



WAT IS EEN STATISTIEK?

Jorre Vannieuwenhuyze



26 november 2025



1	Statistiek als een cijfer	4
1.1	Conceptuele definitie	5
1.2	Gebruikersbehoeften	6
1.3	Operationele definitie	7
1.4	Kwaliteit	11
1.5	Andere administratieve attributen	17
2	Statistiek als een cijfertabel	18
2.1	Concepten & dimensies	19
2.2	Conceptuele en operationele definitie	19
2.3	Soorten attributen	21
3	Statistiek als een tabelreeks	22
4	Conclusie	25
	Bibliografie	27

Voor we bespreken hoe een openbare statistiekdienst zoals Statistiek Vlaanderen statistieken produceert, moeten we eerst duidelijk afbakenen wat we precies verstaan onder een statistiek. Binnen het Vlaams statistieklandschap wordt de term “statistiek” immers op verschillende manieren gebruikt en dat zorgt soms voor verwarring. Grosso modo kunnen we drie niveau’s onderscheiden:

1. Op het eerste niveau wordt de term “statistiek” gebruikt om te verwijzen naar individuele cijfers. Een blik op individuele cijfers is heel nuttig voor de opmaak van goede documentatie. Voor elk afzonderlijk gepubliceerd cijfer wordt vanuit Europa immers duidelijke en ondubbelzinnige documentatie verwacht. Die documentatie vertelt gebruikers waarom en hoe het cijfer werd berekend, hoe het cijfer moet worden geïnterpreteerd, en wat mogelijke valkuilen zijn bij gebruik van het cijfer.
2. Op het tweede niveau wordt de term “statistiek” gebruikt om te verwijzen naar cijfertabellen. Cijfertabellen vormen de basis van gestructureerde databases en zo’n databases zijn handige instrumenten om statistieken te beheren en te ontsluiten in een statistiekproductieproces. Voor sommige entiteiten binnen Statistiek Vlaanderen vormen cijfertabellen nu al de ruggengraat van hun statistiekproductieproces. Het Agentschap Binnenlands Bestuur (ABB), bijvoorbeeld, definieert zijn statistieken voor de Gemeente- en Stadsmonitor aan de hand van cijfertabellen.
3. Op het derde niveau wordt de term “statistiek” gebruikt om te verwijzen naar reeksen van cijfertabellen. Zo’n tabelreeksen maken het mogelijk om statistieken thematisch te ordenen en efficiënter te beheren. De lijst van Vlaamse Openbare Statistieken (VOS’en) in het Vlaams Statistisch Programma (VSP) is een voorbeeld van tabelreeksen.

Een duidelijk onderscheid tussen deze drie niveaus om de term “statistiek” in te vullen, is cruciaal. Het bevordert niet alleen de transparante en efficiënte informatie-uitwisseling tussen departementen en diensten, maar ook tussen collega’s die betrokken zijn bij de productie van statistieken. In dit hoofdstuk gaan we daarom dieper in op deze verschillende niveaus waarop een statistiek kan worden gedefinieerd. Die verkenning maakt het bovendien ook mogelijk om belangrijke basisvragen te beantwoorden rond statistiekproductie: wat beschouwen we als een statistiek en wat niet? Welke informatie is noodzakelijk voor de productie en publicatie ervan? En hoe beoordelen we de kwaliteit van statistieken?

1. STATISTIEK ALS EEN CIJFER

Als je mensen vraagt wat ze verstaan onder een statistiek, zullen velen dit op de eerste plaats definiëren als een cijfer. Die definitie is echter niet volledig. Neem bijvoorbeeld het cijfer “525 935”, een statistiek die we publiceren vanuit Statistiek Vlaanderen. Op zichzelf zegt dit cijfer uiteraard zeer weinig. Waarvan zijn er 525 935? Wat stelt dit getal voor? Waarom publiceren we dat getal? Hoe werd het getal gemeten? Hoe kwalitatief is deze meting? Wie is verantwoordelijk voor dit cijfer?

Statistiek gaat dus niet alleen over het berekenen van cijfers, ook wel **data** genoemd. Statistiek gaat ook over het beheer van allerhande informatie die bij de cijfers hoort. Deze informatie noemen we de **attributen** van de cijfers of **metadata**. Een **statistiek** bestaat dus altijd uit twee onlosmakelijke elementen: het cijfer zelf én de attributen (zie Figuur 1). Beiden zijn even belangrijk en vormen één geheel (Wild, Utts en Horton 2018). In deze paragraaf gaan we dieper in op verschillende belangrijke attributen van een statistiek.

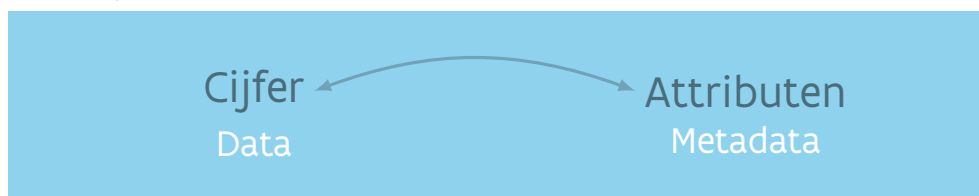


Merk op dat statistieken niet altijd cijfers zijn, het kunnen ook kwalitatieve observaties zijn of zelfs complexere “informatie-objecten”. Zo publiceert het Agentschap Binnenlands Bestuur bijvoorbeeld per gemeente een statistiek over de top 10 andere gemeenten waarnaar kinderen heen pendelen voor school. Voor de stad Antwerpen kan de observatie op deze statistiek bijvoorbeeld de volgende tekstuele vector zijn:

“[Brasschaat, Lier, Brussel, Beveren, Mechelen, ...]”.

Voor de eenvoud verwijzen we in deze tekst echter enkel naar statistieken als eenvoudige cijfers. Cijfers zijn immers veruit de meest voorkomende vorm van statistieken.

Figuur 1: Op het laagste niveau bestaat een statistiek steeds uit een cijfer en attributen van dat cijfer.



1.1. CONCEPTUELE DEFINITIE

Het eerste belangrijke attribuut van een statistiek is uiteraard de betekenis van het cijfer. Die betekenis wordt vormgegeven door de **conceptuele definitie** van de statistiek (Neuman 2014). Een conceptuele definitie is een beschrijving van wat de statistiek als een begrip of concept inhoudt, op een abstract en theoretisch niveau. De conceptuele definitie van het cijfer 525 935 is bijvoorbeeld “het aantal inwoners van de gemeente Antwerpen in 2019”.

Een conceptuele definitie bevat doorgaans verschillende delen of componenten. Bij de statistiek over het aantal inwoners in Antwerpen in 2019 kunnen we bijvoorbeeld drie componenten onderscheiden. Ten eerste kijken we naar een gemeente, ten tweede kijken we naar een bepaald jaar, en ten derde kijken we naar een parameter die we willen meten. Deze componenten noemen we **concepten** (Clark e.a. 2021). Een concept is dus niet meer dan een idee of betekenis die helpt om statistische gegevens te begrijpen en te ordenen (UNECE 2024). Het is als een label dat zegt “Dit is wat we meten”.

Een concept op zich is echter steeds slechts een containerbegrip. In de praktijk neemt een concept steeds een bepaalde **waarde** aan om de betekenis van een statistiek te verduidelijken. In het voorbeeld neemt het concept “gemeente” de waarde “Antwerpen” aan, het concept “jaartal” de waarde ‘2019’, en het concept “parameter” de waarde “aantal inwoners”. De waarden geven dus betekenis aan het cijfer, niet het concept zelf.

Let op, de afbakening van concepten en waarden in een conceptuele definitie is niet altijd duidelijk en is voor een stukje arbitrair. Zo kan je bijvoorbeeld cijfers publiceren over het aantal inwoners in Antwerpen per bevolkingsgroep (mannen, vrouwen, 0- tot 18-jarigen, 19- tot 64- jarigen, 65+’ers, mensen met Belgische nationaliteit, mensen met buitenlandse nationaliteit, ...). Spreken we in deze definitie van één concept “bevolkingsgroep” of hebben we hier te maken met meerdere concepten zoals “geslacht”, “leeftijd” en “nationaliteit”? Concepten kan je ook arbitrair opsplitsen en samenvoegen. Hebben we in het voorbeeld hierboven slechts één concept “parameter” met waarde “aantal inwoners”, of kan je hier niet beter twee aparte concepten onderscheiden, namelijk het concept “statistische parameter” met waarde “aantal” en het concept “eenheid” met waarde “inwoners”? Op deze vragen is er geen éénduidig juist of fout antwoord maar we zullen wel een keuze moeten maken om het productieproces voor deze statistieken in de praktijk te organiseren. Dit wordt verder uitgelegd in hoofdstuk ??.

