


Statistische Analyse Toegankelijkheid Stemlokalen

Eindrapport



Datum	29 augustus 2023
Auteurs	Dr. Seyit Höcük Pradeep Kumar, MSc. Dr. João Guimarães
Versie	2.0

Uitgave

Centerdata
info@centerdata.nl
www.centerdata.nl

Contact

Seyit Höcük
seyit.hocuk@centerdata.nl

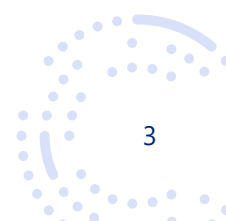
© Centerdata, Tilburg, 2023

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of op enige andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.



Inhoudsopgave

Samenvatting	4
1 Inleiding	6
1.1 Aanleiding en doel	6
1.2 Onderzoeksvragen	6
2 Data en methode	8
2.1 De datasets	8
2.2 Methodes	8
3 Statistische analyses	10
3.1 Verschillen in jaar	10
3.2 Verschillen tussen provincies	12
3.3 Verschillen tussen gemeentegroottes	14
3.4 Verschillen in beoordeling tussen controleurs	16
3.5 Statistische verbanden data Ongehinderd	17
4 Steekproeftrekking	21
4.1 Het algoritme	21
4.2 De steekproeftrekking	22





Samenvatting

Tijdens verkiezingen kunnen kiezers stemmen in stemlokalen verspreid door het hele land. Volgens de kieswet moeten alle stemlokalen toegankelijk zijn voor mensen met een lichamelijke beperking. De toegankelijkheidssituatie wordt beoordeeld via steekproefonderzoeken waarop er conclusies volgen op de toegankelijkheid van stemlokalen. Om de representativiteit van de steekproeven en de subjectiviteit van de beoordelingen te toetsen is er een onderzoek ingesteld door het ministerie van Binnenlandse Zaken en Koninkrijksrelaties. Er zijn zes onderzoeksvragen gesteld door het ministerie. Dit rapport beantwoordt de onderzoeksvragen via een statistische analyse naar de uitkomsten van de steekproefonderzoeken van de gemeenteraadsverkiezingen van 2022 en de Provinciale Staten- en waterschapsverkiezingen van 2023. De samenvattende uitkomsten van de statistische analyses zijn hieronder uitgelicht.

Verschillen in jaar

- Er is een statistisch significant verschil gevonden tussen de gemiddelde beoordelingen van de steekproefonderzoeken van de verkiezingen van 2022 en 2023.
- Er werden in 2022 60 stemlokalen beoordeeld ten opzichte van 120 in 2023. Daarnaast is de steekproeftrekking voor het onderzoek naar toegankelijkheid van stemlokalen in 2023 gedaan naar aanleiding van een statistisch advies voor betere landelijke representativiteit. Dit is een mogelijke verklaring voor het gevonden verschil tussen de twee steekproefonderzoeken.

Verschillen tussen provincies

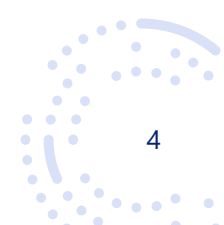
- Er zijn géén statistisch significante verschillen gevonden tussen de gemiddelde waarderingen van de provincies wanneer provincie Friesland niet meegenomen wordt in de analyses.
- Provincie Friesland wijkt sterk af van de rest. De gemiddelde score is significant lager en de spreiding is kleiner ten opzichte van de andere provincies. Een mogelijke verklaring is dat hier tijdens de verkiezingen van 2023 slechts vier stemlokalen zijn beoordeeld en tijdens de verkiezingen van 2022 geen een.

Verschillen tussen gemeenten

- Er zijn géén statistisch significante verschillen gevonden tussen de gemiddelde beoordelingen van kleine (inwoneraantal < 40.000), middelgrote (inwoneraantal 40.000 t/m 150.000), en grote (inwoneraantal > 150.000) gemeenten.

Verschillen tussen controleurs

- Er zijn statistisch significante verschillen gevonden tussen de gemiddelde beoordelingen van controleurs voor de verkiezingen van 2023.
- Voor de verkiezingen van 2022 zijn er geen statistisch significante verschillen gevonden, echter met een kanttekening. Er waren weinig (minder dan 10) bezochte stemlokalen per controleur waardoor de statistische analyse een onzekerder beeld geeft.
- Verschillen in de beoordelingen van controleurs zijn mogelijk te herleiden naar onderlinge verschillen tussen gemeenten, omdat iedere gemeente door een enkele controleur is beoordeeld. Gemeentelijke verschillen zijn statistisch gezien niet te toetsen, want het aantal beoordeelde stemlokalen per gemeente is te laag (gemiddeld 3 per gemeente).

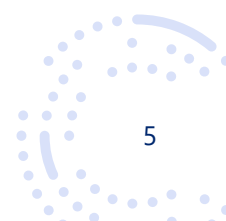




Verbanden tussen toetsingsonderdelen en eindbeoordeling

- Er zijn géén sterke correlaties (onderlinge samenhang) gevonden tussen de beoordelingen van de 14 toetsingsonderdelen. De beoordelingen zijn onafhankelijk van elkaar.
- Er zijn ook géén sterke correlaties gevonden tussen de individuele beoordelingen van de 14 toetsingsonderdelen en de eindwaardering. Een eindwaardering hangt niet (sterk) af van een bepaald toetsingsonderdeel.
- De gemiddelde toegankelijkheidsscore van toetsingsonderdeel 1.2 'Parkeergelegenheid' is sterk gedaald, van 9,44 in 2022 naar 6,67 in 2023. Daarnaast is de spreiding (standaarddeviatie) in de beoordelingen is juist sterk toegenomen, van 1,52 naar 4,73.
- De spreiding (standaarddeviatie) van de toegankelijkheidsscore van toetsingsonderdeel 3.7 'Leesloop' is sterk gestegen, van 0,60 in 2022 naar 3,30 in 2023.

Naast de statistische analyses is er ook een steekproeftrekking van 120 stemlokalen gedaan met een goede afspiegeling van de landelijke situatie voor de Europese Parlementsverkiezingen van 2024. Er is hiervoor een herbruikbaar algoritme geschreven die de trekking doet. De lijst van gemeenten is op 17 juli 2023 opgeleverd aan het ministerie.





