenisloos. Een *statistiek* bestaat daarom altijd uit twee elementen: het cijfer zelf én de betekenis ervan. Beide zijn even belangrijk en vormen één geheel. In dit geval betekent het cijfer 525 935: 'het aantal inwoners van de gemeente Antwerpen in 2019'.

De betekenis van een statistiek noemen we de *conceptuele definitie*. Dit verwijst naar hoe een gebruiker het cijfer interpreteert. Zo'n definitie vloeit voort uit een specifieke *gebruikersbehoefte*. Zo verwijst een bepaald Vlaams decreet misschien naar het aantal inwoners in de Vlaamse gemeenten en zijn we als Statistiek Vlaanderen verplicht het aantal inwoners in Antwerpen te publiceren. Antwerpse beleidsmakers weten daarnaast waarschijnlijk ook graag hoeveel inwoners hun gemeente telt zodat ze hun diensten hierop kunnen afstemmen. Demografen gebruiken dit cijfer dan weer om bevolkingsgroei en -spreiding in Vlaanderen te onderzoeken.



Statistieken hoeven niet per definitie cijfers te zijn, het kunnen ook kwalitatieve observaties zijn of zelf complexe informatie-objecten. Zo publiceert Agentschap Binnenlands Bestuur bijvoorbeeld per gemeente een statistiek over de top 10 andere gemeenten waarnaar kinderen heen pendelen voor school. Voor de stad Antwerpen kan de observatie op deze statistiek bijvoorbeeld de volgende tekstuele vector zijn:

‘[Antwerpen, Brasschaat, Lier, Brussel, Beveren, Mechelen, ...]’.

1.2 OPERATIONELE DEFINITIE

De conceptuele definitie beperkt zich doorgaans tot de betekenis die een doorsnee statistiekgebruiker aan het cijfer geeft. Ze vertelt echter vaak te weinig over hoe het cijfer precies berekend werd en geïnterpreteerd moet worden. Naast de conceptuele definitie heeft elke statistiek daarom ook een operationele definitie. Deze *operationele definitie* beschrijft tot in detail hoe het cijfer exact is gemeten en geeft zo de precieze betekenis ervan weer. De operationele definitie is daarom vaak veel uitgebreider dan de conceptuele definitie.

Neem opnieuw het aantal inwoners van Antwerpen in 2019. Dit cijfer, 525 935, verwijst naar de bevolking op 1 januari 2019. Een vergelijkbaar maar ander cijfer kon worden berekend voor elke andere dag in het jaar of als een jaargemiddelde. Bovendien verwijst 525 935 naar de wettelijke bevolking gerapporteerd door Statbel. In de wettelijke bevolking worden personen uit het wachtregister niet meegeteld, in tegenstelling tot de ‘gewoon verblijvende’ bevolking, ook al wonen deze personen al geruime tijd in België. Verder verwijst het cijfer 525 935 naar de gemeente Antwerpen zoals gedefinieerd door NIS-code 11002 in 2019. Door fusies, splitsingen of grensaanpassingen verandert het grondgebied van sommige gemeenten over de tijd; bijvoorbeeld, in 2025 fuseerde Antwerpen met Borsbeek, waardoor het grondgebied veranderde. De operationele definitie schept volledige duidelijkheid over al dit soort gemaakte keuzes om een statistiek te berekenen. De volledige operationele definitie van 525 935 luidt

“De grootte van de wettelijke bevolking op 1 januari 2019 van de gemeente Antwerpen (NIS 11002 in 2019). De bevolkingsgrootte wordt aangeleverd door Statebl op basis van het Rijksregister van de natuurlijke personen. De wettelijke bevolking telt alle inschrijvingen in het bevolkingsregister (Belgen en buitenlanders die gemachtigd zijn tot vestiging op het Belgisch grondgebied) en het vreemdelingenregister (buitenlanders die toegelaten of gemachtigd zijn tot een verblijf van meer dan 3 maanden op het Belgisch grondgebied, hetzij voor bepaalde of onbepaalde duur). Bepaalde categorieën buitenlanders (vb. diplomatiek en consulaire personeel) zijn vrijgesteld van inschrijving in de bevolkingsregisters. In sommige gevallen kunnen zij op eigen vraag wel ingeschreven worden. Enkel in dat geval worden zij meegerekend in de bevolkingscijfers.

Het Rijksregister omvat verder ook het wachtregister voor asielzoekers (verzoekers om internationale bescherming) waarin asielzoekers ingeschreven worden door de Dienst Vreemdelingenzaken (DVZ) evenals een wachtregister voor EU-burgers in afwachting van woonstcontrole (waarna deze worden ingeschreven in het vreemdelingenregister en worden meegeteld in de bevolkingscijfers). Sinds 1995 worden personen ingeschreven in het wachtregister voor asielzoekers (verzoekers om internationale bescherming) niet meer meegeteld in de bevolkingscijfers van Statbel. Pas nadat asielzoekers worden overgeschreven van het wachtregister voor asielzoekers naar een regulier bevolkingsregister, na erkenning als vluchteling, na toekenning van een statuut subsidiaire bescherming of na verwerving van een verblijfsvergunning om een andere reden, worden zij opgenomen in de bevolkingsstatistieken van Statbel.”

Dit voorbeeld illustreert dat zelfs relatief eenvoudige statistieken al lastige operationele keuzes vereisen. Die keuzes kunnen leiden tot verschillende cijfers. Eurostat, bijvoorbeeld, rapporteert niet de wettelijke bevolking maar de gewoon verblijvende bevolking en publiceert daarom andere cijfers, ook al interpreteren de meeste gebruikers deze cijfers steeds gewoon als de ‘bevolkingsgrootte’. Het doel van de operationele definitie is dus duidelijkheid te scheppen over de gemaakte keuzes zodat iedereen hetzelfde cijfer kan reproduceren. Als twee onderzoekers met dezelfde operationele definitie verschillende cijfers bekomen, is de definitie onvolledig of onnauwkeurig.

De lengte van een operationele definitie kan sterk variëren. Voor een eenvoudig cijfer afgeleid uit officiële registers, zoals het aantal inwoners in Vlaamse gemeenten, kunnen enkele korte paragrafen volstaan. Bij complexere statistieken, zoals bijvoorbeeld de gemiddelde opinie van inwoners gemeten via een bevraging, is daarentegen vaak een uitgebreid rapport nodig met alle details over, onder andere, de enquête-opzet, de steekproeftrekking, het ontwerp van de vragenlijst, de opvolging van respondenten, statistische correcties voor meet- en selectiefouten of de statistische analyses.