1.2. GEBRUIKERSBEHOEFTE

Wanneer een openbare statistiekdienst een statistiek publiceert met een bepaalde betekenis of conceptuele definitie, gebeurt dat steeds als antwoord op specifieke **gebruikersbehoeften**. Die behoeften vormen in feite het vertrekpunt van het hele statistiekproductieproces (UNECE 2025). Niet de conceptuele definitie bepaalt dus welke statistiek wordt ontwikkeld, maar de gebruikersbehoeften geven daartoe de aanzet. Zij bepalen bovendien niet alleen welke statistieken de openbare statistiekdienst moet produceren en publiceren, maar ook hoe dat op de meest kwaliteitsvolle manier kan gebeuren. De gebruikersbehoeften waaraan een statistiek beantwoordt, zijn daarom het tweede essentiële attribuut van een statistiek.

Als we het aantal inwoners van de gemeente Antwerpen in 2019 vanuit Statistiek Vlaanderen publiceren, betekent dit dat sommige beleidsmakers, onderzoekers of burgers daadwerkelijk graag willen weten hoeveel inwoners Antwerpen telde in 2019. Een Vlaams decreet kan bijvoorbeeld verwijzen naar het aantal inwoners in de Vlaamse gemeenten en verplicht ons dit aantal te publiceren zodat de Vlaamse regering hierop beleid kan uitstippelen. Antwerpse beleidsmakers weten daarnaast waarschijnlijk ook graag hoeveel inwoners hun gemeente telt zodat ze hun diensten hierop kunnen afstemmen. Wetenschappelijke onderzoekers gebruiken de Antwerpse bevolkingsgrootte in 2019, samen met gegevens uit andere jaren en gemeenten, dan weer om bevolkingsgroei en spreiding in Vlaanderen te onderzoeken. Burgers willen misschien wel graag weten hoeveel inwoners Antwerpen telt om te beoordelen of bepaalde Vlaamse subsidies aan de stad gerechtvaardigd zijn of niet.

Over het algemeen kunnen gebruikersbehoeften zich op verschillende manieren manifesteren (Eurostat 2021):

1. Er is een **internationale verplichting** om een statistiek te produceren.
2. Er is een **institutionele** vraag naar statistieken vervat in wetten of decreten.
3. Er is een **beleidsondersteunende** vraag naar statistieken vanuit de regering, kabinetten of overheidsadministraties.
4. Er is een **maatschappelijke** vraag naar statistieken omwille van een maatschappelijk debat.
5. Er is een **academische** vraag naar statistieken in functie van wetenschappelijk onderzoek.

In hoofdstuk ?? wordt meer uitleg gegeven hoe je deze behoeften verder kan concretiseren in de metadata.

1.3. OPERATIONELE DEFINITIE

De conceptuele definitie is meestal vaag over wat een cijfer precies weer-geeft. In de zin “het aantal inwoners in Antwerpen neemt toe” is bijvoorbeeld “het aantal inwoners” enkel een abstract idee. Deze woorden vertellen immers niet wie je nu precies meetelt en wie niet. De conceptuele definitie van een cijfer vertelt dus weinig over hoe het cijfer precies werd berekend en exact geïnterpreteerd kan worden. Naast de conceptuele definitie heeft elke statistiek daarom ook een **operationele definitie**. Die 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 veel uitgebreider dan de conceptuele definitie. De vertaling van de conceptuele definitie naar de operationele definitie noemen we **operationaliseren**.

De operationalisering van een conceptuele definitie betekent in de praktijk dat elk concept vertaald wordt van een abstract idee naar een concrete richtlijn om de conceptwaarden te meten. Concepten variëren echter in hoe concreet of abstract ze zijn (Neuman 2014). Een **concreet concept** verwijst naar een fenomeen dat direct observeerbaar of eenvoudig meetbaar is zonder complexe interpretatie of afleiding. Een voorbeeld van een concreet concept is het totale stroomverbruik in een gemeente. Dit is een concreet concept want je kan het rechtstreeks vaststellen door het verbruik af te lezen op alle elektriciteitsmeters binnen de grenzen van deze gemeente. Er is ook weinig discussie over hoe je stroomverbruik meet, namelijk gewoon in kilowattuur (kWh).

Een **abstract concept** verwijst daarentegen naar een idee of fenomeen dat niet rechtstreeks waarneembaar is. Het heeft een zekere mate van theoretische of interpretatieve lading en moet daarom eerst geoperationaliseerd worden naar concretere concepten en conceptwaarden. Een voorbeeld van een abstract concept is sociaal isolement. Sociaal isolement verwijst immers naar een toestand waarin iemand weinig of geen sociale contacten heeft of zich sociaal buitengesloten voelt. Dit kan je niet rechtstreeks meten maar moet je operationaliseren. Dat kan bijvoorbeeld via een lange vragenlijst over, onder andere, het aantal sociale contacten, subjectieve gevoelens van eenzaamheid en deelname aan sociale activiteiten.

In de praktijk zijn concepten niet of concreet of abstract, maar kan je ze op een continue schaal plaatsen op het vlak van abstractheid en concreetheid (Neuman 2014). Het éne concept is al abstracter en moeilijker meetbaar en operationaliseerbaar terwijl een ander concept redelijk concreet is en vrij direct geoperationaliseerd en geobserveerd kan worden.

Neem opnieuw het cijfer 525 935 als voorbeeld om het aantal inwoners van Antwerpen in 2019 weer te geven. De volledige operationele definitie van dit cijfer luidt:

“De grootte van de wettelijke bevolking op 1 januari 2019, 0.00 uur, van de gemeente Antwerpen (NIS 11002 in 2019). De grootte van de wettelijke bevolking wordt gedefinieerd op basis van het Rijksregister van de natuurlijke personen. Dit Rijksregister bevat, onder andere, het bevolkingsregister en het vreemdelingenregister die de wettelijke bevolking bepalen. Het bevolkingsregister bevat alle Belgen en buitenlanders die gemachtigd zijn tot vestiging op het Belgisch grondgebied, en het vreemdelingenregister bevat alle 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 (bvb. diplomatiek en consulaire personeel) zijn vrijgesteld van inschrijving in de bevolkingsregisters en worden daardoor niet meegerekend bij de wettelijke bevolking. 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 een wachtregister voor asielzoekers en een wachtregister voor EU burgers. Het wachtregister voor asielzoekers bevat alle verzoekers om internationale bescherming die worden ingeschreven door de Dienst Vreemdelingenzaken (DVZ). In 1995 besliste Statbel de personen in dit wachtregister niet meer mee te tellen bij de wettelijke bevolking. Pas nadat asielzoekers worden overgeschreven van het wachtregister naar het bevolkingsregister of het vreemdelingenregister, worden zij opgenomen in de bevolkingsstatistieken van Statbel. Zo’n overschrijving naar het bevolkingsregister of het vreemdelingenregister gebeurt na erkenning als vluchteling, na toekenning van een statuut subsidiaire bescherming, of na verwerving van een verblijfsvergunning om een andere reden.

Verder bevat het Rijksregister ook een wachtregister voor EU burgers in afwachting van woonstcontrole. Deze personen worden evenmin meegeteld bij de wettelijke bevolking. Pas na woonstcontrole worden deze personen overgeschreven naar het vreemdelingenregister en worden zij meegeteld in de wettelijke bevolking.”

Deze operationele definitie bevat een duidelijke operationalisering van elk concept afzonderlijk (zie Tabel 2). Het concept “gemeente” wordt geoperationaliseerd als de gemeente volgens de NIS-code indeling in 2019. Het concept “jaar” wordt geoperationaliseerd als de toestand op 1 januari, 0.00

uur, van een kalenderjaar. Het concept “parameter” wordt geoperationa-
liseerd als het aantal personen in de wettelijke bevolking met bijhorende
precieze definitie van wie meegerekend wordt in de wettelijke bevolking
en wie niet.

Dit voorbeeld toont ook dat de concepten in een conceptuele definitie
doorgaans abstract zijn. Zelfs een eenvoudige concept zoals “parameter”
met waarde “aantal inwoners” bevat al een zekere abstractie. De “grootte
van de wettelijk bevolking” is duidelijk een concretere conceptwaarde en
kan worden gebruikt om de abstractere conceptwaarde “aantal inwoners”
te meten.