1 Inleiding

1.1 Aanleiding en doel

Alle stemlokalen in Nederland moeten toegankelijk zijn voor mensen met een lichamelijke beperking. Dit is vastgelegd in de Kieswet (art. J 4).¹ De gemeenten zijn verantwoordelijk voor het aanwijzen, inrichten, en toegankelijk maken van stemlokalen. Om de toegankelijkheidssituatie te beoordelen worden er data verzameld in opdracht van het ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK).

Deze data worden sinds 2022 verzameld door Ongehinderd² op basis van een steekproef van gemeenten. Om te komen tot een representatieve steekproef heeft BZK prof. Ernst Wit gevraagd om een statistisch advies uit te brengen.³ Naar aanleiding van de adviezen van het rapport is de wens ontstaan om een statistische analyse te laten uitvoeren naar de uitkomsten van eerdere steekproeven. Daarnaast wil BZK, met het oog op de verkiezingen voor het Europees Parlement (EP) in 2024, een steekproeftrekking van gemeenten laten doen waarbij 120 stemlokalen op toegankelijkheid zullen worden getoetst.

BZK heeft Centerdata gevraagd om de statistische analyses uit te voeren op de data die zijn verzameld tijdens de gemeenteraadsverkiezingen van 2022⁴ en de Provinciale Staten- en waterschapsverkiezingen van 2023⁵ en daarnaast ook de steekproeftrekking te doen voor de verkiezingen voor het EP in 2024. Hiervoor zijn een aantal onderzoeksvragen gesteld. In dit eindrapport worden de onderzoeksvragen van BZK beantwoord via statistische analyses op de verzamelde data.

1.2 Onderzoeksvragen

Door BZK zijn er zes onderzoeksvragen vastgesteld. Deze zijn:

1. Wat is de betrouwbaarheid en nauwkeurigheid van de uitkomsten van het onderzoek van Ongehinderd naar de toegankelijkheid van stemlokalen in 2023?
2. Welke statistische verbanden zijn er te leggen tussen de data van Ongehinderd?
3. Zijn er controleurseffecten waar te nemen (wordt de uitkomst beïnvloed door wie de controle op toegankelijkheid heeft gedaan)?
4. Zijn er verschillen in toegankelijkheid tussen gemeenten en/of provincies waar te nemen?
5. Hoe verhouden de uitkomsten van het onderzoek van Ongehinderd naar toegankelijkheid van stemlokalen in 2023 zich tot de uitkomsten van hun onderzoek in 2022, en hoe kunnen deze verschillen worden verklaard?

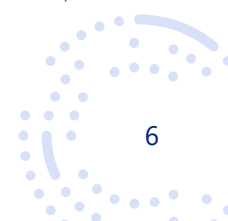
¹ https://wetten.overheid.nl/BWBR0004627/2023-06-20/#AfdelingII_HoofdstukJ_Paragraaf1_ArtikelJ4

² <https://www.ongehinderd.nl/>

³ Statistisch advies voor de steekproef methodologie met betrekking tot de toegankelijkheid van stemlokalen; Prof. Dr. Ernst C. Wit. Università della Svizzera italiana, 5 maart 2023.

⁴ Eindrapport Steekproef Toegankelijkheid Stemlokalen Gemeenteraadsverkiezingen 14, 15, 16 maart 2022; Ongehinderd, gepubliceerd op 13 mei 2022.

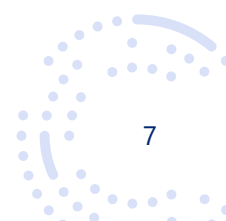
⁵ Eindrapport Steekproef Toegankelijkheid Stemlokalen Provinciale Staten- en Waterschapsverkiezingen 15 maart 2023; Ongehinderd, gepubliceerd op 29 mei 2023.





6. In welke gemeenten kunnen bij de volgende verkiezingen stemlokalen worden onderzocht (en hoeveel per gemeente), rekening houdend met een goede afspiegeling van de landelijke situatie? BZK wil graag dat de steekproeftrekking wordt gedaan op de wijze die is omschreven in het rapport van prof. Wit. Om gemeenten niet te veel te belasten, wil BZK graag dat gemeenten die zijn onderzocht bij de verkiezingen van de Provinciale Staten en waterschappen in 2023 van de steekproef worden uitgesloten. En dat per provincie gemeenten worden gekozen die bij elkaar in de buurt liggen, zodat het voor de controleurs mogelijk is om alle gemeenten uit de steekproef te bezoeken op de dag van de verkiezingen.

De eerste vijf onderzoeksvragen worden behandeld in **hoofdstuk 3**. De laatste onderzoeksvraag wordt beantwoord in **hoofdstuk 4**, waarbij er ook een lijst van gemeenten (inclusief het aantal stemlokalen) wordt gedeeld met BZK voor het steekproefonderzoek van de verkiezingen voor het EP in 2024.





2 Data en methode

2.1 De datasets

De data die zijn gebruikt in dit onderzoek zijn afkomstig van de steekproefonderzoeken naar toegankelijkheid van stemlokalen die uitgevoerd zijn door Ongehinderd. De toegankelijkheid van stemlokalen is getoetst aan de hand van de 'Criteria Integrale Toegankelijkheid Stembureaus' (CITS).⁶ Hiermee is in kaart gebracht in hoeverre mensen met een beperking gebruik kunnen maken van de ruimtes en functies in en rondom de stemlokalen die nodig zijn om een stem uit te brengen.

Zowel de data van het onderzoek naar gemeenteraadsverkiezingen van 2022, met een steekproeftrekking van 60 stemlokalen, als de Provinciale Staten- en waterschapsverkiezingen van 2023, met een steekproeftrekking van 120 stemlokalen, zijn meegenomen in de analyses van dit rapport.

De data van de einduitkomsten zijn verkregen via de opgeleverde eindrapporten van Ongehinderd. De achtergrondgegevens, zoals naam stembureau, gemeente, provincie, en nummer van controleur, zijn opgeleverd door Ongehinderd. Deze twee gegevenssets zijn aan de hand van een koppelvariabele ('Rapport ID') aan elkaar gekoppeld.

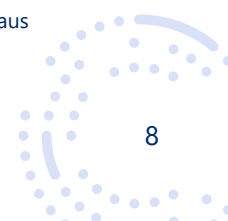
De datasets bevatten beoordelingen op zowel kwalitatief als kwantitatief niveau. Dat wil zeggen dat een deel van de beoordelingen (onderdelen 1.1, 1.2, en 3.3 t/m 3.7 van CITS) aan de hand van punten zijn gewaardeerd en een ander deel van de beoordelingen (onderdelen 1.3 t/m 3.2 van CITS) aan de hand van de categorieën *Goed* of *Fout*.