In theorie leidt een operationele definitie steeds tot één exact resultaat. In de praktijk is dit echter niet steeds het geval. Als het productieproces bijvoorbeeld een willekeurige steekproef of simulatie bevat, kan het cijfer variëren. Toch blijft de operationele definitie als concept ook in zo’n situaties overeind. Twee onderzoekers zouden in zo’n situatie op basis van dezelfde operationele definitie gemiddeld steeds tot hetzelfde cijfer moeten komen wanneer zij hun productieproces blijven herhalen. Het gebied van de inferentiële statistiek biedt hiervoor het theoretisch kader, maar dat gaat voorbij het doel van deze nota.

Waarom gebruiken we conceptuele definities als ze, in tegenstelling tot de operationele definitie, niet accuraat zijn? De nuances in de operationele definitie hebben meestal weinig invloed op beleidskeuzes of onderzoeksconclusies die gebruikers

maken op basis van de cijfers. Bovendien is de conceptuele definitie een pak handiger in gebruik voor narratieven en theorievorming. Statistiekgebruikers beperken zich daarom tot de conceptuele definitie om cijfers te interpreteren, ook al zijn er binnen dezelfde conceptuele definitie meerdere operationele definities mogelijk. Of je nu de feitelijke of de verblijvende bevolking gebruikt om de bevolkingsomvang in Antwerpen te meten, de meeste gebruikers gaan het cijfer enkel interpreteren als “het aantal inwoners in Antwerpen in 2019” en zullen in beide gevallen dezelfde resultaten boeken.

1.3 KWALITEIT

In sommige situaties kan er een wanverhouding ontstaan tussen gebruikersbehoeften en de operationele definitie van een statistiek. In zo’n situatie hebben we een *kwaliteitsprobleem* en moeten we onze operationele definitie aanpassen aan de behoeften. Als we bij het cijfer 525 935 bijvoorbeeld merken dat bijna alle gebruikers deze statistiek interpreteren als de grootte van de verblijvende Antwerpse bevolking in plaats van de wettelijke bevolking, moeten we dringend onze operationele definitie herzien en dit cijfer aanpassen. Daarom is het essentieel om voor elke statistiek niet alleen de conceptuele en operationele definitie, maar ook de gebruikersbehoeften grondig in kaart te brengen. Wie gebruikt deze statistiek? Welk decreet dwingt ons ertoe deze statistiek te verzamelen? Voor welke beleidsbeslissingen of welk onderzoek is deze statistiek relevant? ...? Hoewel onderzoek naar deze vragen geen exacte wetenschap is, is het noodzakelijk om de kwaliteit van een statistiek te kunnen beoordelen.

Bovendien kunnen er zich situaties voordoen waarin de operationele definitie voldoet voor de meeste gebruikers maar niet voor een selecte groep gebruikers met specifieke behoeften. Een beleidsmaker die het woonbeleid in Antwerpen moet bepalen, kan bijvoorbeeld exacte cijfers nodig hebben over de verblijvende bevolking in plaats van de wettelijke bevolking. In zo’n geval ontstaat er een gebruikersbehoefte waar we met de statistiek ‘Aantal inwoners in Antwerpen’ niet aan voldoen. Als we besluiten in te spelen op de behoefte van deze selecte groep gebruikers, moeten we onze publicatiestrategie herzien. We produceren dan niet langer één statistiek over “het aantal inwoners in Antwerpen”, maar twee afzonderlijke statistieken: “de wettelijke bevolkingsgrootte van Antwerpen” en “de verblijvende bevolkingsgrootte van Antwerpen”. In zo’n situatie veranderen dus zowel de conceptuele definities als de bijhorende operationele definities.

Merk op dat kwaliteit in de brede zin van het woord kan worden bekeken. Een operationele definitie kan inhoudelijk aansluiten bij een gebruikersbehoefte, maar als de berekening te lang duurt en de statistiek te laat beschikbaar is, verlaagt dit alsnog de kwaliteit. Het is dus belangrijk een evenwicht te vinden tussen verschillende vormen van kwaliteit zoals inhoudelijke nauwkeurigheid, gebruiksvriendelijkheid

en tijdige levering. Een overzicht van alle kwaliteitseisen voor statistieken wordt gegeven in de *Praktijkcode voor Europese Statistieken*¹.



Gebruikersbehoeften, conceptuele en operationele definitie komen weliswaar aan bod in informatiestandaarden zoals de SIMS en de GSBPM maar op een zeer onduidelijke en onoverzichtelijke manier. In de SIMS worden de gebruikersbehoeften pas opgelijst in categorie 12 terwijl dit de start is van een statistiekproductieproces. De operationele definitie zit in de SIMS dan weer op een zeer onsamenhangende manier verspreid over verschillende categorieën terwijl de conceptuele definitie volledig ontbreekt. De GSBPM doet het op dit vlak beter. Het beschrijft een proces dat vertrekt vanuit gebruikersbehoeften en eindigt met een evaluatie. Het expliciteert echter onvoldoende dat de evaluatie gestoeld moet zijn op de gebruikersbehoeften en alle operationele keuzes die onderweg werden gemaakt.

1.4 EXPERIMENTELE STATISTIEKEN

Sommige statistieken ontstaan niet vanuit geobserveerde gebruikersbehoeften, maar vanuit veronderstellingen hierover. Dit kan voortkomen uit de inhoudelijke interesse van een collega of door nieuwe dataverzamelings- of analysemethoden. Deze statistieken noemen we *experimentele statistieken*. Bij gebrek aan concrete gebruikersbehoeften kan de kwaliteit van dergelijke statistieken niet direct worden geëvalueerd.

Daarom verbinden we ons ertoe binnen vijf jaar na de eerste publicatie een evaluatieonderzoek uit te voeren naar de gebruikersbehoeften. Dit onderzoek beantwoordt vragen zoals:

- Wordt de statistiek gebruikt, en zo ja, waarvoor?
- Sluit het gebruik en de interpretatie aan bij de operationele definitie?
- Zijn er suggesties van gebruikers om de statistiek beter af te stemmen op hun behoeften?

Eens de gebruikersbehoeften in kaart zijn gebracht, kan de kwaliteit van de statistiek verder worden geëvalueerd. Afhankelijk van de resultaten kan de statistiek nadien:

- worden geüpgraded tot een volwaardige statistiek;
- worden aangepast op basis van de bevindingen; of
- worden stopgezet indien de kwaliteit onvoldoende is.