Verder illustreert dit voorbeeld dat zelfs relatief eenvoudige statistieken
al een tamelijk lange operationele definitie vereisen. De lengte van een
operationele definitie kan dan ook sterk variëren. Voor een eenvoudig
cijfer afgeleid uit officiële registers, zoals het aantal inwoners in Vlaamse
gemeenten, kunnen enkele paragrafen volstaan. Bij complexere statistie-
ken, zoals bijvoorbeeld de gemiddelde opinie van inwoners gemeten via
een bevraging, is daarentegen een uitgebreide rapportering nodig met alle
details over, onder andere, het enquêtedesign, de steekproeftrekking, het
ontwerp van de vragenlijst, de opvolging van respondenten, statistische
correcties voor meet- en selectiefouten en de statistische analysemetho-
den.

Als laatste illustreert dit voorbeeld ook dat tijdens de vertaalslag van con-
ceptuele naar operationele definitie keuzes moeten worden gemaakt. Die
keuzes kunnen ervoor zorgen dat andere cijfers worden bekomen. Het
cijfer 525 935 verwijst bijvoorbeeld naar het jaar 2019, maar wordt wel
enkel gemeten op 1 januari 2019. Een vergelijkbaar maar ander cijfer kon
worden berekend als “jaar” op een andere wijze werd geoperationaliseerd,
bijvoorbeeld door een andere dag in het jaar te kiezen in plaats van 1 janu-
ari of door het jaargemiddelde te berekenen over alle dagen. Bovendien
verwijst 525 935 naar de wettelijke bevolking gerapporteerd door Stat-
bel en in die wettelijke bevolking worden personen uit de wachtregisters

Tabel 2: Operationaliseren betekent dat alle abstracte concepten in de conceptuele de-
finitie worden vertaald naar concrete en observeerbare concepten in de operationele
definitie

	Concept 1	Concept 2	Concept 3
conceptuele definitie	aantal inwoners	in Antwerpen	in 2019
	↓	↓	↓
operationele definitie	aantal inwoners volgens de wettelijke bevolking	in Antwerpen volgens NIS-code 11002 in 2019	op 1 januari 2019, 0.00 uur

niet meegeteld. Eurostat, daarentegen, rapporteert niet de wettelijke bevolking maar de “gewoonlijk verblijvende bevolking” inclusief personen in het wachtregister en publiceert daarom andere cijfers dan Stabel en Statistiek Vlaanderen, ook al interpreteren de meeste gebruikers zowel de Europese als de Belgische en Vlaamse cijfers gewoon als de “bevolkingsgrootte”. 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. We zouden opnieuw een vergelijkbaar maar ander cijfer kunnen berekenen voor alle personen in 2019 die op dat moment woonden in wat we nu als de gemeente Antwerpen beschouwen, inclusief Borsbeek dus.

Een operationele definitie heeft dus twee doelen. Het eerste doel is om duidelijkheid te scheppen over de gemaakte keuzes om een statistiek te berekenen. Hierdoor kan iedereen de cijfers reproduceren en creëer je transparantie. Als twee onderzoekers met dezelfde operationele definitie verschillende cijfers bekomen, is de definitie onvolledig of onnauwkeurig omdat bepaalde concepten nog te abstract zijn.



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 idee 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 handleiding.

Het tweede doel van de operationele definitie is duidelijkheid te scheppen over de vergelijkbaarheid van cijfers. Je kan bijvoorbeeld niet zomaar de grootte van de wettelijke Antwerpse bevolking in 2019 vergelijken met de grootte van de gewoonlijk verblijvende Antwerpse bevolking in 2020. Dit zou immers kunnen leiden tot onjuiste conclusies over bevolkingsgroei of -krimp, ook al worden beide cijfers conceptueel geïnterpreteerd als de Antwerpse “bevolkingsgrootte”.

Je zou je nu kunnen afvragen waarom we eigenlijk conceptuele definities opstellen als ze, in tegenstelling tot de operationele definitie, niet concreet en voldoende accuraat zijn? Toch zijn er goede redenen om voor elke statistiek een goede conceptuele definitie te bepalen. Ten eerste is de conceptuele definitie een pak handiger in gebruik dan de operationele definitie om cijfers te rapporteren. Statistiekgebruikers beperken zich doorgaans

tot de conceptuele definitie om cijfers te benoemen en te interpreteren, ook al zijn er binnen dezelfde conceptuele definitie meerdere operationele definities mogelijk. Iemand die praat over het aantal inwoners in Antwerpen in 2019 zal dit eenvoudigweg de “Antwerpse bevolkingsgrootte” noemen en niet de hele operationele definitie afratelen. Ten tweede hebben de nuances in de operationele definitie vaak weinig invloed op beleidskeuzes of onderzoeksconclusies die gebruikers maken op basis van de cijfers. Of je nu de wettelijke of de gewoonlijk verblijvende bevolking gebruikt om de bevolkingssomvang in Antwerpen te meten, en of je dit nu doet aan het begin of in het midden van het kalenderjaar, de meeste gebruikers zullen min of meer dezelfde conclusies trekken over bevolkingsgroei als ze deze cijfers over de jaren heen vergelijken.

1.4. KWALITEIT

We zagen reeds dat een statistiek ontstaat vanuit gebruikersbehoeften en dat deze behoeften worden vertaald naar een conceptuele definitie en vervolgens een operationele definitie (zie Figuur 3). In een ideale situatie zijn de conceptuele en operationele definitie afgestemd op de gebruikersbehoeften. Wanneer dat niet het geval is, heeft de statistiek een **kwali-teitsprobleem**.

Kwaliteitsproblemen bij statistieken kunnen voorkomen in verschillende vormen. Een overzicht hiervan wordt gegeven in de “Praktijkcode voor Europese statistieken” (zie Eurostat 2017). Deze praktijkcode schept een kwaliteitskader voor het Europees statistisch systeem (ESS) en legt 16 beginselen vast voor de ontwikkeling, productie en verspreiding van Europese statistieken. Vijf van deze beginselen beschrijven kwaliteitscriteria voor statistische output en worden hieronder verder besproken.

Relevantie

Volgens de praktijkcode moeten openbare statistiekdiensten op de eerste plaats statistieken produceren die **relevant** zijn. Relevantie draait steeds om de vraag of gebruikersbehoeften inhoudelijk vervuld worden.

Figuur 3: De kwaliteit van een statistiek wordt gedefinieerd door de verhouding tussen drie elementen.



Er is een probleem op het vlak van relevantie als de conceptueel gedefinieerde statistieken van een openbare statistiekdienst inhoudelijk niet in overeenstemming zijn met de gebruikersbehoeften. Dit probleem kan optreden in twee richtingen. Enerzijds kan er behoefte zijn aan bepaalde statistieken maar produceert de openbare statistiekdienst ze niet. Dit zou het geval zijn mocht Statistiek Vlaanderen bijvoorbeeld geen bevolkingscijfers publiceren terwijl daar wel een maatschappelijke vraag naar is. Anderzijds kan de openbare statistiekdienst ook statistieken produceren waar helemaal geen behoefte aan is. Ook in deze laatste situatie kan de kwaliteit van de statistiekproductie verbeterd worden. De openbare statistiekdienst verliest dan immers tijd en energie aan irrelevante statistieken terwijl ze die tijd en energie kon inzetten op een andere manier. In beide situaties vervult de openbare statistiekdienst haar rol als publieke en onafhankelijke dienstverlener onvoldoende.

Een relevantieprobleem kan ook ontstaan als er een wanverhouding bestaat tussen gebruikersbehoeften en de operationele definitie van een statistiek. In ons Antwerps voorbeeld kunnen gebruikers bijvoorbeeld het cijfer 525 935 systematisch interpreteren als de grootte van de gewoonlijk verblijvende Antwerpse bevolking in plaats van de wettelijke bevolking omdat ze gewoonweg beleidsmatig voornamelijk de grootte van de gewoonlijk verblijvende bevolking nodig hebben. In dit soort situaties moet de statistiekdienst onderzoeken of de operationele definitie aangepast kan worden aan de behoeften. In ons hypothetisch voorbeeld is het misschien aangeraden dat Statistiek Vlaanderen voortaan Eurostat volgt en de grootte van de gewoonlijk verblijvende bevolking publiceert in plaats van de wettelijke bevolking.