2.2 Methodes

Voor de gemeenteraadsverkiezingen in 2022 werd een van de CITS toetsingsonderdelen, de stoelen, buiten beschouwing gelaten aangezien deze vanwege de Covid maatregelen niet mochten worden gebruikt. In het steekproefonderzoek voor de Provinciale Staten- en waterschapsverkiezingen van 2023 zijn stoelen wel meegenomen. Om consistent te blijven in dit onderzoek in de vergelijking tussen de steekproefonderzoeken van 2022 en 2023 is het toetsingsonderdeel stoelen uit de statistische analyses weggelaten.

In de rapporten van Ongehinderd is de eindwaardering van een stemlokaal een totaalscore indien alle kwalitatieve beoordelingen als *Goed* worden gewaardeerd. Indien een of meer van de kwalitatieve beoordelingen als *Fout* worden gewaardeerd, dan wordt de eindwaardering per definitie 'Voldoet niet', zonder dat er een totaalscore wordt afgegeven. De conclusies in de eindrapporten van 2022 en 2023 zijn gebaseerd op welk percentage van de stemlokalen als 'Voldoet' zijn gewaardeerd. De focus lag dus op een kwalitatieve beoordeling en er werd geen sterke nadruk gelegd op de (waarde van de) totaalscores.

⁶ <https://www.rijksoverheid.nl/documenten/brochures/2020/09/30/criteria-integrale-toegankelijkheid-stembureaus>





In dit onderzoek worden zowel de kwalitatieve als de kwantitatieve waarderingen gebruikt voor de analyses. De kwalitatieve beoordelingen zijn overgenomen van Ongehinderd. Voor de kwantitatieve beoordelingen zijn de waarderingen *Goed* en *Fout* omgezet in een numerieke waarde waarbij *Goed* staat voor +1 en *Fout* voor -1. De keuze van de gekozen waardes zijn echter niet bepalend, omdat alle waarderingen in een vervolgstap gestandaardiseerd zijn tot een score tussen de 0 en 1.

Standaardisatie resulteert in een gelijkwaardige beoordeling van alle onderdelen waardoor geen onderdeel belangrijker is dan de ander. Dit heeft als grote voordeel dat arbitraire keuzes niet uitmaken en dat de variatie in het maximum aantal mogelijke punten per onderdeel geen impact meer heeft op het eindresultaat. Deze nieuwe, gestandaardiseerde methode van puntenwaardering resulteert in een toegankelijkheidsscore. De toegankelijkheidsscore, die de som is van alle individuele scores van de onderdelen, is genormaliseerd naar een getal tussen 0 en 10.

Naast een kwalitatieve score van het percentage van toegankelijke stemlokalen, creëren we een nieuwe kwantitatieve score gebruik makend van alle individuele waarderingen resulterend in een toegankelijkheidsscore tussen 0 en 10.

Via de kwantitatieve beoordeling, de toegankelijkheidsscore, kan worden bepaald hoe goed een stemlokaal presteert of hoe dicht een stemlokaal is bij een eindwaardering 'Voldoet'. Dit geeft meer inzicht in hoe ver een stemlokaal is verwijderd van een 'perfect grade'.

Kwantitatieve beoordelingen zijn ook geschikter voor statistische analyses, zoals toetsen voor significantie en correlatie. De verschillende dimensies van de data (verkiezingsjaar, provincie, gemeente, controleur) worden getoetst op significante verschillen in zowel de gemiddelde waarden als de spreiding van de steekproefverdeling aan de hand van meerdere statistische toetsen. De toegepaste technieken zijn: *Shapiro-Wilk* normaliteitstoets⁷, tweezijdig *Welch* t-toets⁸, ANOVA⁹ en als er verschillen zijn in de ANOVA de *Tukey's HSD*¹⁰, en *Pearsons* correlatie¹¹.

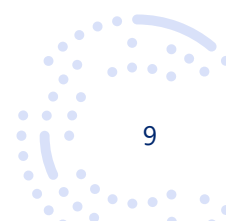
⁷ https://en.wikipedia.org/wiki/Shapiro-Wilk_test

⁸ https://en.wikipedia.org/wiki/Welch%27s_t-test

⁹ https://en.wikipedia.org/wiki/One-way_analysis_of_variance

¹⁰ https://en.wikipedia.org/wiki/Tukey%27s_range_test

¹¹ https://en.wikipedia.org/wiki/Pearson_correlation_coefficient





3 Statistische analyses

Er zijn vier dimensies in het onderzoek waarop statistische analyses worden uitgevoerd. Deze dimensies zijn: (verkiezings)jaar, provincie, gemeente, en controleur. Op iedere dimensie wordt er gekeken naar de verschillen in de data via exploratieve verkenning en of er statistisch significante verschillen zijn in de gemiddelde waarden van de toegankelijkheidsscores, de spreiding in deze, en de verdeling van de data. De verbanden worden weergegeven via grafieken, tabellen, en statistische toetsuitkomsten.

3.1 Verschillen in jaar

Op het eerste oog lijken er grote verschillen te zijn tussen de uitkomsten van de onderzoeken van Ongehinderd van 2022 en 2023. De conclusies die getrokken zijn komen erop neer dat in Nederland 37% van de stemlokalen toegankelijk zijn in 2022 terwijl er 'slechts' 21% van de stemlokalen toegankelijk zijn in 2023. Dit is een daling in toegankelijkheid van 16%. Hierbij wel de kanttekening dat er twee keer zoveel stemlokalen (120) zijn bezocht en beoordeeld in 2023.

Als er gekeken wordt naar de gemiddelde toegankelijkheidsscores van de twee jaren (de score volgens de nieuwe methode), dan is er eveneens een daling in toegankelijkheid. De gemiddelde toegankelijkheidsscore is gedaald van 8.70 (2022) naar 7.72 (2023), een daling van 11%. Dit komt overeen met de kwalitatieve beoordeling. Dit is een minder grote, maar nog steeds een sterke daling. De eindwaarderingen en de toegankelijkheidsscores worden in **Tabel 1** weergegeven.

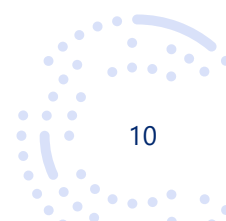
Tabel 1: Eindwaarderingen en toegankelijkheidsscores.

