

Handleiding ICT-B

Mei 2026

1. Inleiding

De enquête 'ICT-gebruik bij bedrijven' meet jaarlijks gegevens over de automatisering en de toepassing van informatie- en communicatietechnologie (ICT) bij bedrijven in Nederland. Ze beschrijft onder andere het internetgebruik, de in- en verkoop via elektronische netwerken (i.e. e-commerce), het gebruik van software en andere ICT-applicaties zoals clouddiensten en kunstmatige intelligentie, en de genomen maatregelen en het aantal voorgekomen incidenten met betrekking tot ICT-veiligheid bij bedrijven. De uitkomsten van de enquête zijn van belang voor beleidsmakers, onderzoekers, bedrijven en brancheorganisaties.

De enquête wordt in eerste instantie uitgevoerd voor Eurostat. Daarnaast worden er vragen en subpopulaties toegevoegd voor het ministerie van Economische Zaken (t.b.v. StatLine) en het ministerie van Justitie en Veiligheid (t.b.v. StatLine en de Cybersecuritymonitor).

2. Documentatie

4.1.4

In 2025 is samen met 5.1.2.e een Business Model Canvas (BMC) opgesteld:

4.1.4

Een schematische weergave van de code is te vinden op Miro:

4.1.4

3. Populatie en steekproeftrekking

3.1 Populatie

Bedrijfstak

De populatie bestaat uit bedrijven uit de bedrijfstakken:

- C-N en Q (SBI2008)
- C-O en R (SBI2025).

Merk op dat de bedrijfstakken K en S (SBI2008) slechts *gedeeltelijk* zijn geïncludeerd t/m ICTBEDR2025. Vanaf ICTBEDR2026 is alleen de bedrijfstak T (SBI2025) gedeeltelijk geïncludeerd, want vanaf dat jaar wordt L (SBI2025) in opdracht van Eurostat **wel** volledig meegenomen. De overige bedrijfstakken worden ook volledig meegenomen.

Merk op dat de groep S 95.11-95.12 (SBI2008)/T 95.1 (SBI2025) niet tot de populatie behoort. Deze groep behoort tot de ICT-sector (Eurostat), waardoor het wel wordt meegenomen. Deze groep wordt **niet** meegenomen in de StatLine-totalen, **wel** in de Eurostat-totalen.

Bedrijfs grootte

Vanaf 2021 worden alle bedrijven meegenomen met ten minste 1 werkzame persoon. In 2020 lag de grens nog op 2 werkzame personen.

Zzp'ers

Een zzp'er is gedefinieerd als een persoon die voor eigen rekening of risico arbeid verricht in een eigen bedrijf of praktijk (zelfstandig ondernemer) en geen personeel in dienst heeft. Een zzp'er wordt alleen geïnccludeerd als hij het kenmerk BeRechtsvormCode == 1 (Eenmanszaak) heeft in het ABR.

3.2 Steekproeftrekking

In december van jaar t-1 wordt de steekproef getrokken voor het onderzoek van jaar t. De steekproef bestaat uit een integraal deel en een gedeelte op steekproefbasis:

1. **Integraal:**
 - o Bedrijven met 100 of meer werkzame personen ($gk_sbs \geq 71$) → hoofdzakelijk om goede cijfers te krijgen voor de module 'E-commerce'
 - o Datacenters
2. **Gestratificeerde aselecte steekproef:**
 - o Bedrijven met <100 werkzame personen → gestratificeerde aselecte steekproef op basis van $gk_sbs \times sbi_digit2$.

Steekproefontwerp

Voor ICTBEDR2026 is een nieuw steekproefontwerp gemaakt door 5.1.2.e en 5.1.2.e o.b.v. SBI2025. Dit steekproefontwerp bevat iets minder bedrijven dan het ontwerp dat in eerdere jaren is gebruikt. Het is niet meer terug te vinden hoe het oude steekproefontwerp tot stand is gekomen die t/m ICTBEDR2025 is gebruikt. In het bestand S:\ALR26\ALG_26\02_ONTWERP_26\Steekproef\ICTBEDR allocatie 05022026 alle variabelen.xlsx zijn alle aantallen per stratum etc. van het nieuwe ontwerp te vinden.

Voor het maken van dit onderwerp hebben wij geleverd:

- Een aantal variabelen (bijv. biometrisch)
- De belangrijkste indelingen (bijv. voor Eurostat en StatLine)
- De criteria voor betrouwbaarheid (i.e. 5%-regel)

De steekproefgroep heeft doorgerekend wat de minimale steekproefaantallen zijn om o.b.v. deze informatie niet boven de betrouwbaarheidscriteria uit te komen. Het is niet onvermijdelijk dat de resultaten altijd betrouwbaar zijn; soms heb je te maken met kleine strata en een lage respons.