In vele situaties is een evaluatie van de operationele definitie in functie van gebruikersbehoeften echter geen zwart-wit-verhaal. Er kunnen zich immers situaties voordoen waarin de gebruikte operationele definitie voldoet voor de ene gebruiker maar niet voor een andere gebruiker. Zo kan de statistiek over bevolkingsgrootte op basis van de wettelijke bevolking voldoende zijn voor heel wat gebruikers, behalve voor een specifieke beleidsmaker die het woonbeleid in Antwerpen moet bepalen. Deze beleidsmaker heeft misschien nood aan exacte cijfers over de gewoonlijk verblijvende bevolking in plaats van de wettelijke bevolking. Ook in zo'n geval is er een gebruikersbehoefte waar de statistiek "Aantal inwoners in Antwerpen" niet aan voldoet, ook al voldoet hij wel voor de behoeften van de meeste andere gebruikers.

In situaties met verschillende niet-overlappende gebruikersbehoeften kan een openbare statistiekdienst zoals Statistiek Vlaanderen besluiten niet langer één statistiek te publiceren maar verschillende statistieken. Zo kan worden besloten niet één statistiek te publiceren over "het aantal inwoners in Antwerpen", maar wel twee afzonderlijke statistieken: "de wet-

telijke bevolkingsgrootte van Antwerpen” en “de verblijvende bevolkingsgrootte van Antwerpen”. Er wordt dan voldaan aan de specifieke noden van alle gebruikers. Merk op dat in zo’n situaties zowel de conceptuele definities als de bijhorende operationele definities verder verfijnd en aangepast moeten worden. Uiteraard moeten dit soort keuzes steeds gebeuren vanuit praktische en pragmatische overwegingen in functie van de beschikbare middelen en het personeel. Er kan in bovenstaande situatie bijvoorbeeld ook besloten worden om niet te voldoen aan de noden van deze éne specifieke beleidsmaker omdat dit te veel inspanning zou vragen voor wat het oplevert.

Nauwkeurigheid

Op de tweede plaats moeten statistieken zo **nauwkeurig** mogelijk zijn. Er is een probleem op het vlak van **nauwkeurigheid** als een statistiek niet meet wat ze moet meten of, anders verwoord, de operationele definitie onvoldoende aansluit bij de conceptuele definitie. De nauwkeurigheid van statistieken hangt steeds af van de gebruikte methoden en technieken om data te verzamelen en te analyseren.

Zo produceert Statistiek Vlaanderen bijvoorbeeld statistieken over de energiescore van bestaande woningen. Administratieve registers bevatten echter enkel informatie over geregistreerde energiescores, meer bepaald op het moment dat een woning verkocht of verhuurd wordt. Er is dus, onder andere, geen informatie beschikbaar over energiescores van woningen die werden gerenoveerd na verkoop of voor woningen die al heel lange tijd niet werden doorverkocht. Dit kan er toe leiden dat de geproduceerde cijfers niet genoeg nauwkeurig zijn om een beeld te krijgen van de algemene energiezuinigheid van woningen in Vlaanderen.

In dit soort situaties heeft een openbare statistiekdienst twee taken te vervullen. Ten eerste communiceert ze zo transparant mogelijk over de manier waarop de cijfers werden verzameld en de mogelijke invloed van de beperkingen van de data en analysemethoden op de nauwkeurigheid. Indien mogelijk verfijnt ze ook de conceptuele definitie van de statistiek zodat het risico op verwarring verkleint. Zo kan ze in bovenstaand voorbeeld de conceptuele definitie aanpassen van “energiescore van bestaande woningen” naar “energiescores van geregistreerde woningen”. Uiteraard kan de aanpassing van de conceptuele definitie er toe leiden dat de statistiek haar relevantie verliest omdat de conceptuele definitie niet meer in lijn ligt met de gebruikersbehoeften. Beleidsmakers weten immers waarschijnlijk liefst hoe het zit met de energiescore van alle woningen in Vlaanderen, niet enkel diegenen die geregistreerd werden.

Ten tweede moet de statistiekdienst op zoek gaan naar bijkomende informatie, analysemethoden en modeleringstechnieken die de tekortko-

mingen van de beschikbare data helpen corrigeren. Indien ze geschikte methoden vindt om zo'n correcties uit te voeren, is het misschien mogelijk om de statistiek toch onder de oorspronkelijke conceptuele definitie te blijven produceren, in het voorbeeld hierboven de "energiescore van bestaande woningen". Uiteraard vertrekken analysemethoden vaak van niet-verifieerbare aannames die alsnog ervoor kunnen zorgen dat de geproduceerde cijfers niet volledig het beoogde concept accuraat weergeven. Het blijft dan belangrijk om volledig transparant te zijn over de gebruikte analysemethoden en hun beperkingen, en om de keuze voor de analysemethoden zo goed mogelijk te beargumenteren op basis van theoretisch en empirisch wetenschappelijke studies.



*Wanneer een statistiek wordt berekend op basis van een willekeurige steekproef of simulatie, kan je nauwkeurigheid verder opsplitsen in twee componenten, **geldigheid** en **betrouwbaarheid**. Geldigheid duidt aan of de statistiek gemiddeld genomen meet wat het moet meten. Als je een productieproces dus voortdurend herhaald, moet je gemiddeld gezien wel op het juiste cijfer komen. Betrouwbaarheid verwijst naar hoe groot de willekeurige schommelingen zijn in de statistiek. Als je het productieproces voortdurend herhaalt zullen de schattingen bij een betrouwbare statistiek steeds dicht tegen de gemiddelde waarde liggen, terwijl ze bij een onbetrouwbare statistiek daar sterk van kunnen afwijken. Merk op dat je een willekeurige steekproef heel ruim kan beschouwen. In een bevraging kunnen respondenten bijvoorbeeld ook van dag tot dag willekeurig andere antwoorden geven afhankelijk van de context en de gemoedstoestand waarin ze de vragenlijst invullen. Ook dit kan een oorzaak zijn van onbetrouwbare statistieken.*

Tijdigheid en Punctualiteit

Naast relevantie en nauwkeurigheid, somt de praktijkcode nog andere kwaliteitscriteria op. Volgens de praktijkcode horen statistieken ook **tijdig** en **punctueel** te worden gepubliceerd. Tijdig betekent dat de statistieken zo snel mogelijk worden gepubliceerd na de datum waarop de data werden verzameld. Punctueel betekent dat de statistieken op duidelijke tijdstippen worden gepubliceerd. Dit laat beleidsmakers toe om snelle beleidsbeslissingen te nemen of om beleidsbeslissingen te nemen gebonden aan duidelijke juridische deadlines.

Wanneer er problemen ontstaan rond tijdigheid en punctualiteit van de statistiek is dit doorgaans omwille van een verkeerd gekozen operationele definitie om de statistiek te meten. Zo kan een operationele definitie inhoudelijk aansluiten bij een gebruikersbehoefte, maar als de berekening

te lang duurt en de statistiek daardoor te laat beschikbaar is, verlaagt dit alsnog de kwaliteit.

Vergelijkbaarheid

De praktijkcode vermeldt ook **vergelijkbaarheid** van statistieken als kwaliteitscriterium. Dit betekent dat de gebruiker de statistieken kan vergelijken met eerder verschenen cijfers of met cijfers uit andere landen of regio's. Statistiekgebruikers hebben uiteraard geen boodschap aan nauwkeurig gemeten statistieken als deze helemaal niet vergelijkbaar zijn met andere gepubliceerde cijfers.

Opnieuw worden problemen in verband met vergelijkbaarheid veroorzaakt door slecht gekozen conceptuele en operationele definities. Hoe sterker de definities van twee cijfers van elkaar afwijken, hoe minder deze cijfers vergelijkbaar zijn. Merk op dat streven naar vergelijkbaarheid en streven naar nauwkeurigheid in sommige situaties tegengestelde doelen kunnen zijn. De nauwkeurigheid van een statistiek kan in bepaalde situaties bijvoorbeeld verbeterd worden door nieuwe dataverzamelings- en -verwerkingsmethoden. Zo'n nieuwe methoden kunnen echter tot gevolg hebben dat de statistiek niet meer vergelijkbaar is met eerder gepubliceerde cijfers. De statistiekdienst zal in zo'n situatie een gedragen evenwicht moeten vinden tussen nauwkeurigheid en vergelijkbaarheid.