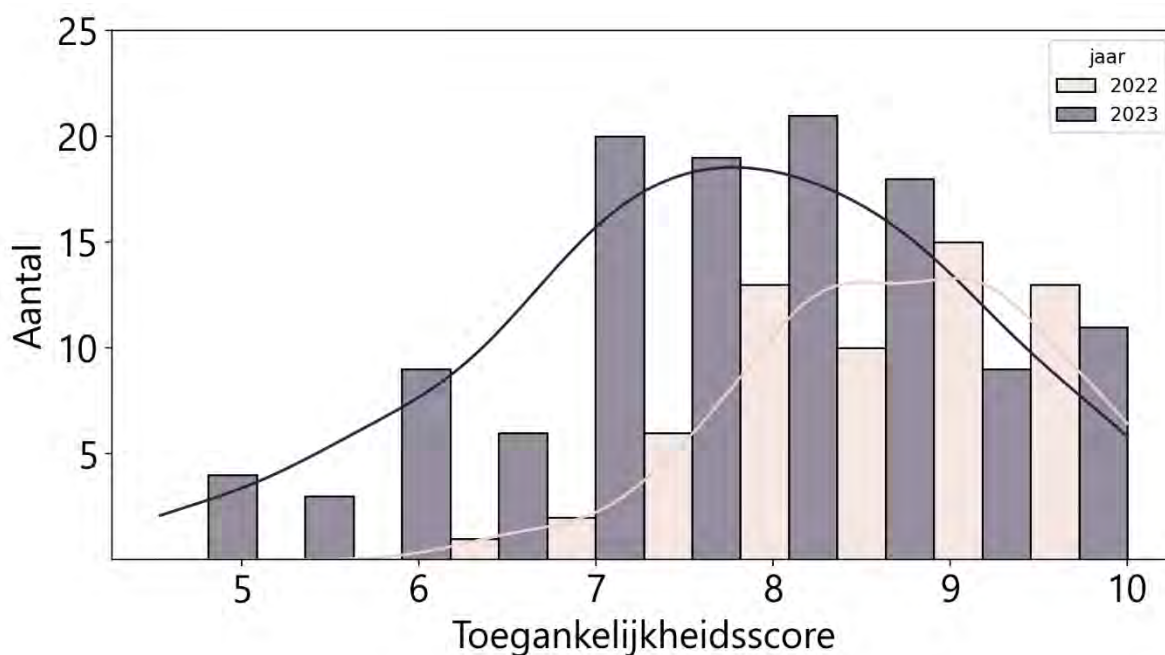
Jaar	Eindwaardering "Voldoet"	Gemiddelde toegankelijkheidsscore
2022	37%	8.70 (0.84)
2023	21%	7.72 (1.29)

De standaarddeviatie is tussen haakjes weergegeven.

Een gedetailleerder beeld wordt geschetst via een histogram van de toegankelijkheidsscores voor de twee jaren, waarmee de spreiding van de scores van stemlokalen goed zichtbaar wordt, zie **Figuur 1**.

In de grafiek wordt evident dat er, naast het verschil in gemiddelde waarde, ook een groot verschil is in de spreiding van de toegankelijkheidsscores tussen de twee jaren. Er is een grotere spreiding in de scores van 2023, met een minimum score van 5. De vraag blijft echter of de verschillen statistisch significant zijn. In de volgende paragraaf wordt daarop ingegaan.





Figuur 1: Histogram van toegankelijkheidsscores voor twee verschillende jaren.

Significantietest

De toegepaste *Shapiro-Wilk* normaliteitstest laat zien dat de toegankelijkheidsscores van ieder jaar afzonderlijk een normale verdeling volgen, zoals ook te zien is in **Figuur 1**. Een normale verdeling van de data is het verwachte patroon. De p-waarden van de toetsen zijn 0.1102 (2022) en 0.0930 (2023).¹² Dit betekent dat de nulhypothese van een normale verdeling is aangenomen voor beide jaren.¹³

Uit de twee jaren gecombineerd volgt echter geen normale verdeling. De *Shapiro-Wilk* toetsuitkomst heeft een p-waarde van $3.376e-4$, waarmee de nulhypothese wordt verworpen. Dit geeft al een hint dat de twee verdelingen niet afkomstig zijn van dezelfde onderliggende verdeling.

De tweezijdige *Welch* t-toets voor de populatiegemiddelden van beide jaren geeft aan dat de twee sets significant van elkaar verschillen. De berekende p-waarde is $6.06e-09$, waardoor de hypothese dat de gemiddelde waarden niet van elkaar verschillen verworpen is.

Een kanttekening hierbij is dat de steekproefgrootte van 2022 een kleinere set van 60 stemlokalen betreft. Daarnaast is de steekproeftrekking van 2022 uitgevoerd vóór de adviezen van prof. Wit. De kleinere steekproef van 60 stemlokalen heeft daardoor mogelijk te maken met selectie-effecten.

De steekproefonderzoeken van de verkiezingen van 2022 en 2023 verschillen statistisch significant van elkaar.

¹² Een p-waarde kleiner dan 0.01 wordt beschouwd als sterk significant en een p-waarde tussen de 0.01-0.05 als mild significant.

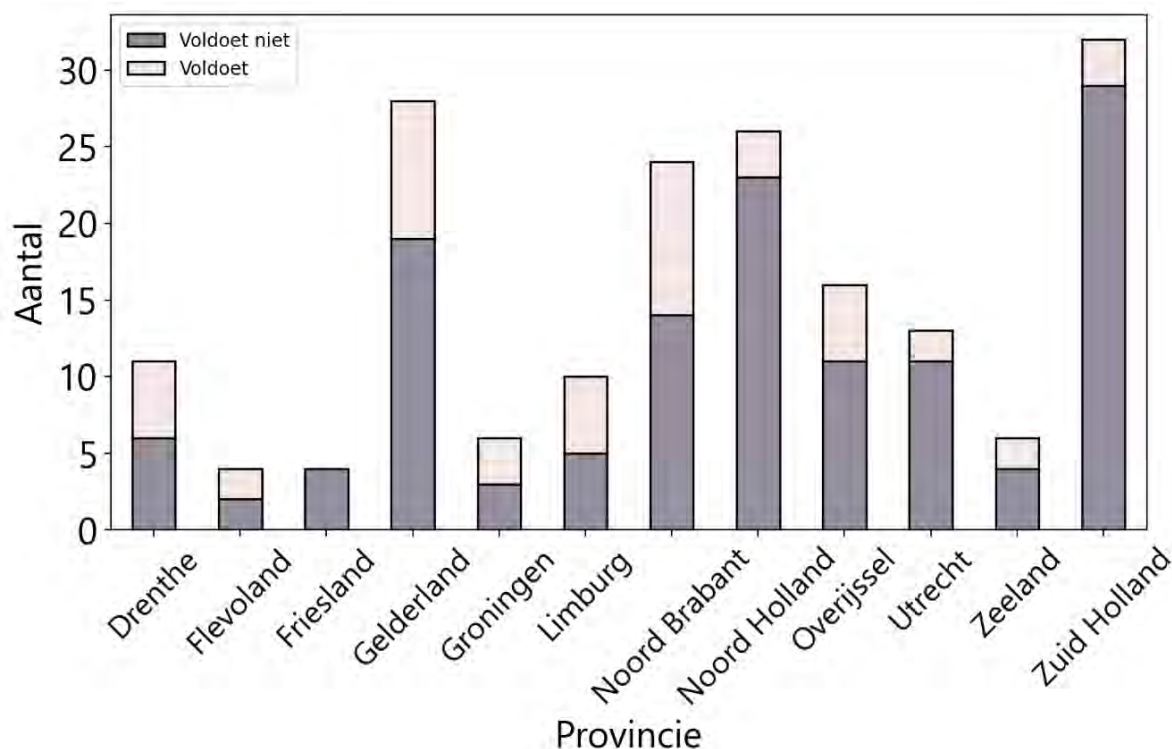
¹³ De nulhypothese is dat de verdeling volgens een normale verdeling is. Een p-waarde kleiner dan 0.05 verwerpt de hypothese.