Het merendeel van de statistieken van de VSA zijn vermoedelijk experimenteel, aangezien gebruikersbehoeften niet systematisch zijn vastgesteld. Hierdoor is een kwaliteitsbeoordeling van deze statistieken momenteel niet mogelijk. Onderzoek naar gebruikersbehoeften dringt zich op bij deze statistieken.

¹ zie ec.europa.eu/eurostat/documents/4031688/9394211/KS-02-18-142-NL-N.pdf

Tabel 1: Een statistiek bestaat uit een cijfer, een conceptuele definitie die voortvloeit uit gebruikersbehoeften, en een operationele definitie die vertelt hoe het cijfer exact wordt gemeten. Een wanverhouding tussen gebruikersbehoeften en de operationele definitie bepaalt de kwaliteit van de statistiek.

Gebruikersbehoeften:	→ Decreet XX.XX verwijst naar het aantal inwoners in Vlaamse gemeenten in de context van ... → Agentschap Binnenlands Bestuur publiceert cijfers over het aantal inwoners in gemeenten als ... → Academische onderzoekers vragen cijfers over het aantal inwoners in gemeenten om onderzoek te voeren over ... → ...
Conceptuele definitie:	Aantal inwoners van de gemeente Antwerpen in 2019
Operationele definitie:	De grootte van de wettelijke bevolking op 1 januari 2019 in de gemeente Antwerpen (Statbel NIS-code 11002 in 2019). De data worden aangeleverd door Statbel op basis van het Rijkregister van de natuurlijke personen. De wettelijke bevolking verwijst naar ...
Cijfer:	525 935
Kwaliteitsevaluatie:	De gebruikers hebben eerder nood aan de grootte van de verblijvende bevolking in plaats van de wettelijke bevolking

1.5 SAMENGEVAT

Samengevat, voor elke statistiek die we produceren en publiceren hebben we nood aan:

1. de gebruikersbehoeften,
2. een conceptuele definitie op basis van de gebruikersbehoeften,
3. een operationele definitie,
4. het cijfer, en
5. een kwaliteitsevaluatie van de operationele definitie in functie van de gebruikersbehoeften.

Tabel 1 vat deze basisinformatie samen aan de hand van het Antwerps voorbeeld. Als de gebruikersbehoeften van een statistiek nog niet in kaart zijn gebracht, is de statistiek een experimentele statistiek.

2 STATISTIEKREEKSEN

Tabel 2: *We publiceren doorgaans geen individuele statistieken maar statistiekreeksen.*

Conceptuele definitie	Cijfer
Aantal inwoners in Aalst in 2019	86 445
Aantal inwoners in Aalter in 2019	28 906
Aantal inwoners in Aarschot in 2019	30 115
Aantal inwoners in Aartselaar in 2019	14 293
Aantal inwoners in Affligem in 2019	13 228
Aantal inwoners in Alken in 2019	11 499
Aantal inwoners in Alveringem in 2019	5 047
Aantal inwoners in Antwerpen in 2019	525 935
Aantal inwoners in Anzegem in 2019	14 716
...	...
Aantal inwoners in Zwijndrecht in 2019	19 056

2.1 WAT IS EEN STATISTIEKREEKS?

In de vorige paragraaf werd uitgelegd hoe een statistiek verwijst naar één enkel cijfer met een bijbehorende conceptuele en operationele definitie ontstaan vanuit specifieke gebruikersbehoeften. In de praktijk produceren we echter veel statistieken met sterk overlappende gebruikersbehoeften en definities. Zo publiceren we niet alleen het aantal inwoners voor de gemeente Antwerpen, maar ook voor alle andere Vlaamse gemeenten, de drie gewesten en heel België. Aparte definities opstellen en beheren voor al deze cijfers is niet efficiënt.

Om efficiëntie te verhogen werken we daarom met *statistiekreeksen* of cijferreeksen: verzamelingen van statistieken die grotendeels dezelfde conceptuele en operationele definities delen. Tabel 2 toont bijvoorbeeld een statistiekreeks met de inwonersaantallen van alle Vlaamse gemeenten.

2.2 CONCEPTEN, DIMENSIES EN ATTRIBUTEN

Aangezien de conceptuele en operationele definities van individuele cijfers in statistiekreeksen overlappen, kunnen we deze definities opsplitsen in verschillende onderdelen. Deze verschillende onderdelen van de definities worden de *concepten* genoemd. Concepten zijn eigenlijk gewoon variabelen, kenmerken waar de statistieken naar refereren en die kunnen variëren tussen statistieken. Bij de statistiekreeks in tabel 2 kunnen we bijvoorbeeld drie concepten onderscheiden: (1) de parameter 'aantal inwoners', (2) het jaartal '2019', en (3) de gemeente. Niet alle statistieken die we publiceren geven immers het aantal inwoners als parameter weer, en evenmin geven ze allemaal informatie voor het jaar 2019. De statistiekreeks in tabel 2 toont bovendien al aan hoe gemeente varieert over verschillende statistieken.

Binnen één statistiekreeks wordt een onderscheid gemaakt tussen twee soorten concepten. Sommige concepten helpen elk cijfer uniek te identificeren, terwijl anderen enkel bijkomende informatie geven over de cijfers. De concepten die de cijfers helpen identificeren binnen een statistiekreeks bepalen de *dimensies* van de statistiekreeks. Ze worden ook wel *identificeerders* (identifiers) genoemd. In tabel 2 is er slechts één dimensie, namelijk de gemeente, aangezien elk cijfer het aantal inwoners van een andere gemeente weergeeft en de cijfers enkel van elkaar verschillen omdat ze naar een andere gemeente verwijzen.

De concepten die enkel bijkomende informatie geven zonder de cijfers verder te identificeren worden daarentegen *attributen* genoemd. Attributen worden daarom ook soms *beschrijvers* (describers) genoemd. In tabel 2 zijn de parameter en het jaartal de attributen, omdat elk cijfer verwijst naar een inwonersaantal in 2019 en niets anders. Attributen kunnen informatie bevatten voor de hele statistiekreeks, maar kunnen evengoed enkel informatie bevatten voor één enkel cijfer of een beperkte groep van cijfers binnen de statistiekreeks.