Legaliteit

Als laatste moeten statistieken ook steeds **legaal** zijn. Dit betekent dat de statistieken worden gepubliceerd conform de wet en regelgeving en bijvoorbeeld geen inbreuk opleveren tegenover de privacywetgeving. De legaliteit van een statistiek is uiteraard een conditio sine qua non voordat een gebruiker de statistiek kan raadplegen. Statistieken die verzameld werden op illegale manieren kunnen per definitie niet ontsloten worden en hebben zo geen enkele waarde voor gebruikers.

De legaliteit van statistieken hangt opnieuw voor een groot deel af van de operationele definitie. Als deze definitie geen maatregelen incorporeert om gevoelige gegevens te versluieren of onterechte toegang tot data te vermijden, zijn de opgeleverde statistieken per definitie illegaal.



Dat cijfers legaal moeten zijn staat niet expliciet vermeld in de Europese praktijkcode als beginsel. Dit is een hiaat in de praktijkcode. Deze vereiste wordt wel impliciet vermeld via beginselen over een institutioneel kader rond statistische geheimhouding en gegevensbescherming. Dit beginsel dwingt echter geen beschrijving af van concreet genomen maatregelen om

een bepaalde dataset te versluieren en inbreuken tegen privacy te vermijden.

Kwaliteitsevaluatie

Concluderend is het dus belangrijk dat we als Statistiek Vlaanderen diepgaand onderzoek uitvoeren naar gebruikersbehoeften zodat we deze kunnen gebruiken voor [kwaliteitsevaluatie](#) (United Nations Statistics Division 2025). Wie heeft er nood aan statistieken? Wie gebruikt deze statistieken? Welke wet of welk decreet ondersteunt de productie van bepaalde statistieken? Voor welke beleidsbeslissingen of welk onderzoek zijn deze statistieken relevant? Daarnaast is er ook voortdurend onderzoek nodig naar de kwaliteit van statistieken. Zijn de statistieken nog steeds nauwkeurig en kosten-efficiënt? Kunnen we de statistieken voldoende tijdig opleveren? Kunnen we de operationele definitie niet verbeteren door gebruik te maken van nieuwe innovatieve dataverzamelings- en analysemethoden?

Hoewel onderzoek naar bovenstaande vragen niet altijd een exacte wetenschap is, vormt ze een belangrijke stap in de productie en publicatie van openbare statistieken. Voor zo'n onderzoek moeten we bovendien verder kijken dan de kwantitatieve methoden waarmee statistiek vaak geassocieerd worden. Binnen gebruikersonderzoek kan ook veel kennis worden vergaard door tekstuele analyse van, onder andere, wetteksten, beleidsdocumenten, persartikels, opiniestukken of audits van ondernemingen, of door kwalitatief onderzoek via diepte-interviews, focusgroepen, participerende observaties en case studies bij gebruikersgroepen. De literatuur rond evaluatie-onderzoek biedt hierbij een handige kapstok (zie bijvoorbeeld Rossi, Lipsey en Henry 2018; Mertens, Wilson en Hall 2025).



De verschillende kwaliteitscriteria voor statistieken staan in de "Praktijkcode voor Europese statistieken" opgesomd als beginselen rond statistische output. Naast de opgenoemde criteria, vermeldt de praktijkcode echter ook toegankelijkheid van de statistiek als kwaliteitscriterium. Dit is echter geen kenmerk van een statistiek zelf, maar wel van het systeem waarmee statistieken worden ontsloten.

Naast de beginselen rond statistische output, bevat de "Praktijkcode voor Europese statistieken" bovendien ook nog tien andere beginselen rond het institutioneel kader waarbinnen een openbare statistiekdienst functioneert en de statistische processen die een openbare statistiekdienst installeert. Deze "beginselen" kan je echter bekijken als hulpmiddelen om de beginselen rond statistische output waar te maken.

Zo vereist beginsel 1 dat openbare statistiekdiensten professioneel onafhankelijk opereren. Dit betekent dat er geen politieke of

andere externe invloeden zijn op het werk van de openbare statistiekdienst. Beginsel 6 is hier nauw mee verbonden en vraagt dat openbare statistiekdiensten onpartijdig en objectief zijn. Deze beginselen zorgen er dus voor dat openbare statistiekdiensten alle gebruikersbehoeften maatschappijbreed in kaart brengen, hun statistiekproductie hierop afstemmen en daardoor relevante en nauwkeurige statistieken opleveren. Relevante en nauwkeurige statistieken zijn dus de beginselen, onafhankelijkheid en onpartijdigheid zijn enkel hulpmiddelen om deze beginselen waar te maken.

Beginsel 2 bepaalt dat een openbare statistiekdienst een wettelijk mandaat krijgt voor de verzameling van en toegang tot gegevens. Beginsel 3 stelt op haar beurt dat openbare statistiekdiensten over voldoende personele, financiële en technische middelen moeten beschikken. Deze beginselen garanderen dat gebruikersbehoeften optimaal in kaart worden gebracht door de openbare statistiekdienst en dat aan deze behoeften voldaan wordt met relevante en nauwkeurige statistieken.

Beginsel 4 vraagt om een sterk kwaliteitsbewustzijn bij openbare statistiekdiensten. Dit beginsel hangt nauw samen met beginsels 7 en 8, die vereisen dat openbare statistiekdiensten deugdelijke methoden en geschikte statistische procedures gebruiken om kwaliteit voortdurend te optimaliseren. Deze beginselen zeggen niet meer en niet minder dan dat openbare statistiekdiensten moeten streven naar maximale nauwkeurigheid in hun statistieken en zijn dus volledig redundant.

1.5. ANDERE ADMINISTRATIEVE ATTRIBUTEN

We zagen al dat een statistiek bestaat uit een cijfer en een heleboel attributen zoals de gebruikersbehoeften, de conceptuele en operationele definitie, en een grondige kwaliteitsevaluatie op basis van verschillende criteria. Daarnaast kan een cijfer echter nog heel wat andere attributen hebben. Een belangrijk bijkomend attribuut is bijvoorbeeld de vertrouwelijkheid van een cijfer. Kan het cijfer worden gedeeld met het brede publiek of met collega's of is het vertrouwelijk en kan enkel wordend geraadpleegd door een beperkte groep van gemachtigde analisten? Hoewel de vertrouwelijkheid van een cijfer geen betekenis geeft aan het cijfer zelf, is het uiteraard belangrijk om dit soort informatie ook nauwgezet te registreren en te bewaren.

Een ander belangrijk attribuut is het eigenaarschap van een cijfer. Het eigenaarschap van een statistiek verwijst naar de verantwoordelijkheid en het mandaat dat een bepaalde instantie heeft over het verzamelen, produceren, publiceren en beheren van die statistiek. In de context van officiële statistieken houdt dit eigenaarschap ook de verantwoordelijkheid in om

kwaliteit en betrouwbaarheid te garanderen, een mandaat om het cijfer te publiceren, en de verantwoordelijkheid om definities, classificaties, bronnen en methoden te beheren die bij de statistiek horen. Het is belangrijk steeds duidelijkheid te hebben over het eigenaarschap van cijfers omdat dit coherentie en vertrouwen garandeert in openbare statistieken, dubbel werk en tegenstrijdige cijfers voorkomt tussen instellingen en duidelijk maakt wie aanspreekbaar is bij vragen of kritiek.

Samengevat, voor elke statistiek die we produceren berekenen we een cijfer maar registreren we ook alle relevante attributen die bij dat cijfer horen zoals:

1. de gebruikersbehoeften,
2. een conceptuele definitie op basis van de gebruikersbehoeften en met afbakening van alle individuele concepten,
3. een operationele definitie met de operationalisering van alle concepten,
4. informatie over de kwaliteit van het cijfer, en
5. bijkomende administratieve attributen.

In de volgende hoofdstukken bespreken we hoe we dat concreet in de praktijk aanpakken.

2. STATISTIEK ALS EEN CIJFERTABEL

In de vorige paragraaf werd uitgelegd hoe een statistiek kan verwijzen naar één enkel cijfer met bijhorende informatie-attributen waaronder de conceptuele en operationele definitie ontstaan vanuit specifieke gebruikersbehoeften. In de praktijk produceren we echter vaak verschillende cijfers met sterk overlappende gebruikersbehoeften en definities. Zo publiceren we uiteraard niet alleen het aantal inwoners voor de gemeente Antwerpen, maar ook voor alle andere Vlaamse gemeenten, de drie gewesten en heel België. Het is dan niet efficiënt om attributen cijfer per cijfer te beheren en te ontsluiten. Om efficiëntie te verhogen groeperen we daarom cijfers bij elkaar. Dat doen we via cijfertabellen. Als we praten over statistieken, kan dit dus ook verwijzen naar zo'n cijfertabel in plaats van individuele cijfers.