3.2 Verschillen tussen provincies

Nu de verschillen tussen de jaren bewezen zijn, wordt er gekeken naar de toegankelijkheid van stemlokalen in de twaalf provincies. Dit wordt zowel aan de hand van de kwalitatieve beoordelingen als de kwantitatieve toegankelijkheidsscores gedaan.

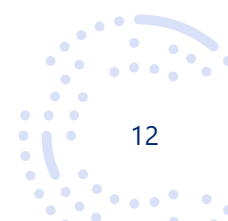
Eerst wordt de kwalitatieve situatie in beeld gebracht voor de complete dataset. In **Figuur 2** wordt het aantal wel en niet toegankelijke stemlokalen per provincie via een staafdiagram gevisualiseerd voor de samengevoegde data van de jaren 2022 en 2023.



Figuur 2: Toegankelijkheidspercentages van provincies van de steekproefonderzoeken.

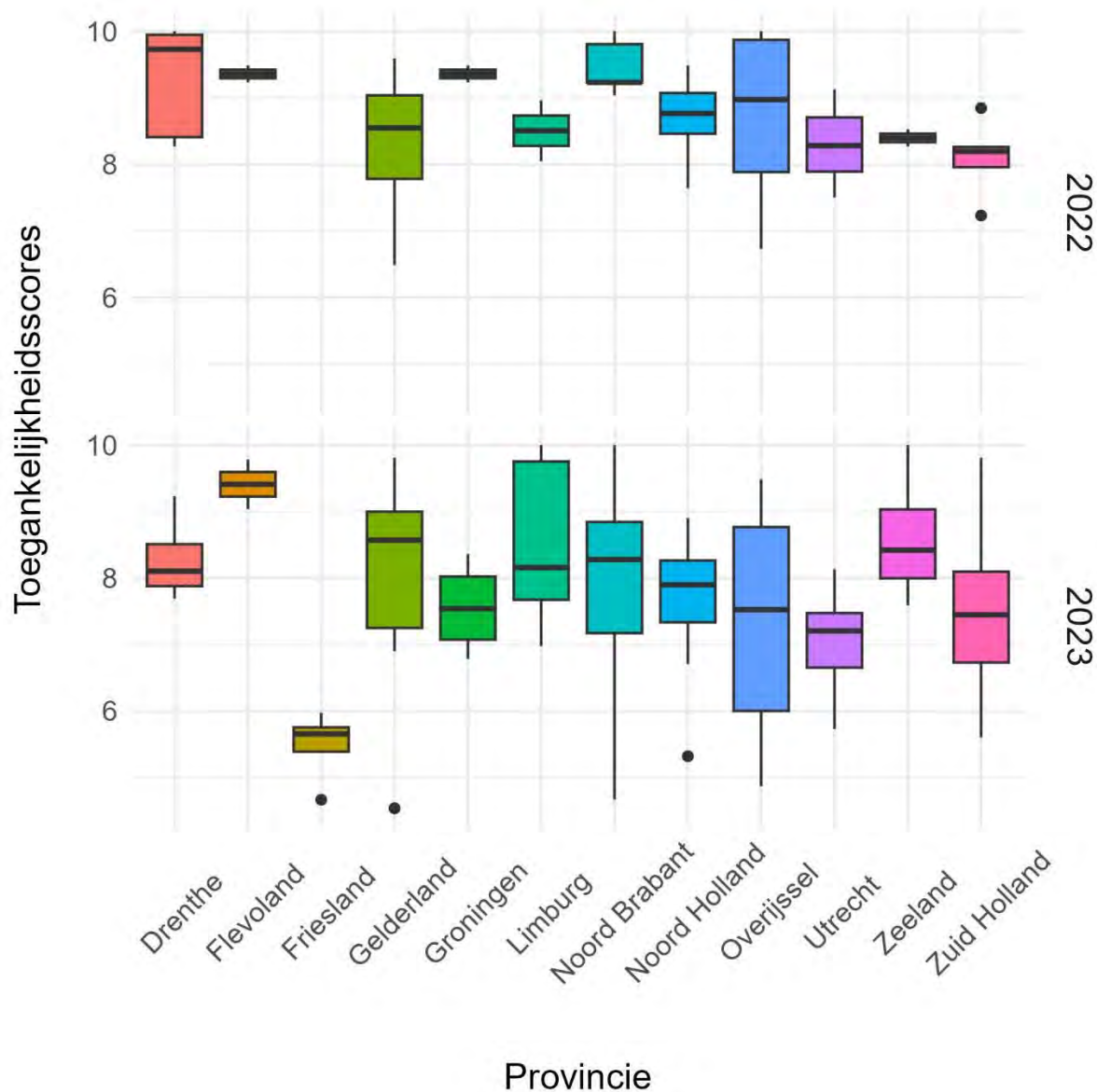
Er zijn redelijk grote verschillen te zien tussen de provincies in de verhouding van stemlokalen die wel en niet voldoen volgens de criteria van CITS. Friesland, Noord-Holland, Utrecht, en Zuid-Holland hebben verhoudingsgewijs lagere percentages van toegankelijke stemlokalen (minder dan 20%). Het totaal aantal bezochte stemlokalen van Flevoland (4) en Friesland (4) is echter laag, waardoor er een vertekend beeld kan ontstaan voor deze twee provincies. Via toegankelijkheidsscores in plaats van een categorisatie van wel of niet toegankelijk, kan een kritischer beeld worden geschetst.