Onthoud echter dat attributen dimensies kunnen worden wanneer verschillende statistiekreeksen worden samengevoegd. Om bijvoorbeeld bevolkingsdichtheden te berekenen, moeten de cijfers uit tabel 2 worden gecombineerd met cijfers over de oppervlakte van alle gemeenten. In die samengestelde tabel wordt de parameter een dimensie in plaats van een attribuut, omdat sommige cijfers verwijzen naar de parameter 'aantal inwoners', terwijl anderen verwijzen naar de parameter 'oppervlakte' van de gemeenten. Hetzelfde gebeurt met het concept jaartal wanneer we aantallen inwoners combineren over verschillende jaren heen voor een tijdreeks. Omgekeerd kan een dimensie ook een attribuut worden wanneer we een tabel opsplitsen volgens die dimensie. Bijvoorbeeld, als we longitudinale data op jaarbasis opsplitsen in aparte tabellen per jaartal, wordt jaartal in die nieuwe tabellen slechts een attribuut in plaats van een dimensie. Het verschil tussen attributen en dimensies is dus relatief want het hangt af van welke tabel je precies bekijkt. Over al mogelijke statistieken bekeken zijn alle concepten dimensies want geen enkele operationele definitie is dezelfde.



We hebben hier een hiaat in het SDMX model blootgelegd. De documentatie van SDMX bespreekt nergens hoe attributen dimensies kunnen worden of vice versa wanneer datasets met statistiekreeksen worden samengevoegd of uitgesplitst. Om deze reden is de uitwerking van attributen in het model ook veel beperkter dan de uitwerking van dimensies. Hierdoor kan SDMX moeilijker gelinkt worden aan de ideeën rond conceptuele definities, operationele definities en kwaliteit van statistiekreeksen.

2.3 CONCEPTUELE EN OPERATIONELE DEFINITIE

Net zoals elke afzonderlijke statistiek worden statistiekreeksen ook beschreven via een conceptuele en een operationele definitie. Dit vergemakkelijkt informatiebeheer over de statistiekreeksen.

De *conceptuele definitie* verwijst naar de interpretatie die een doorsnee gebruiker aan de statistiekreeks toekent. De statistiekreeks in tabel 2 kan conceptueel worden gedefinieerd als het “aantal inwoners in Vlaamse gemeenten in 2019”. Deze definitie sluit aan bij specifieke gebruikersbehoeften, zoals bijvoorbeeld beschreven in tabel 1.

De *operationele definitie* van een statistiekreeks beschrijft hoe de cijferreeks precies is verzameld en reproduceerbaar is. Bovendien specificeert de operationele definitie welke waarden de attributen en dimensies binnen de reeks precies aannemen. Voor tabel 2 luidt de operationele definitie bijvoorbeeld: “De grootte van de wettelijke bevolking op 1 januari 2019 per gemeente volgens de NIS-code-indeling in 2019. De wettelijke bevolking omvat ...” Deze definitie benoemt heel duidelijk ‘gemeente’ als dimensie en geeft aan hoeveel gemeenten de reeks omvat, namelijk de 300 Vlaamse gemeenten volgens de NIS-code-indeling van 2019, en bijvoorbeeld niet de 285 Vlaamse gemeenten vanaf 2025. Daarnaast beschrijft de definitie ook heel duidelijk de waarde van de attributen. De reeks bevat namelijk cijfers over de grootte van de wettelijke bevolking en bijvoorbeeld niet de verblijvende bevolking, en voor het jaar 2019 en geen ander kalenderjaar.

Door de exacte beschrijving van dimensies bepaalt de operationele definitie bovendien strikt uit hoeveel cijfers een statistiekreeks bestaat, zelfs als sommige cijfers een missende of versluierde waarde hebben. De reeks in tabel 2 bestaat uit 300 cijfers, voor elke Vlaamse gemeente één. Als er in de reeks ook cijfers zouden worden gepubliceerd voor Vlaanderen, Wallonië, Brussel en België, verandert de dimensie ‘gemeente’ naar het ruimer concept ‘geografisch gebied’ met 304 mogelijke waarden en dikt de reeks met evenveel cijfers aan. Splitsen we de cijfers verder uit naar mannen en vrouwen, dan wordt de dimensie ‘geslacht’ toegevoegd met twee mogelijke waarden (man of vrouw) en bevat de reeks 608 cijfers. Inclusief de totalen over beide geslachten komt het totaal op 912 cijfers. Voeg daar nog het percentage mannen en vrouwen per gemeente aan toe, dan bevat de reeks 1520 cijfers.

Een kleine tip om de operationele definitie duidelijk uit te schrijven: Maak slim gebruik van voorzetsels zoals ‘voor’, ‘in’ of ‘per’ om de attributen en dimensies aan te duiden. Zo kan je bijvoorbeeld een statistiekreeks definiëren als het “aantal inwoners IN het jaar 2019 PER geografisch gebied (Vlaamse gemeenten, Belgische gewesten en heel België volgens de NIS-code-indeling van 2019), VOOR elk geslacht (mannen, vrouwen, totaal), en PER leeftijdsgroep (0-18, 19-30, 31-65, 66+ jaar, totaal).”

Hoewel het bij de conceptuele definitie minder strikt is om alle attributen en dimensies expliciet te vermelden, is dit ook een goede praktijk. Voor de statistiekreeks

Tabel 3: De conceptuele definitie van een statistiekreeks verwijst naar de interpretatie van een doorsnee gebruiker terwijl de operationele definitie alle concepten (attributen en dimensies) van de reeks duidelijk beschrijft.

AANTAL MANNEN EN VROUWEN IN VLAAMSE GEMEENTEN IN 2019. Het aantal en het percentage mannen en vrouwen in de wettelijke bevolking op 1 januari 2019 per gemeente volgens de NIS-code-indeling in 2019.			
Gemeente	Geslacht	Aantal	Percentage
Aartselaar	man	7089	49.6
Aartselaar	vrouw	7204	50.4
Antwerpen	man	262921	50.0
Antwerpen	vrouw	263014	50.0
Boechout	man	6506	49.0
Boechout	vrouw	6760	51.0
Boom	man	9024	49.5
Boom	vrouw	9220	50.5
Borsbeek	man	5270	48.6
Borsbeek	vrouw	5584	51.4

in tabel 2 kunnen we bijvoorbeeld kiezen tussen de conceptuele definities ‘bevolkingsomvang’ of ‘aantal inwoners in de Vlaamse gemeenten in 2019’. De tweede definitie is uiteraard een pak duidelijker als omschrijving. De eerste definitie is anderszids een pak beknopter en gemakkelijker te citeren.