Een **cijfertabel** is een verzameling van cijfers die grotendeels dezelfde conceptuele en operationele definitie hebben. Meer specifiek bestaan de conceptuele definities van deze cijfers uit exact dezelfde concepten maar variëren sommige conceptwaarden. Tabel 4 toont bijvoorbeeld een cijfertabel waarin we niet alleen het inwonersaantal van Antwerpen in 2019

zien, maar ook de inwonersaantallen van alle andere Vlaamse gemeenten in 2019. Deze tabel bevat voor elk cijfer een aparte conceptuele definitie maar je merkt al gauw dat er veel overlap zit tussen deze definities. Elke definitie bestaat immers nog steeds uit de drie concepten “parameter”, “gemeente” en “jaartal” en bij elk cijfer is de parameter het aantal inwoners en het jaartal 2019. Omwille van deze overlap, zullen we de attributen van een cijfertabel dan ook anders organiseren dan de attributen van een individueel cijfer.

2.1. CONCEPTEN & DIMENSIES

In een cijfertabel kunnen we een onderscheid maken tussen twee soorten concepten. In een cijfertabel zullen de waarden van sommige concepten variëren over de cijfers en helpen elk cijfer uniek te identificeren, terwijl de waarden van andere concepten constant blijven over alle cijfers heen. De concepten met variërende waarden noemen we de **dimensies** van de cijfertabel (Eurostat 2023), maar ze worden ook wel **identificeerders** (identifiers) genoemd (UNECE 2024). In tabel 4 is er slechts één dimensie, namelijk de gemeente, aangezien elk cijfer het aantal inwoners van een andere gemeente weergeeft in 2019. De andere concepten, waarvan de waarde constant blijft, worden **beschrijvers** (describers) genoemd (UNECE 2024) omdat deze concepten enkel bijkomende informatie geven zonder de cijfers verder te identificeren. In tabel 4 zijn de beschrijvers de parameter en het jaartal, omdat elk cijfer verwijst naar een inwonersaantal in 2019 en niets anders.

Merk op dat beschrijvers in een cijfertabel dimensies kunnen worden wanneer verschillende cijfertabellen worden samengevoegd en omgekeerd. Als we bijvoorbeeld een tabel met inwonersaantallen definiëren over verschillende jaren heen in plaats van enkel 2019, wordt jaartal een bijkomende dimensie van die cijfertabel. Omgekeerd kan een dimensie ook een loutere beschrijver 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 beschrijver in plaats van een dimensie. Het verschil tussen dimensies en beschrijvers is dus relatief want het hangt af van welke tabel je precies bekijkt.

2.2. CONCEPTUELE EN OPERATIONELE DEFINITIE

Aangezien een cijfertabel bestaat uit cijfers met sterk overlappende conceptuele en operationele definities, kunnen we deze definities veralgeme-

Tabel 4: *We publiceren statistieken doorgaans niet als individuele cijfers maar als cijfertabellen.*

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

nen naar het niveau van de hele tabel. Dit maakt informatiebeheer efficiënter.

De **conceptuele definitie** van een cijfertabel verwijst nog steeds naar de interpretatie die een doorsnee gebruiker aan de cijfers toekent. De conceptuele definitie bevat daarom nog steeds een verwijzing naar alle concepten. Voor beschrijvers verduidelijkt de conceptuele definitie steeds welke waarde deze concepten aannemen. Voor dimensies benoemt de conceptuele definitie echter enkel het concept zelf, de waarden van de dimensies variëren immers over de cijfers en zijn terug te vinden in de cijfertabel zelf. De cijfertabel in Tabel 4 kan bijvoorbeeld conceptueel worden gedefinieerd als het “aantal inwoners in verschillende gemeenten in 2019”. De concepten in deze definitie zijn nog steeds de parameter, het jaartal en de gemeente. Zowel de parameter en het jaartal zijn loutere beschrijvers en krijgen een duidelijke waarde meegegeven in de definitie zelf, namelijk het ‘aantal inwoners’ en ‘2019’. Gemeente is echter een dimensie. Voor de exacte interpretatie van de individuele cijfers moet de gebruiker kijken naar de waarden op deze dimensie in de cijfertabel zelf.

De **operationele definitie** van een cijfertabel beschrijft hoe de cijfertabel precies werd verzameld en hoe ze gereproduceerd kan worden. Bovendien preciseert de operationele definitie steeds exact welke waarden de attributen en dimensies binnen de tabel aannemen. Voor tabel 4 luidt de operationele definitie bijvoorbeeld: “De grootte van de wettelijke bevolking op 1 januari 2019, 0.00 uur, per Vlaamse gemeente volgens de NIS code indeling in 2019. De wettelijke bevolking omvat...”

Door de expliciete en exacte beschrijving van dimensies bepaalt de operationele definitie steeds ondubbelzinnig uit hoeveel cijfers een cijfertabel precies bestaat, zelfs als sommige cijfers een missende of versluierde waarde hebben. De operationele definitie van Tabel 4 benoemt bijvoorbeeld 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 beschrijvers. De reeks bevat namelijk cijfers over de grootte van de wettelijke bevolking en bijvoorbeeld niet de gewoonlijk verblijvende bevolking, en voor het jaar 2019 en geen ander kalenderjaar.

De presentatie van een statistiekreeks gebeurt doorgaans niet zoals in tabel 4. Beschrijvers worden vaak uit de rijen verwijderd en opgenomen in conceptuele definitie van de cijfertabel, die gebruikt wordt als titel. Tabel 5 toont een meer gangbare weergave van een uitgebreidere cijfertabel. In deze tabel zie je dat het concept jaartal met waarde “2019” enkel wordt vermeld in de titel van de tabel. De tabel telt verder drie dimensies. Dimensies gemeente en geslacht staan in de rijen. De dimensie parameter, die hier een verschil maakt tussen aantallen en percentages, staat in de kolommen.

2.3. SOORTEN ATTRIBUTEN

In paragraaf 1 zagen we dat statistieken nog heel wat andere attributen hebben naast de conceptuele en operationele definitie. Net zoals deze definities, kan informatie over deze andere attributen in een cijfertabel vaak veralgemeend worden van het niveau van het cijfer naar het niveau van de hele tabel. In de praktijk kunnen we attributen zelfs op drie verschillende niveaus toewijzen.

Een eerste soort attributen zijn attributen op het niveau van de hele cijfertabel. Deze attributen bevatten informatie die geldt voor alle cijfers in de tabel. In Tabel 4 is het concept jaartal een voorbeeld van zo’n attribuut aangezien alle cijfers verwijzen naar bevolkingsgroottes in jaartal 2019. Informatie over dit soort attributen hoeft uiteraard enkel op tabelniveau te bewaren en niet apart per cijfer.

De tweede soort attributen zijn attributen op het niveau van individuele cijfers. Zij bevatten informatie die slechts geldt voor één enkel cijfer en niet voor alle andere cijfers in de tabel. Een voorbeeld van zo’n attribuut is informatie over de vertrouwelijkheid van elk cijfer. Bepaalde cijfers in een tabel kunnen beschouwd worden als vrij te publiceren terwijl andere cijfers niet openbaar gemaakt mogen worden omdat ze vertrouwelijke informatie kunnen ontsluiten. Opgelet, de aanduiding van deze vertrou-

Tabel 5: Bij cijfertabellen worden beschrijvers opgenomen in de conceptuele definitie in de titel. De rijen en kolommen bevatten enkel de dimensies.

Aantal en percentage inwoners per geslacht in Vlaamse gemeenten 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
...

weljkheid identificeert de cijfers niet en is dus geen dimensie van de tabel, ook al varieert deze informatie wel van cijfer tot cijfer.

Een derde soort attributen bevinden zich op een tussenniveau, namelijk attributen die betrekking hebben op bepaalde groepen van cijfers. Zo kunnen, hypothetisch, de cijfers in Tabel 4 uit één bepaalde provincie misschien op een andere manier zijn verzameld. De cijfers uit de verschillende provincies vormen dan groepen en per groep heb je een andere operationele definitie als attribuut.