Vervolgens wordt gekeken naar de kwantitatieve uitkomsten via de toegankelijkheidsscores. De provincies voor de twee steekproefonderzoeken (2022, 2023) worden hierna afzonderlijk beoordeeld.





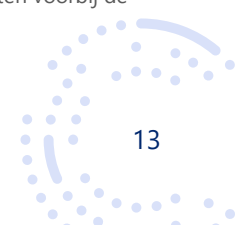
In **Figuur 3** wordt de verdeling van toegankelijkheidsscores van stemlokalen per provincie per jaar gevisualiseerd via een boxplot.¹⁴ De verdeling geeft inzicht in de spreiding van de scores van stemlokalen. De hoogte van een doos (grootte in de verticale richting) staat in verhouding tot hoe gespreid de scores van stemlokalen zijn op toegankelijkheid.



Figuur 3: Boxplot voor de spreiding van de toegankelijkheidsscores van stemlokalen per provincie en jaar.

De boxplot laat zien dat er grote verschillen zijn in de spreiding van de toegankelijkheidsscores van stemlokalen tussen de provincies. Er zijn ook een aantal uitschieters, te zien aan de zwarte punten. Ook hier zijn de aantallen van beoordeelde stemlokalen echter van belang. Een significantietest moet uitsluitsel geven op hoe verschillend de beoordelingen tussen de provincies zijn.

¹⁴ Een boxplot of een doosdiagram heeft drie horizontale strepen (onder, binnen, en boven) en twee verticale strepen. De drie horizontale strepen van de gekleurde dozen staan voor de drie kwartielen: 25%, 50%, en 75% (van onder naar boven). Het 50% kwartiel staat gelijk aan de mediaan. De verticale strepen buiten de dozen staan voor de inter quartile range (IQR) vermenigvuldigd met 1.5 en zijn begrensd door de uiterste datapunten die daar binnenin vallen. De zwarte punten voorbij de strepen zijn per definitie uitschieters in de data.





Significantietest

Een eenzijdige ANOVA toets is uitgevoerd op de data van de provincies. Volgens de ANOVA toets is er geen significant verschil in de gemiddelde toegankelijkheidsscores tussen de provincies in 2022. Er wordt een p-waarde van 0.0519 verkregen.

Voor 2023 is er wel een significant verschil in de variantie van de toegankelijkheidsscores, echter met een kanttekening. De verkregen p-waarde is 0.0022. Dit heeft te maken met de scores van provincie Friesland die veel afwijken van de rest. Friesland was niet meegenomen in de steekproef van 2022 en in 2023 zijn slechts vier stemlokalen beoordeeld, waarvan drie in gemeente Leeuwarden.

Zonder Friesland verandert het hele beeld. Door Friesland uit te sluiten zijn de verschillen tussen de provincies niet significant wat betreft de uitkomst van de ANOVA toets. De p-waarde stijgt naar 0.0901. Geen andere provincie, behalve Flevoland (waarbij het ook een klein aantal stemlokalen betreft), heeft zo'n sterke impact op de ANOVA toetsuitkomst.

Provincie Friesland uitgezonderd zijn er geen statistisch significante verschillen tussen de beoordelingen van de provincies. De afwijking door de provincie Friesland ligt waarschijnlijk aan de lage aantallen bezochte stemlokalen.

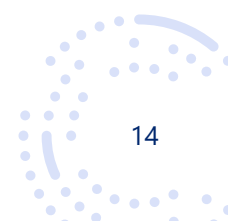
3.3 Verschillen tussen gemeentegroottes

Een statistische analyse op gemeenteniveau is niet goed mogelijk. Er zijn namelijk 25 gemeenten in de steekproef van 60 stemlokalen in 2022 en 41 gemeenten in de steekproef van 120 stemlokalen in 2023. Per gemeente zijn er maar enkele stemlokalen beoordeeld. Dit zijn te lage aantallen voor een goede statistische vergelijking. In plaats van te kijken naar de verschillen tussen individuele gemeenten, wordt er hier gekeken naar de verschillen tussen de drie gemeentegroottes. De drie gemeentegroottes zijn: klein (inwoners < 40.000), middel (inwoners 40.000 t/m 150.000), en groot (inwoners > 150.000). De aantallen worden per jaar uitgesplitst en worden weergegeven in **Tabel 2**.

Tabel 2: Beschrijvende statistiek van de drie gemeentegroottes.