Een uitgebreidere beschrijving van het nieuwe steekproefontwerp is hier te vinden:

Commented [EM1]: Tot en met ICTBEDR2025 werden ook integraal uitgevraagd:
• **Bedrijven met 10 t/m 99 werkzame personen in de SBI-groepen 47911 t/m 47919 (G, SBI2008).** Dit zijn webwinkels en worden uitgevraagd voor de module 'E-commerce' Vanaf SBI2025 bestaat deze uitsplitsing niet meer en daarmee is dit deel komen te vervallen.
• **DTC-bedrijven.** Het ASD-project over DTC-bedrijven is in 2025 voor het laatst uitgevoerd, omdat het DTC daarna is gestopt en overgegaan in het NCSC.

Commented [MV(2)]: Deze bedrijven worden vermoedelijk vanaf 2018 integraal uitgevraagd.

Commented [MV(3)]: De datacenters zijn pas later toegevoegd, vanaf ICTBEDR2023 (ALR24).

ICT-holdings

ICT-holdings zijn bedrijven die de vragenlijst geconsolideerd invullen. In de settingsfile (in: **holdings**) staat een lijst van deze bedrijven en welke BE's daarbij horen.

Stel, de volgende twee BE's horen bij hetzelfde overkoepelende bedrijf:

be_id	wp_actueel	omzet	anti_virus
123	2	10	1
456	5	80	1

Om de responslast verminderen, vult dit bedrijf (op hun eigen verzoek) één vragenlijst in namens beide BE's. We krijgen dan op één BE (bijv. 123) de gegevens binnen met o.a. de gecombineerde omzet (in dit geval dus 90).

In de code wordt BE 456 toegevoegd aan de data. De respons van de invullende BE wordt gekopieerd voor alle BE's die bij de holding horen. Omdat BE 123 heeft aangegeven dat anti_virus = 1, krijgt BE 456 dus ook anti_virus = 1.

Vervolgens worden variabelen van het type 'float' en 'int' herschaald. Stel dat de BE 123 een omzet invult van 90, dan gaan we deze waarde verdelen o.b.v. het aantal werkzame personen. De BE 123 krijgt daardoor uiteindelijk een omzet van $90 \cdot (2/7) = 26$, en de BE 456 een omzet van $90 \cdot (5/7) = 64$.

Portie 1 en portie 2

Omdat de steekproef van ICT-B erg groot is, wordt de steekproef willekeurig verdeeld over twee porties. Er zit vaak een week tussen (i.e. portie 1 krijgt de brief etc. een week eerder). Van alle bestanden zijn er dus 2 versies (bijv. STP_ICTBEDR2026JP01.csv, STP_ICTBEDR2026JP02.csv).

Verrijgingsbestand (AVV)

De steekproef wordt getrokken door de steekproefgroep (DVZ). Dit resulteert in een lijst met BE's. Een bepaalde BE kan echter meerdere CBS-personen (i.e. KVK-nummers) hebben. In de vragenlijst zetten we daarom voor zulke respondenten neer voor welke bedrijven de vragenlijst moet worden ingevuld. Daarnaast voeren we extra handelingen uit voor ICT-holdings. ICT-holdings bestaan uit meerdere BE's, maar slechts één BE (de 'hoofd-BE') daarvan ontvangt de vragenlijst. De berichtgever krijgt vervolgens wel een lijst van alle CBSP's en BE's te zien die daarbij horen. Voor ICT-holdings moeten we dus alléén de hoofd-BE opnemen in het AVV-bestand, maar wel de CBSP's van alle bijbehorende sub-BE's noemen.

AVV-bestand maken

Er zijn meerdere manieren om het AVV-bestand te maken:

- Gebruik het script **run_koppel_steekproef.ps1** van 5.1.2.e (S:\ALR26\ALG_26\02_ONTWERP_26\Steekproef\run_koppel_steekproef.ps1) i.c.m. de ict_koppel_tool (onderdeel van de ICT-analyser) en het configuratiebestand **koppel_abr_cbsp_settings.yml**.
- Gebruik het script **maken_AVV_bestand_emsk** van Elisca: S:\ALR26\ALG_26\02_ONTWERP_26\Steekproef\maken_AVV_bestand_emsk.R
- Na afloop van het proces wordt vanaf ICTBEDR2026 alle records met Sturing == 2 (zzp'ers) verwijderd uit het AVV-bestand. In 2026 waren dit maar ongeveer 8 records.

Let erop dat de december-stand van het jaar t-1 (dus 2512 voor ICTBEDR2026) wordt gebruikt.