3. STATISTIEK ALS EEN TABELREEKS

Cijfertabellen bevatten cijfers met sterk overlappende definities maar om statistieken te definiëren kunnen we nog een stap verder gaan. Het kan voorkomen dat we cijfers publiceren die nog steeds sterk overlappende conceptuele en operationele definities hebben maar die desondanks niet exact dezelfde concepten delen. Daardoor zijn ze niet gemakkelijk te organiseren in één enkele gestructureerde dimensionale cijfertabel. In die situatie kunnen we ervoor kiezen om statistieken te organiseren in een [reeks van cijfertabellen](#). Als we praten over statistieken, kan dit dus ook verwijzen naar zo'n tabelreeks in plaats van één enkele cijfertabel of één individueel cijfer.

Zoals we in vorige voorbeelden al toonden, publiceert Statistiek Vlaanderen heel wat cijfers over het aantal inwoners in Vlaamse gemeenten. Dat doet ze niet alleen per gemeente en per jaar, ze publiceert ook cijfers volgens verschillende demografische groepen. Zo publiceert ze bijvoorbeeld cijfers over het aantal mannen en vrouwen en over het aantal inwoners in verschillende leeftijdscategorieën voor elke Vlaamse gemeente. Aangezien geslacht en leeftijd twee verschillende concepten zijn worden deze cijfers best georganiseerd in twee verschillende cijfertabellen zoals te zien in Tabel 6. Al deze cijfers worden echter op een gelijkaardige manier afgeleid uit het Rijksregister en kunnen dus op één en dezelfde plaats gedocumenteerd en beheerd.

Door de overlap tussen de conceptuele definities van cijfers in een tabelreeks kunnen we opnieuw deze definities generaliseren van het niveau van de cijfertabellen naar het niveau van de tabelreeks. Tabel 6 heeft bijvoorbeeld een overkoepelende conceptuele definitie als titel waarin nog steeds alle concepten aan bod komen, namelijk de parameter (het aantal inwoners), gemeente (alle Vlaamse gemeenten) en het jaartal (2019). Daarnaast vermeldt het ook de nieuwe concepten geslacht en leeftijd. Elke cijfertabel in de reeks heeft echter ook nog steeds haar eigen conceptuele definitie ter aanvulling van de titel van de hele tabelreeks.

Net zoals de conceptuele definities, kunnen we uiteraard ook de operationele definities generaliseren naar de hele tabelreeks. De operationele definitie van de tabelreeks in Tabel 6 luidt bijvoorbeeld: “De grootte van de wettelijke bevolking op 1 januari 2019 per gemeente volgens de NIS code indeling in 2019 volgens (1) geslacht zoals geregistreerd in het Rijksregister (man/vrouw), en (2) leeftijd zoals geregistreerd in het Rijksregister en opgedeeld in vier categorieën (0-18 jaar, 19-35 jaar, 36-64 jaar en 65+ jaar). De wettelijke bevolking omvat...”. Merk op dat de operationele definitie een heel duidelijke opsomming geeft van de cijfertabellen in de tabelreeks, namelijk een tabel volgens geslacht en een tabel volgens leeftijd. Daarnaast is deze definitie nog steeds duidelijk expliciet over hoe de verschillende concepten worden gemeten en opgedeeld. Zo beschrijft ze heel duidelijk dat geslacht wordt opgedeeld in de twee waarden “man” en “vrouw” en leeftijd in de vier waarden “0-18 jaar”, “19-35 jaar”, “36-64 jaar” en “65+ jaar”. Uit de operationele definitie van de tabelreeks moet je dan ook de operationele definitie van elke individuele cijfertabel in de reeks kunnen afleiden.

Het spreekt voor zich dat we nu ook attributen op een vierde niveau kunnen definiëren. Naast attributen van individuele cijfers, groepen cijfers en cijfertabellen, kunnen er immers ook attributen zijn die gelden voor alle cijfers in de hele tabelreeks. Dit geldt uiteraard nog steeds voor alle soorten attributen zoals concepten en conceptwaarden in de conceptuele en operationele definities, kwaliteitsevaluatie, gebruikersbehoeften,

Tabel 6: Tabelreeksen zijn verzamelingen van cijfertabellen met sterke overlap tussen de conceptuele en operationele definities.

Aantal inwoners naar geslacht en leeftijd in Vlaamse gemeenten in 2019.					
Volgens geslacht.			Volgens leeftijdsgroep.		
Gemeente	Geslacht	Aantal	Gemeente	Leeftijd	Aantal
Aartselaar	man	7089	Aartselaar	0-18 jaar	3475
Aartselaar	vrouw	7204	Aartselaar	19-35 jaar	2938
Antwerpen	man	262921	Aartselaar	36-64 jaar	5948
Antwerpen	vrouw	263014	Aartselaar	65+ jaar	3928
Boechout	man	6506	Antwerpen	0-18 jaar	133495
Boechout	vrouw	6760	Antwerpen	19-35 jaar	95092
Boom	man	9024	Antwerpen	36-64 jaar	157700
Boom	vrouw	9220	Antwerpen	65+ jaar	102493
Borsbeek	man	5270	Boechout	0-18 jaar	2302
Borsbeek	vrouw	5584	Boechout	19-35 jaar	2505
...

en andere administratieve metadata. Net zoals de operationele definitie, moeten gebruikersbehoeften en kwaliteitsevaluatie nog steeds gedefinieerd zijn zodat ze alle cijfers in de tabelreeks rechtvaardigen en evalueren. In Tabel 6 zijn de parameter “aantal inwoners” en het jaartal “2019” attributen op het niveau van de tabelreeks, want alle cijfers in deze tabelreeks verwijzen naar aantallen inwoners in 2019.



Tabelreeksen zijn erg geschikt voor de weergaven van meervoudige concepten (Bhattacharjee 2019). Naast concreet en abstract, kan je concepten immers ook opsplitsen in enkelvoudige en meervoudige concepten.

- Een *enkelvoudig concept* verwijst naar een idee of kenmerk dat op één aspect betrekking heeft. Het is relatief eenvoudig te definiëren en te meten, en het omvat één afgebakend betekenisveld. Leeftijd is bijvoorbeeld een enkelvoudig concept: het verwijst enkel naar het aantal jaren sinds iemands geboorte. Het is met één cijfer te meten en wordt in de praktijk meestal als een numerieke variabele weergegeven (in jaren, maanden, ...).
- Een *meervoudig concept* verwijst naar een idee dat uit meerdere dimensies of onderliggende componenten bestaat. Je hebt meerdere indicatoren nodig om het concept volledig te vatten. Welzijn is bijvoorbeeld een meervoudig concept, want het kan verschillende onderdelen bevatten zoals fysieke gezondheid, mentale gezondheid, sociale relaties of economische situatie omvatten. Om het te meten zijn verschillende indicatoren of meetinstrumenten nodig – bijvoorbeeld via samengestelde indexen of uitgebreide vragenlijsten.

Zowel enkelvoudige als meervoudige concepten kunnen concreet en abstract zijn, al zullen abstracte concepten in de praktijk doorgaans vaker meervoudig zijn dan enkelvoudig.

- *Leeftijd in jaren is een **concreet en enkelvoudig** concept. Dit is een direct observeerbaar, meetbaar kenmerk (concreet) en het gaat slechts over één dimensie (enkelvoudig). De meting is objectief en eenduidig.*
- *kwiteit van een woning is een **concreet en meervoudig** concept. Het bestaat uit verschillende woningkenmerken die stuk voor stuk direct observeerbaar (concreet) zijn. Ze beschrijven echter enkel allemaal samen het meervoudig concept “kwiteit van een woning”. Je hebt dus meerdere concrete indicatoren nodig om het volledig te vatten.*
- *Subjectieve gezondheid is een **abstract en enkelvoudig** concept. Dit is een subjectief, niet rechtstreeks observeerbaar gevoel (abstract), maar het wordt in onderzoek meestal benaderd als één enkel kenmerk, bijvoorbeeld via een enkele vraag: “Hoe is uw gezondheid over het algemeen?”. Je kan natuurlijk discussiëren of zo’n enkele vraag de nuances in subjectieve gezondheid wel volledig vat en of het niet als meervoudig concept moet worden bekeken. Opgelet, objectieve gezondheid is sowieso een meervoudig concept dat je meet via allerhande medische testen.*
- *Sociaal isolement is een **abstract en meervoudig** concept. Dit is een abstract concept dat meerdere onderliggende aspecten omvat, zoals subjectieve eenzaamheid, frequentie van sociaal contact of deelname aan sociale activiteiten. Het vereist een set van vragen of indicatoren om het volledig te meten.*

4. CONCLUSIE

Samenvattend, een statistiek kan dus op meerdere niveaus gedefinieerd worden. Een statistiek verwijst immers naar individuele cijfers die ook georganiseerd kunnen worden in cijfertabellen, en zo’n cijfertabellen kunnen dan weer gebundeld worden in tabelreeksen (zie Figuur 7). Alle drie niveaus zijn nuttig om in het achterhoofd te houden.