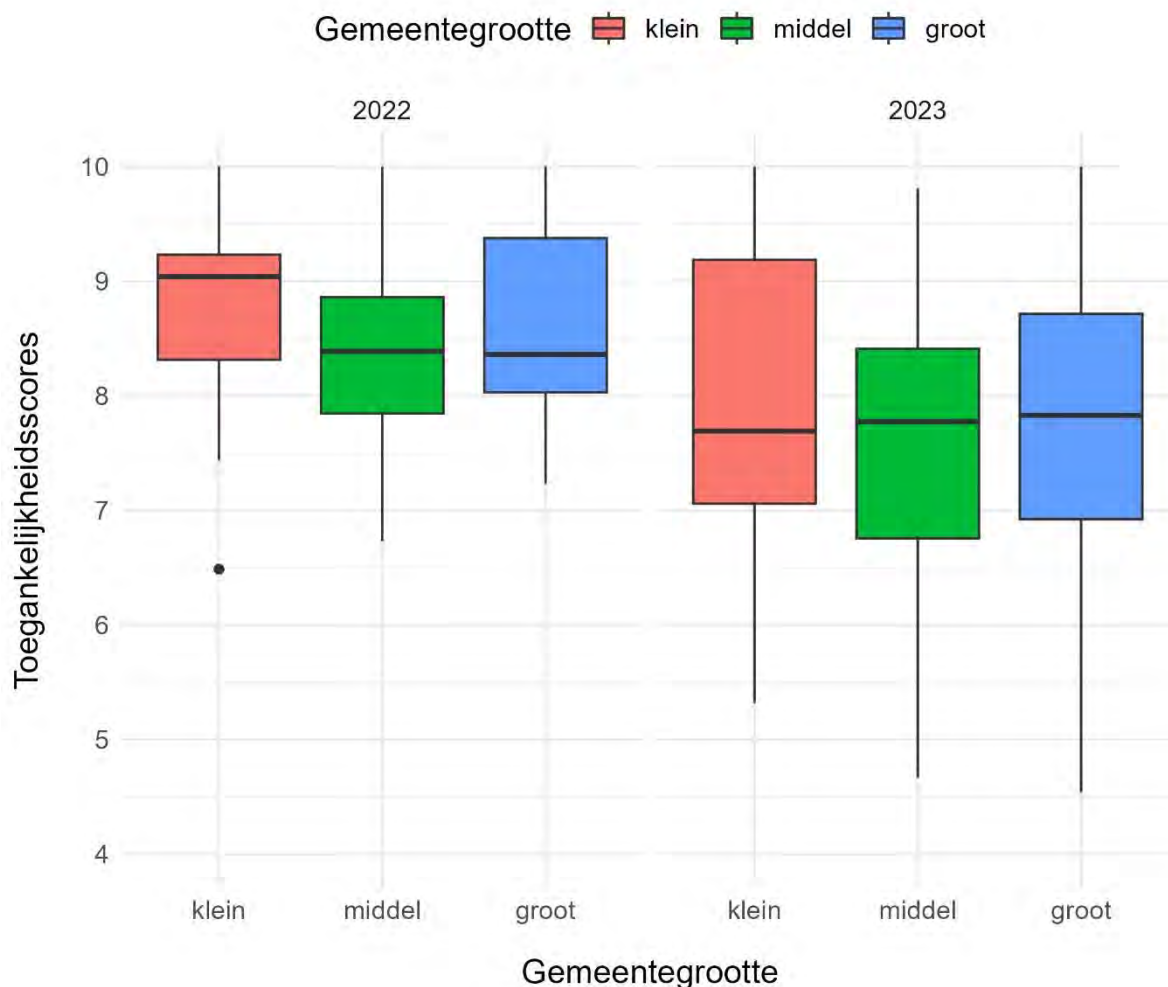
Gemeentegrootte (aantal inwoners)	Jaar	Aantal stemlokalen	Percentage voldoet	Toegankelijkheidsscore	Standaarddeviatie
Klein (< 40.000)	2022	35	34,3%	8,84	0,77
	2023	35	34,3%	7,95	1,41
Middel (40.000 – 150.000)	2022	15	46,7%	8,42	0,89
	2023	55	16,4%	7,54	1,23
Groot (> 150.000)	2022	10	30,0%	8,64	0,96
	2023	30	20,0%	7,79	1,25

Er zijn relatief meer kleine gemeenten in de steekproef van 2022 (35 van 60) dan in 2023 (35 van 120). Er zijn genoeg aantal stemlokalen in de uitsplitsingen van gemeentegrootte met een minimum aantal van 10. Ook hier wordt gekeken naar de toegankelijkheidsscores met de spreiding via een boxplot, dit is weergegeven in **Figuur 4**.





In de boxplot is te zien dat de verschillen tussen de gemeentegroottes per jaar klein zijn qua spreiding van de toegankelijkheidsscores. De kleine gemeenten van 2022 wijken iets af met een hoger gemiddelde en een hogere mediaan dan de rest. Daarom is een significantietest uitgevoerd (zie hieronder).

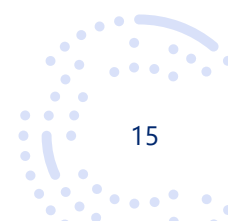


Figuur 4: Boxplot voor de spreiding van de toegankelijkheidsscores van stemlokalen per gemeentegrootte en jaar.

Significantietest

Een eenzijdige ANOVA toets is uitgevoerd op de data van de gemeentegroottes. Volgens de ANOVA toets is er geen significant verschil in de gemiddelde toegankelijkheidsscores tussen de drie gemeentegroottes in 2022 en ook geen in 2023. De verkregen p-waarden zijn respectievelijk 0.256 (2022) en 0.319 (2023).

Er zijn geen statistisch significante verschillen tussen de beoordelingen van kleine, middelgrote, en grote gemeenten.





3.4 Verschillen in beoordeling tussen controleurs

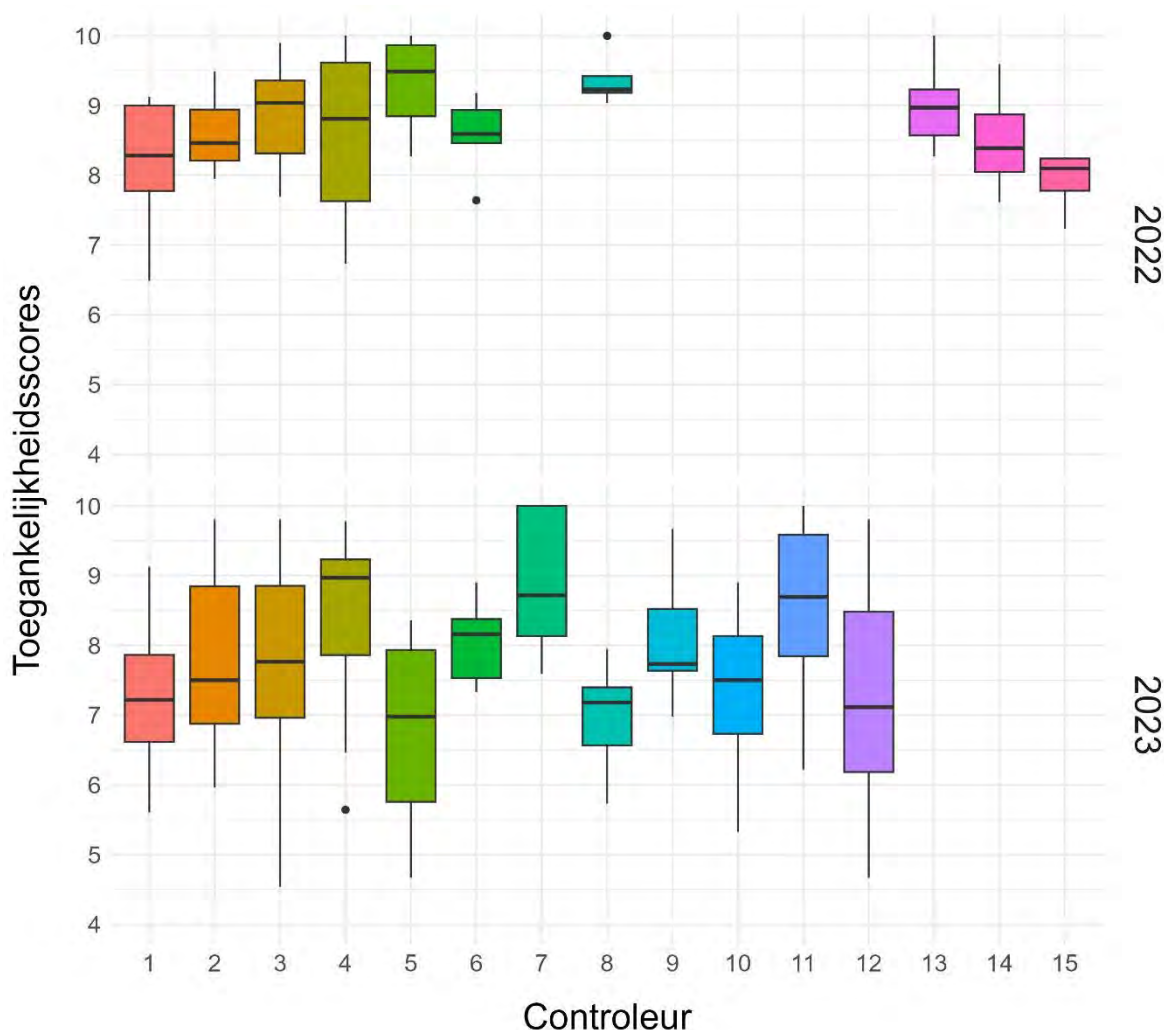
Er waren in 2022 tien verschillende controleurs die de stemlokalen beoordeelden. In 2023 waren dat er twaalf. Zeven controleurs waren betrokken bij beide steekproefonderzoeken. De aantallen bezochte stemlokalen per controleur zijn weergegeven in **Tabel 3**.