Net zoals bij individuele statistieken bepaalt een wanverhouding tussen gebruikersbehoeften en de operationele definitie de kwaliteit van een statistiekreeks. Als de operationele definitie niet in lijn ligt met de behoeften van gebruikers moet de reeks worden aangepast of worden geschrapt.

De presentatie van een statistiekreeks gebeurt doorgaans niet zoals in tabel 2. Attributen worden vaak uit de rijen verwijderd en opgenomen in de titel of conceptuele definitie van de statistiekreeks. Tabel 3 toont een meer gangbare weergave van een uitgebreidere cijferreeks. In deze tabel zie je dat de concepten ‘aantal inwoners’ en jaartal ‘2019’ enkel worden vermeld in de titel van de tabel. De tabel telt verder drie dimensies. Dimensies gemeente en geslacht staan in de rijen. De derde dimensie is de statistische parameter die een verschil maakt tussen aantallen en percentages. Deze dimensie staat in de kolommen.



Binnen de huidige SDMX-standaard stellen we statistiekreeksen voor met slechts één kolom voor de cijfers. De structuur in tabel 2 zou hiervoor gepivoteerd moeten worden. Hoogstwaarschijnlijk versoepelt de SDMX-standaard echter in de toekomst waardoor verschillende kolommen cijfers kunnen bevatten.

2.4 CONCEPTENLIJST

Om consistentie te waarborgen tussen statistieken en statistiekreeksen, hanteren we een algemene conceptenlijst conform de SDMX-standaard. Deze conceptenlijst biedt een volledig overzicht van alle concepten die voorkomen in de door ons geproduceerde statistieken. Elk concept wordt voorzien van een korte identificatiecode, die kan worden gebruikt als snelle referentie in communicatie of als variabelenaam in datasets en datatabellen.

Grosso modo onderscheiden we twee soorten concepten:

- Algemene concepten: Dit zijn terugkerende concepten die vaak voorkomen in verschillende statistieken, zoals jaartal of geografisch gebied. De beschrijving van deze concepten is doorgaans beperkt.
- Het restconcept 'statistiekreeks': Dit concept omvat de rest van de operationele definities en dekt daarmee alle details die niet onder de algemene concepten vallen.

Aan elk concept wordt bovendien een codelijst gekoppeld, eveneens conform de SDMX-standaard. Een codelijst bepaalt welke waarden een concept kan aannemen in een statistiekreeks. Voor het concept 'geslacht' bevat de codelijst bijvoorbeeld de waarden 'man', 'vrouw', en 'andere', maar ook bijvoorbeeld de waarde 'totaal', om totaalcijfers over beide geslachten aan te duiden.

Voor het restconcept statistiekreeks kunnen de waarden zeer uitgebreid zijn. Bij statistiekreeksen die gebaseerd zijn op bevragingen omvat dit restconcept in principe de volledige operationele beschrijving van de bevraging. Wanneer het ontwerp van de bevraging wordt aangepast, leidt dit tot een nieuwe statistiekreeks en een nieuwe waarde voor het restconcept. Net als bij de concepten voorzien we ook korte identificatiecodes voor elke waarde.

Merk op dat de codelijst voor het concept statistiekreeks automatisch een overzicht biedt van alle statistiekreeksen die we produceren.

2.5 HOE BEPALEN WE CONCEPTEN EN DIMENSIES?

De bepaling van statistiekreeksen op basis van concepten is geen exacte wetenschap, maar eerder een arbitrair proces waarin verschillende keuzes gemaakt moeten worden. Een eerste keuze is welke concepten je gebruikt als attribuut om statistiekreeksen te onderscheiden en welke als dimensies binnen deze statistiekreeksen. In het ene uiterste gebruik je alle concepten als attributen waardoor een aparte reeks voor elke individuele statistiek ontstaat, zoals in tabel 4a. In het andere uiterste gebruik je alle concepten als dimensies waardoor één enkele reeks ontstaat met alle cijfers erin, zoals in tabel 4b. Beide benaderingen zijn uiteraard onpraktisch. In de praktijk zoeken we een evenwicht dat zowel overzichtelijk als werkbaar is.

Tabel 4: Concepten kunnen worden gebruikt om aparte cijferreeksen te definiëren of als dimensie in één grote cijferreeks.

(a) Elk cijfer in aparte statistiekreeksen.

Aantal inwoners in Vlaanderen
6 821 770
Percentage Vlamingen met gevorderde digitale vaardigheden
26,1
Aantal zeugen in Vlaanderen
339 103
Ondergrens 95%-CI gemiddelde vertrouwen in provinciale overheid
2,34
...
...

(b) Één statistiekreeks voor alle cijfers.

Parameter	Cijfer
Aantal inwoners in Vlaanderen	6 821 770
Percentage Vlamingen met gevorderde digitale vaardigheden	26,1
Aantal zeugen in Vlaanderen	339 103
Ondergrens 95%-CI gemiddelde vertrouwen in provinciale overheid	2,34
...	...

De tweede keuze is welke concepten binnen een statistiekreeks worden samengevoegd of uitgesplitst. Bijvoorbeeld: cijfers over mannen en vrouwen in gemeenten kun je uitsplitsen op basis van ‘parameter’ en ‘geslacht’, zoals in tabel 5b. Je kunt er ook voor kiezen om beide concepten te combineren in één dimensie, zoals in tabel 5a. Ook hier gaat het om een afweging tussen eenvoud en gebruiksgemak.

Een derde keuze is de definitie van de concepten zelf. Stel dat je cijfers publiceert over inwoners van Vlaamse gemeenten, uitgesplitst naar geslacht en leeftijd, zonder deze kenmerken te kruisen. Je kunt hiervoor aparte dimensies definiëren voor ‘geslacht’ en ‘leeftijd’, zoals in tabel 6a. Alternatief kun je deze herdefiniëren tot één dimensie ‘bevolkingsgroep’, die alle waarden van geslacht en leeftijd bevat, zoals in tabel 6b.

Deze voorbeelden tonen aan dat de bepaling van statistiekreeksen arbitrair is. Richtlijnen zijn daarom essentieel om consistentie en werkbaarheid te waarborgen.

Richtlijn 1. Cijfers die enkel verschillen in geografische indeling worden gecombineerd in één statistiekreeks. Gebruik de geografische indeling als dimensie. Bij voorkeur wordt de NIS-code-indeling van Statbel gevolgd. Hierbij moet steeds de versie van de NIS-code-indeling vermeld worden. Naast NIS-codes kunnen echter ook andere geografische gebieden, zoals postcodes, toeristische regio’s of ziekenhuisnetwerken, worden opgenomen. Voor hiërarchische indelingen kunnen voor de volledigheid extra dimensies worden toegevoegd zoals in tabel 7.