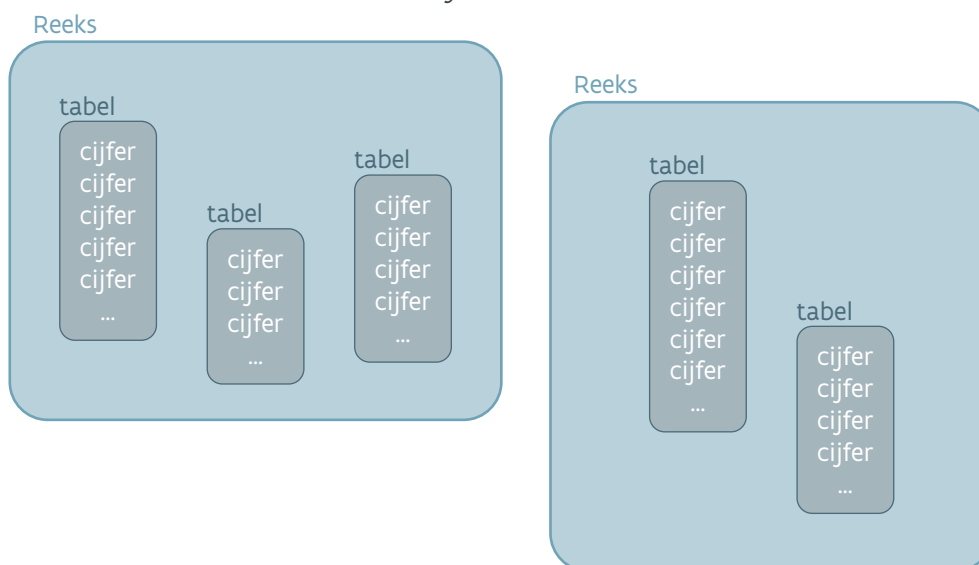
- Focus op statistieken als individuele cijfers is nuttig want voor elk individueel gepubliceerd cijfer moet juiste metadata worden gepubliceerd, zonder uitzondering. Meer specifiek moet voor elk cijfer duidelijk zijn wat de conceptuele en operationele definitie is, voor welke gebruikersbehoeften dit cijfer werd gepubliceerd, en hoe de kwaliteit van het cijfer wordt geëvalueerd.

- Focus op cijfertabellen is nuttig omdat tabellen de bouwstenen zijn van gestructureerde databases. Zo'n databases vormen wereldwijd de standaard om cijfers te bewaren en te organiseren. De term "tabel" is bovendien een standaardbegrip dat verwijst naar een dimensioneel gestructureerde reeks van cijfers, bijvoorbeeld in veel programmeertalen en in wetenschappelijke disciplines zoals wiskunde, maar ook in internationale standaarden voor data- en metadata-uitwisseling zoals SDMX (Eurostat 2023).
- Focus op tabelreeksen is nuttig omdat we ons momenteel binnen Statistiek Vlaanderen op dit niveau organiseren. Meer specifiek wordt de lijst van openbare statistieken in het Vlaams Statistisch Programma (VSP) gedefinieerd via tabelreeksen. Voor de term "reeks" kunnen we opnieuw de parallel leggen naar veel programmeertalen en wetenschappelijke disciplines waarin deze term gebruikt wordt om te verwijzen naar ongestructureerde of semigestructureerde verzamelingen van cijfers.

Hoewel de drie niveaus duidelijk van elkaar verschillen, is het belangrijk op te merken dat tabelreeksen geen internationale standaard zijn en ook niet gebruikt worden in kaders zoals SDMX (Eurostat 2023). Historisch gezien werken we bij Statistiek Vlaanderen op dit niveau, maar op langere termijn is het wenselijk de focus te verleggen van tabelreeksen naar cijfertabellen. Dit vereenvoudigt toekomstige afspraken met netwerkpartners en maakt informatiebeheer efficiënter.

In de volgende hoofdstukken zullen we de begrippen uit dit hoofdstuk gebruiken om te bespreken hoe we in de praktijk ons werk kunnen inrichten als openbare statistiedienst. Hoofdstuk ?? beantwoordt de vraag hoe we statistieken afbakenen? We leggen hierin het verschil uit tussen officiële, experimentele en niet-officiële statistieken, en we geven richtlijnen om tabelreeksen en cijfertabellen af te bakenen. Hoofdstuk ?? beantwoordt de vraag hoe we statistieken documenteren via metadatafiches. Hoofdstuk ?? beantwoordt de vraag hoe we statistieken beheren als datasets en hoe we structurele metadata gebruiken. Hoofdstuk ?? beantwoordt de vraag hoe we statistieken ontsluiten naar het brede publiek via databases, cijferapplicaties, cijferpagina's en nieuwsberichten.

Figuur 7: Een tabelreeks is een reeks van gestructureerde cijfertabellen, die op hun beurt bestaan uit een reeks van individuele cijfers.



BIBLIOGRAFIE

- Bhattacharjee, Anol (2019). [Social Science Research: Principles, Methods, and Practices](https://jwu.pressbooks.pub/socialscienceresearch/). Revised. Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License. University of Southern Queensland / Open Textbooks. URL: <https://jwu.pressbooks.pub/socialscienceresearch/>.
- Clark, Tom e.a. (2021). [Bryman's Social Research Methods](#). 6th edition. Oxford: Oxford University Press.
- Eurostat (2017). [European Statistics Code of Practice – Revised Edition 2017](#). Accessed: 2025-07-18. URL: <https://ec.europa.eu/eurostat/web/quality/european-quality-standards/european-statistics-code-of-practice>.
- (2021). [European Statistical System handbook for quality and metadata reports: 2021 re-edition](#). Collection: Manuals and guidelines. Theme: General and regional statistics. Luxembourg: Publications Office of the European Union. ISBN: 978-92-76-44408-4. DOI: 10.2785/616374. URL: <https://ec.europa.eu/eurostat/documents/3859598/13925930/KS-GQ-21-021-EN-N.pdf>.
 - (2023). [SDMX 3.0 API - Structure Queries Guidelines](#). Accessed: 2025-09-19. European Commission. URL: <https://ec.europa.eu/eurostat/web/user-guides/data-browser/api-data-access/api-detailed-guidelines/sdmx3-0/structure-queries>.

- Mertens, Donna M., Amy T. Wilson en Jori N. Hall (2025). [Program Evaluation Theory and Practice: A Comprehensive Guide](#). 3rd. New York: Guilford Press, p. 631. ISBN: 978-1-4625-5590-1.
- Neuman, W. Lawrence (2014). [Social Research Methods: Qualitative and Quantitative Approaches](#). 7th edition. Boston: Pearson.
- Rossi, Peter H., Mark W. Lipsey en Gary T. Henry (2018). [Evaluation: A Systematic Approach](#). 8th. Thousand Oaks, CA: SAGE Publications, Inc.
- UNECE (2024). [Generic Statistical Information Model \(GSIM\) Version 2.0: User Guide](#). Accessed: 2025-09-19. United Nations Economic Commission for Europe. Geneva. URL: <https://unece.org/sites/default/files/2025-03/GSIM%20User%20Guide.pdf>.
- United Nations Economic Commission for Europe (UNECE) (mei 2025). [Generic Statistical Business Process Model \(GSBPM\), Version 5.2](#). Endorsed by the Conference of European Statisticians, June 2025. URL: <https://unece.org/statistics/documents/2025/07/standards/generic-statistical-business-process-model-gsbpm-version-52>.
- United Nations Statistics Division (2025). [Handbook on Management and Organization of National Statistical Systems](#). 4th. New York: United Nations. URL: <https://projects.officialstatistics.org/hb-mgmt-org-nss/handbook/intro.html>.
- Wild, C.J., J.M. Utts en N.J. Horton (2018). "What Is Statistics?" In: [International Handbook of Research in Statistics Education](#). Red. door D. Ben-Zvi, K. Makar en J. Garfield. Springer International Handbooks of Education. Cham: Springer, p. 5–36. DOI: 10.1007/978-3-319-66195-7_1.