AVV-bestand aanleveren

Het AVV-bestand moet worden aangeleverd in het aanleverloket:

4.1.4 Kies hierbij voor de optie 'AVV met **herhaling**'. Door de twee porties zijn er eigenlijk twee bestanden:

- AVV_ICTBEDR2026JP01_Variant1_verHH.csv
- AVV_ICTBEDR2026JP02_Variant1_verHH.csv

Logistieke test

De logistieke test vindt plaats vóór de trekking van de steekproef. Tijdens de logistieke test testen we of alles naar behoren werkt. Van DVZ ontvangen we een testbestand voor de steekproef (bijv. STP_ICTBEDR2026JP01_LT.csv). Vanuit onze kant testen wij of het aanleveren van een AVV-bestand lukt.

Hotspots

Hotspots zijn bedrijven tot 10 medewerkers (GK 0-3) die meer dan 3 verschillende surveyonderzoeken per 12 maanden hebben ontvangen, of bedrijven tot 20 medewerkers (GK 4) met meer dan 4 surveyonderzoeken per 12 maanden. Vanaf 2026 heeft DVZ er een nieuwe manier voor het uitsluiten van hotspots (4.1.4)

4.1.4 Het proces gaat als volgt in zijn werk:

1. Ze maken een lijst van alle bestaande hotspots, én van bedrijven die hotspot worden als ze nog één onderzoek krijgen.
2. De betrokken statistiekmakers kijken vóór de steekproeftrekking of deze bedrijven echt nodig zijn binnen de steekproef. Zo wordt al vóór de steekproeftrekking waar mogelijk voorkomen dat bedrijven op deze lijst worden geselecteerd. Hierdoor neemt het aantal hotspots af, zonder dat er afbreuk wordt gedaan aan de kwaliteit van de statistiek.

De criteria voor ICT-B voor het wel of niet includeren van hotspots zijn:

1. Een bedrijf is **cruciaal** als het een ICT-holding is.
2. Een bedrijf is **cruciaal** als het een datacenter is.
3. Een bedrijf is **cruciaal** als de steekproef fractie 1 is in een stratum.
4. Een bedrijf kan **cruciaal** zijn als het aantal hotspots in een stratum groter is dan het verschil tussen de populatieomvang en de steekproefomvang.

Bijvoorbeeld, stel we hebben de volgende hotspots in een bepaald stratum:

BE	pop	stp
123	27	25
456	27	25
789	27	25

In dit geval kunnen we niet alle 3 de hotspots uitsluiten, want dan is het aantal in de steekproef 24 i.p.v. 25. Dit betekent dat we er 1 moeten behouden. Let hierbij ook op evt. geblokeerde bedrijven in een stratum, want die vallen ook nog eens uit.

5. Een bedrijf is **niet cruciaal** als het voldoet aan één van de bovenstaande voorwaarden maar een actieve blokkade (EENH of OND) heeft.
6. Alle andere bedrijven zijn **niet cruciaal**.

Maken hotspots-bestand

In 2026 is hiervoor het script **maken_bestanden_hotspots_emsk.R** gebruikt
(S:\ALR26\ALG_26\02_ONTWERP_26\Steekproef\maken_bestanden_hotspots_emsk.R).

Nog meer zeggen over:

- Witte vlekken
- Wat is de kans dat bedrijven nog een keer worden getrokken?

4. Dataverzameling

4.1 Enquêtestructuur

De vragenlijst is elk jaar anders. Eurostat hanteert voornamelijk wisselende modules, die het ene jaar wel, en het andere jaar niet voorkomen, maar voegt soms ook nieuwe vragen of modules toe. Daarnaast worden jaarlijks (nieuwe) vragen toegevoegd voor andere opdrachtgevers (ministeries).

4.2 Enquêteontwikkeling

De vragenlijstontwikkeling voor het jaar t begint in de tweede helft van het jaar t-2. Dus bijv. de vragenlijstontwikkeling voor ICT-gebruik bij bedrijven 2027 begint in de tweede helft van 2025. Het begint met de Task Force/Working Groups van Eurostat. In het jaar t-1 begint vervolgens de samenwerking met DVZ, waarna de vraagteksten, de routing, en het testen van de vragenlijst volgt.

4.3 Waarneming

De waarneming vindt plaats van februari t/m begin oktober in het jaar t. Elk jaar worden er zo'n XX COM-berichten afgehandeld door het ICT-B-team.

5. Weging

De weging wordt vanaf verslagjaar 2025 uitgevoerd in R met behulp van het R package *survey*. Vóór 2025 werden de weegfactoren handmatig berekend met Pythoncode van 5.1.2.e Hij had de berekening op zijn beurt weer overgenomen vanuit de SPSS-code van vóór 2018.