Tabel 5: De indeling van dimensies is arbitrair. Kenmerken kunnen gecombineerd worden in één dimensie of als aparte dimensies worden gespecificeerd.

(a) Geslacht en parameter worden gecombineerd als één dimensie.

Bevolkingsaantal 2019		
Gemeente	Parameter	Observatie
A	totaal aantal	300
A	aantal vrouwen	180
A	percentage vrouwen	60
A	aantal mannen	120
A	percentage mannen	40
B	totaal aantal	500
B	aantal vrouwen	100
B	percentage vrouwen	20
B	aantal mannen	400
B	percentage mannen	80

(b) Geslacht en parameter zijn aparte dimensies.

Bevolkingsaantal 2019			
Gemeente	Geslacht	Parameter	Observatie
A	totaal	aantal	300
A	vrouwen	aantal	180
A	vrouwen	percentage	60
A	mannen	aantal	120
A	mannen	percentage	40
B	totaal	aantal	500
B	vrouwen	aantal	100
B	vrouwen	percentage	20
B	mannen	aantal	400
B	mannen	percentage	80

Tabel 6: Concepten in een statistiekreeks kunnen arbitrair worden hergedefinieerd.

(a) Aparte dimensies voor geslacht en leeftijd.

Bevolkingsaantal 2019			
Gemeente	Geslacht	Leeftijd	Observatie
A	mannen	totaal	90
A	vrouwen	totaal	130
A	totaal	18-30 jaar	60
A	totaal	31-65 jaar	120
A	totaal	66+ jaar	40
B	mannen	totaal	350
B	vrouwen	totaal	150
B	totaal	18-30 jaar	20
B	totaal	31-65 jaar	400
B	totaal	66+ jaar	80

(b) Geslacht en leeftijd worden gecombineerd in de dimensie 'bevolkingsgroep'.

Bevolkingsaantal 2019		
Gemeente	Bevolkingsgroep	Observatie
A	mannen	90
A	vrouwen	130
A	18-30 jarigen	60
A	31-65 jarigen	120
A	66+ jarigen	40
B	mannen	350
B	vrouwen	150
B	18-30 jarigen	20
B	31-65 jarigen	400
B	66+ jarigen	80

Tabel 7: Voor hiërarchische dimensies kunnen bijkomende dimensies worden toegevoegd, hoewel niet strikt noodzakelijk.

Bevolkingsaantal 2019			
Geografisch gebied	Gewest	Provincie	Aantal inwoners
Vlaams Gewest			6 589 069
Waals Gewest			3 633 795
Brussels Gewest			1 208 542
Prov. Antwerpen	Vlaams Gewest		1 857 986
Aartselaar	Vlaams Gewest	Prov. Antwerpen	14 293
Antwerpen	Vlaams Gewest	Prov. Antwerpen	525 935
Boechout	Vlaams Gewest	Prov. Antwerpen	13 266
Boom	Vlaams Gewest	Prov. Antwerpen	18 244
Borsbeek	Vlaams Gewest	Prov. Antwerpen	10 854
...

Richtlijn 2. Cijfers die alleen verschillen in tijdsperiode worden samengevoegd in één statistiekreeks, met de tijdsperiode als dimensie.



De keuze om tijdsperiodes als dimensie te behandelen, betekent niet dat alle cijfers steeds in één dataset opgeslagen moeten worden. Zo is het op een file server handiger datasets per jaar te bewaren, terwijl deze datasets toch statistieken bevatten uit dezelfde statistiekreeks.

Richtlijn 3. Combineer cijfers die op dezelfde manier zijn berekend en uit dezelfde bron komen in één reeks. Dit vereenvoudigt de verwerking en documentatie.

Richtlijn 4. Statistische parameters zoals frequenties en gemiddelden worden zo veel mogelijk opgesplitst in aparte reeksen. Dit vergemakkelijkt de beschrijving van de statistiekreeks. Parameters die op een natuurlijke manier bij elkaar horen, vormen een uitzondering en worden wel gecombineerd in dezelfde statistiekreeks:

- Proporties/percentages berekend op frequenties uit één enkele statistiekreeks, worden ook aan deze statistiekreeks toegevoegd.
- Inferentiële statistieken zoals de onder- en bovengrens van betrouwbaarheidsintervallen, standaardfouten of p -waarden worden toegevoegd aan de statistiekreeks met de bijhorende puntschattingen.



Percentages kunnen op verschillende manieren worden gedefinieerd. Dit moet altijd duidelijk worden omschreven in de operationele definitie en de parameterdimensie (bijv. het percentage mannen in gemeente X is niet hetzelfde als het percentage Vlaamse mannen die in gemeente X wonen).

Richtlijn 5. Ontwerp statistiekreeksen waarin dimensies zo goed als volledig worden gekruist. Als dit niet mogelijk is, kunnen dimensies worden samengevoegd, bijvoorbeeld een dimensie 'bevolkingsgroep' in plaats van aparte dimensies 'geslacht', 'leeftijd' en 'nationaliteit'. Anders is het eleganter om de statistiekreeks te splitsen in twee aparte reeksen.

Richtlijn 6. Cijfers die gebaseerd zijn op meerdere basisreeksen afgeleid uit verschillende databronnen, zoals bevolkingsdichtheid (gebaseerd op inwonersaantallen en geografische oppervlakten), worden altijd in aparte reeksen ondergebracht. Dit bevordert overzichtelijkheid en documentatie.

Richtlijn 7. Alle concepten worden verzameld in een centrale conceptenlijst die wordt beheerd door enkele collega's verantwoordelijk voor het datamodel. Bij elk concept wordt ook een codelijst voorzien die alle mogelijke waarden op het concept bevat.

2.6 VERSIES VAN STATISTIEKREEKSEN

Statistiekreeksen kunnen doorheen de tijd veranderen om verschillende redenen:

- De operationele definitie van een statistiekreeks wordt aangepast omdat dimensies veranderen, bijvoorbeeld de gemeentelandeling verandert door gemeentefusies.
- De operationele definitie wordt aangepast in lijn met nieuwe gebruikersbehoeften.
- Er werd een fout ontdekt in de berekening van cijfers en deze fout wordt gecorrigeerd.

Omdat we verwacht worden elk gepubliceerd cijfer beschikbaar te houden, creëren we in elk van bovenstaande situaties een nieuwe versie van de statistiekreeks. Deze versie duiden we aan door een versienummer terwijl de cijfers en documentatie van de oude versie bewaard blijven. In het geval van een berekeningsfout voegen we een disclaimer toe aan de documentatie van de oude versie met uitleg over deze fout. Bij wijzigende definities, documenteren we eveneens waarom deze wijziging nodig was. Bij een nieuwe versie van een statistiekreeks publiceren we cijferreeksen ook retrospectief indien dat mogelijk is en gewenst wordt geacht.

2.7 SAMENGEVAT

Samengevat, een statistiekreeks is een reeks van statistieken waarbij gebruikersbehoeften, conceptuele en operationele definities overlappen. Elke statistiekreeks is nauwkeurig gedefinieerd via haar operationele definitie. Deze operationele definitie vertelt exact hoeveel cijfers de statistiekreeks bevat, wat deze cijfers betekenen en hoe de cijfers werden bekomen. Een wanverhouding tussen de operationele definitie en de gebruikersbehoeften bepaalt de kwaliteit van de statistiekreeks.

Met Statistiek Vlaanderen leggen we op voorhand vast welke statistiekreeksen we publiceren. Deze statistiekreeksen zijn terug te vinden in het Vlaams Statistisch Programma (VSP-lijst). Voor elke statistiek bevat de VSP-lijst ten minste volgende informatie:

1. de conceptuele definitie (die gebruikt wordt als handig label voor de statistiekreeks),
2. het versienummer
3. de operationele definitie inclusief een beschrijving van alle concepten (attributen en dimensies) inclusief de exacte waarden van die concepten,
4. een overzicht van de gebruikersbehoeften, en
5. een kwaliteitsevaluatie van de operationele definitie in functie van de gebruikersbehoeften.

Tabel 8 toont hoe de VSP-lijst er minimaal zou moeten uitzien. In de praktijk zullen we uiteraard informatie over statistiekreeksen niet bewaren in één gigantische tabel maar werken met informatiefiches per statistiekreeks gebaseerd op de single integrated metadata structure (SIMS)².

3 ONTSLUITING STATISTIEKEN

We ontsluiten statistieken op twee manieren:

1. via volledige statistiekreeksen
2. via samenvattende cijferpagina's

Beide manieren worden hieronder besproken.

3.1 STATISTIEKREEKSEN

Een statistiekreeks is in essentie een dataset of datatabel die eenvoudig openbaar gepubliceerd kan worden. Het is mogelijk dat sommige cijfers in zo'n dataset een ontbrekende waarde hebben, bijvoorbeeld vanwege ontbrekende informatie of als gevolg van versluiering.

De structuur van een dataset met een statistiekreeks wordt bepaald volgens de SDMX-standaard. Deze standaard stelt dat een dataset met een statistiekreeks drie soorten informatie moet bevatten:

- Measures: Een kolom met de cijfers of statistieken. (In een toekomstige versie van SDMX kan een dataset meerdere kolommen met measures bevatten, waarbij het onderscheid tussen deze kolommen wordt vastgelegd als één van de dimensies).
- Dimensies: Kolommen met kenmerken van de cijfers die nodig zijn om elk cijfer uniek te identificeren.
- Attributen: Aanvullende informatie over de cijfers. Deze kan betrekking hebben op de gehele dataset, op een groep van cijfers of op één enkel cijfer.

De combinatie van alle dimensies en attributen bepaalt altijd de volledige operationele definitie van de cijfers. Daarnaast heeft elke dataset een heldere en beknopte conceptuele definitie, oftewel een titel.

Het overzicht van de measures, dimensies en attributen wordt vastgelegd in de Data Structuur Definitie (DSD). De DSD bevat ook een verwijzing naar de codelijst van elke measure, dimensie en attribuut. Elke statistiekreeks heeft precies één DSD.

² ec.europa.eu/eurostat/web/metadata/reference-metadata-reporting-standards

Tabel 8: Het Vlaams Statistisch Programma (VSP) definieert een lijst van statistiekreeksen die Statistiek Vlaanderen publiceert. Elke statistiekreeks in deze reeks beschikt over een duidelijke conceptuele en operationele definitie, opgelijste gebruikersbehoeften en een kwaliteitsevaluatie.