Vanaf 2025 wordt het volgende weegmodel gebruikt: $gk_vanaf_1_zzp + (sbi_digit2_weging \times gk_weging) + (sbi_digit1_weging \times gk3_label) + (ictsector \times gk3_label)$.

Deze GK- en SBI-variabelen worden in de Pythoncode aangemaakt. Voor de definities van deze variabelen, zie de settingsfile of voer een query uit in SQL. Bijvoorbeeld:

```
SELECT DISTINCT sbi, sbi_digit2_weging  
FROM records_df_25_2
```

In 2025 is, op advies van 5.1.2.e van methodologie, gekozen voor een ander weegmodel. De hoofdreden was dat de weging plaatsvond op een te laag niveau (i.e. kleiner dan het publicatieniveau). Er was daardoor nog potentie tot het verlagen van de correctiegewichten. Daarnaast hield de oude weegmethode geen rekening met het feit dat sommige subgroepen integraal werden uitgevraagd (bijv. G47), waardoor de eindgewichten te groot waren voor deze bedrijven. Tot slot werd er niet goed opgehoogd naar de populatie van de ICT-holdings. Alle ICT-holdings kregen namelijk een eindgewicht van 1, maar er werd geen rekening gehouden met non-responderende holdings, waardoor deze buiten de populatie vielen.

In de praktijk wordt in R eerst de functie `svydesign()` gebruikt om een design te specificeren, waarbij de initiële insluitgewichten en de finite population correction (FPC) worden meegegeven. Vervolgens worden de populatietotalen berekend, en worden de gewichten met de functie `calibrate()` gekalibreerd naar deze bekende populatietotalen. Deze stappen leveren de eindgewichten op.

De weging wordt drie keer uitgevoerd om op te hogen naar het aantal bedrijven (**units**), het aantal werkzame personen (**wp_actueel**), en de totale omzet (**omzet_totaal**).

6. Imputatie

Wanneer imputatie?

Soms is het nodig dat er waarden worden geïmputeerd, namelijk wanneer:

1. *Een vraag leeggelaten kan worden door respondenten.*
Dit zijn vragen die géén eigenschap 'NOEMPTY' hebben in de vragenlijst van DVZ.
2. *Een vervolgvraag waarvan de filtervraag tijdens het gaafmaken op 1 wordt gezet.*
Tegenwoordig nemen we aan dat elk bedrijf toegang heeft tot internet. Maar soms geeft een bedrijf aan geen vast internet (**xdsl**) én geen mobiel internet (**mobint**) te hebben. In dat geval wordt **xdsl** alsnog op 1 gezet tijdens het gaafmaken. Bedrijven die aangeven vast internet te hebben, krijgen in de vragenlijst de vraag te zien wat de maximale downloadsnelheid is van de vaste internetverbinding (**download_snelheid2**). De bedrijven die hadden aangegeven geen vast internet te hebben, hebben deze vraag niet gekregen. Maar omdat we **xdsl** op 1 hebben gezet, moet er nog wel een waarde voor **download_snelheid2** geïmputeerd worden.

Type imputatie

Welke waarde wordt geïmputeerd, hangt af van het type variabele.

Type variabele	Imputatiemethode
Dict, bool	Random donorimputatie
Float, percentage	Mediaan

Voor variabelen van het type 'float' of 'percentage' wordt voor ieder bedrijf met een ontbrekende waarde wordt via loting een willekeurig bedrijf gezocht in hetzelfde stratum (dus die vergelijkbaar is qua grootte en activiteit) waarvan de betreffende waarde wordt overgenomen. Alle variabelen worden onafhankelijk van elkaar geïmputeerd.

Voor variabelen van het type 'dict' of 'bool' wordt de mediaan van de waarden in desbetreffende stratum gebruikt.

Als een stratum te weinig donoren heeft (<5 bedrijven), wordt de zoektocht naar donoren uitgebreid door in een groter stratum te kijken. De strata die worden gebruikt zijn: het stratum (**gk6_label x sbi_digit2**), daarna (**gk3_label x sbi_digit1**), daarna (**gk3_label**), en tot slot alle bedrijven bij elkaar. Eerder geïmputeerde waarden worden niet gezien als geldige donorwaarden.

Commented [MV(4)]: Dit wordt gespecificeerd in de settingsfile.

7. Data-analyse

Beschrijvende statistieken: Welke statistieken worden gebruikt om de enquêtegegevens te analyseren (bijv. gemiddelde, mediaan, frequenties)?

Betrouwbaarheid van de gegevens: Hoe wordt de betrouwbaarheid van de gegevens beoordeeld (bijv. betrouwbaarheidsinterval, foutmarge)?

8. Statistische beveiliging

9. Publicatie