<p>Aantal mannen en vrouwen, versie 1</p> <p><i>gebruikersvoorwaarden:</i></p> <ul style="list-style-type: none"> → Decreet XX.XX verwijst naar het aantal inwoners in Vlaamse gemeenten om beleid te voeren over ... → Agentschap Binnenlands Bestuur publiceert cijfers over het aantal inwoners in gemeenten als ... → Academische onderzoekers vragen cijfers over het aantal inwoners in gemeenten om onderzoek te voeren over ... → ... <p><i>Operationele definitie:</i> Het aantal en het percentage mannen en vrouwen in de feitelijke bevolking (aantal geregistreerde inwoners in het Rijksregister inclusief personen in het wachtregister en ambassadeurs) op 1 januari van elk kalenderjaar vanaf 2005 per Vlaamse gemeente volgens de NIS-code-indeling in 2019 = 600 cijfers per jaar.</p> <p><i>Kwaliteit:</i> De gebruikers hebben eerder nood aan de grootte van de verblijvende bevolking in plaats van de feitelijke bevolking.</p> <p><i>Nota:</i> Stopgezet in 2022, wegens aanpassing definitie.</p>
<p>Aantal mannen en vrouwen, versie 2</p> <p><i>gebruikersvoorwaarden:</i></p> <ul style="list-style-type: none"> → Decreet XX.XX verwijst naar het aantal inwoners in Vlaamse gemeenten om beleid te voeren over ... → Agentschap Binnenlands Bestuur publiceert cijfers over het aantal inwoners in gemeenten als ... → Academische onderzoekers vragen cijfers over het aantal inwoners in gemeenten om onderzoek te voeren over ... → ... <p><i>Operationele definitie:</i> Het aantal en het percentage mannen en vrouwen in de verblijvende bevolking (aantal geregistreerde inwoners in het Rijksregister inclusief personen in het wachtregister en ambassadeurs en personen die minder dan drie maanden in België verblijven) op 1 januari van elk kalenderjaar vanaf 2005 per Vlaamse gemeente volgens de NIS-code-indeling in 2019 = 600 cijfers per jaar.</p> <p><i>Kwaliteit:</i> Onderzoek XX toont aan dat er geen problemen zijn met deze statistiekreeks.</p> <p><i>Nota:</i> Stopgezet in 2025, wegens aanpassing gemeenten door fusies.</p>
<p>Aantal mannen en vrouwen, versie 3</p> <p><i>gebruikersvoorwaarden:</i></p> <ul style="list-style-type: none"> → Decreet XX.XX verwijst naar het aantal inwoners in Vlaamse gemeenten om beleid te voeren over ... → Agentschap Binnenlands Bestuur publiceert cijfers over het aantal inwoners in gemeenten als ... → Academische onderzoekers vragen cijfers over het aantal inwoners in gemeenten om onderzoek te voeren over ... → ... <p><i>Operationele definitie:</i> Het aantal en het percentage mannen en vrouwen in de verblijvende bevolking (aantal geregistreerde inwoners in het Rijksregister inclusief personen in het wachtregister en ambassadeurs en personen die minder dan drie maanden in België verblijven) op 1 januari van elk kalenderjaar vanaf 2005 per Vlaamse gemeente volgens de NIS-code-indeling in 2025 = 570 cijfers per jaar.</p> <p><i>Kwaliteit:</i> Onderzoek XX toont aan dat er geen problemen zijn met deze statistiekreeks.</p>
<p>Tewerkstelling in hoogtechnologische sector, versie 1</p> <p><i>gebruikersvoorwaarden:</i></p> <ul style="list-style-type: none"> → Decreet XX.XX verwijst naar Tewerkstelling in hoogtechnologische sector in kader van ... → De cijferpagina met cijfers over tewerkstelling in hoogtechnologische sector wordt X aantal keer per jaar geraadpleegd. → ... <p><i>Operationele definitie:</i> Percentage van de hele werkende bevolking aan de slag in de hoogtechnologische sector op 1 januari van elk kalenderjaar vanaf 2005 per Vlaamse gemeente en voor het hele Vlaamse gewest volgens de NIS-code-indeling in 2019. De werkende bevolking omhelst ...De hoogtechnologische sector bestaat uit- bedrijven ...De data worden verzameld via de Enquête naar de Arbeidskrachten (EAK) door Statbel. In deze enquête wordt data verzameld door ...= 301 cijfers per jaar.</p> <p><i>Kwaliteit:</i> Onderzoek YY toont aan dat er geen problemen zijn met deze statistiekreeks.</p>
<p>Drinkwaterkwaliteit, versie 1</p> <p><i>gebruikersvoorwaarden:</i></p> <ul style="list-style-type: none"> → Decreet XX.XX verwijst naar drinkwaterkwaliteit in kader van ... → ... <p><i>Operationele definitie:</i> Conformiteitspercentage van het kraantjeswater in heel Vlaanderen per jaar. Het conformiteitspercentage wordt berekend door Vlaamse Milieumaatschappij (VMM) op basis van het totale aantal analyses en het totale aantal vastgestelde normoverschrijdingen voor volgende parameters: ...= 1 cijfer per jaar.</p> <p><i>Kwaliteit:</i> Volgens rapport ZZ ontstaan er kleine onzekerheidsfouten in de meting van parameter x, waardoor het werkelijke conformiteitspercentage kan afwijken met 0,2%. Deze fluctuatie heeft slechts beperkte invloed op de kwaliteit waardoor geen herdefiniëring nodig is.</p>
<p>....., versie ...</p> <p><i>gebruikersvoorwaarden:</i> ...</p> <p><i>Operationele definitie:</i> ...</p> <p><i>Kwaliteit:</i> ...</p> <p>.....</p>

Bovendien moet elke statistiekreeks volledig vrij raadpleegbaar zijn. Ook oude versies van een statistiekreeks blijven beschikbaar. Deze datasets kunnen op verschillende manieren gepubliceerd worden, bijvoorbeeld:

- als een eenvoudige downloadbare dataset of een reeks van datasets (bijvoorbeeld in csv-bestandsformaat), of
- via een interactieve cijferapplicatie waarin gebruikers selecties kunnen maken op basis van dimensies en alleen relevante delen van een statistiekreeks kunnen visualiseren en downloaden.

3.2 CIJFERPAGINA'S

Cijferpagina's zijn webpagina's waarop specifieke statistieken in de schijnwerpers worden gezet. Het doel van een cijferpagina is niet het publiceren van volledige statistiekreeksen, maar het presenteren van een toegankelijk en boeiend narratief rond enkele opvallende cijfers uit deze reeksen. Bij het opstellen van een cijferpagina wordt er dus selectief te werk gegaan, om zo een helder en overzichtelijk verhaal te creëren dat aantrekkelijk is voor een breed publiek. Vanwege deze opzet kunnen niet alle beschikbare statistieken op een cijferpagina worden besproken of gevisualiseerd. Dit zou de pagina immers onnodig complex en minder gebruiksvriendelijk maken.

Cijferpagina's bevatten vaak verschillende tabellen die een selectie van cijfers uit de statistiekreeksen presenteren. De structuur van deze tabellen is echter zelden uniform. Zo kan de eerste tabel bijvoorbeeld een tijdreeks met totaalcijfers tonen, waarbij de dimensie het concept 'tijdsperiode' is. Een tweede tabel kan daarentegen cijfers weergeven voor slechts één enkel jaar, uitgesplitst naar verschillende bevolkingsgroepen. In dat geval is er gefilterd op één tijdsperiode en vormt 'bevolkingsgroep' de dimensie. Deze variatie in tabelopmaak maakt duidelijk dat cijferpagina's niet bedoeld zijn als basis voor het vaststellen van volledige statistiekreeksen.

De inhoud van een cijferpagina wordt samengesteld door een team dat verantwoordelijk is voor het toegankelijk maken van onze statistieken. Dit proces staat los van de totstandkoming van de statistiekreeksen zelf, wat ervoor zorgt dat de focus blijft liggen op het creëren van begrijpelijke en informatieve verhalen voor een breed publiek.

Ideeën voor verdere ontwikkeling van deze handleiding

- Handleiding over hoe je een duidelijke operationele definitie schrijft, opgedeeld in de verschillende concepten?
- Diepgaandere uitwerking van definitie gebruikersbehoeften?
- Handleiding SDMX?
- Handleiding SIMS/GSBPM/...?
- Handleiding over hoe kernstatistieken worden aangeleverd aan de VSA (als een gefilterde statistiekreeks).
- ...?