Tabel 3: Aantal bezochte stemlokalen per controleur.

Jaar	Controleur														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2022	7	7	7	8	7	5		4					6	5	4
2023	12	12	14	10	10	10	9	9*	10	9	8*	7*			

In het steekproefonderzoek van 2023 zijn 24 van de 120 stemlokalen door twee mensen beoordeeld. Dit is in de tabel aangegeven met een asterisk (*). Er wordt hiermee verder niets gedaan.

De spreiding van de toegankelijkheidsscores wordt weergegeven via een boxplot in **Figuur 5**. Er lijken verschillen te zijn tussen de controleurs. Ook hier is een significantietest uitgevoerd (zie hieronder).



Figuur 5: Boxplot voor de spreiding van de toegankelijkheidsscores van stemlokalen per controleur en jaar.





Significantietest

Een eenzijdige ANOVA toets is uitgevoerd op de data van de controleurs. Volgens de ANOVA toets is er geen significant verschil in de gemiddelde toegankelijkheidsscores tussen de controleurs in 2022. De verkregen p-waarde is 0.114.

Een kanttekening die hier geplaatst moet worden is dat er relatief weinig datapunten zijn per controleur. Het gemiddeld aantal bezochte stemlokalen per controleur is zes. Dit is minder dan het aantal dat in het rapport van prof. Wit wordt geadviseerd.

Voor de data van 2023 is er wel een significant verschil gevonden tussen de controleurs. De toets op variantie heeft een p-waarde van 0.0045. Via een meervoudige vergelijkingsprocedure met de *Tukey's HSD*¹⁰ statistische toets wordt duidelijk dat controleur 7 het meest afwijkt van de rest. Uitsluiten van controleur 7 levert alsnog een significantie op met een p-waarde van 0.0441. Dit geeft aan dat het probleem niet ligt aan een enkele controleur.

Een extra stap die is genomen is om te kijken naar alleen de 'ervaren' controleurs in 2023. Dit zijn de controleurs die ook mee hebben gedaan in 2022. Controleurs 1 tot en met 6 en controleur 8 waren betrokken bij beide steekproefonderzoeken. Door vijf van de twaalf controleurs uit te sluiten vermindert de dataset met bijna de helft voor deze analyse. Een analyse op alleen de 'ervaren' controleurs in 2023 levert een p-waarde van 0.0448. Ook met 'ervaren' controleurs is er een significant verschil te vinden. Hiermee versterkt het beeld dat er significante verschillen zijn tussen controleurs.

Een belangrijke noot is dat de gevonden verschillen niet direct te herleiden zijn naar de subjectiviteit van de beoordelingen. Omdat om praktische redenen iedere gemeente door een enkele controleur is beoordeeld, kunnen onderlinge verschillen tussen gemeenten nog impact hebben op de resultaten. Ook al zijn er geen significante verschillen gevonden tussen de gemeentegroottes (zie **hoofdstuk 3.3**), dan kunnen er nog steeds verschillen zijn tussen individuele gemeenten die doorwerken naar de gevonden verschillen tussen controleurs. Het is niet mogelijk om de verschillen tussen individuele gemeenten statistisch te toetsen vanwege de kleine hoeveelheid data op gemeenteniveau.

Er zijn statistisch significante verschillen gevonden in de beoordelingen van controleurs in 2023. Een nuancering hier is dat de significante verschillen tussen controleurs mogelijk beïnvloed zijn door onderlinge verschillen tussen gemeenten, omdat om praktische redenen iedere gemeente door een enkele controleur is beoordeeld.

3.5 Statistische verbanden data Ongehinderd

Om te zien of er afhankelijkheden zijn in de data van Ongehinderd wordt er gekeken naar de verbanden tussen de beoordelingen van de individuele toetsingsonderdelen van het CITS.⁶ De verbanden worden weergegeven aan de hand van statistische gemiddelden en standaarddeviaties van de toegankelijkheidsscores en aan de hand van correlaties (samenhang) tussen de beoordelingen van de verschillende toetsingsonderdelen.





Tabel 4 geeft de gemiddelde toegankelijkheidsscores met de standaarddeviaties tussen haakjes weer.

Tabel 4: Gemiddelde toegankelijkheidsscores per toetsingsonderdeel van CITS per jaar.

CITS onderdeel	Korte naamgeving	Gemiddelde score 2022	Gemiddelde score 2023
1.1	Openbare weg	9,21 (2,34)	8,63 (3,22)
1.2	Parkeergelegenheid	9,44 (1,52)	6,67 (4,73)
1.3	Route entree	6,83 (4,44)	5,42 (4,73)
1.4	Hoofdentree	5,53 (4,46)	4,11 (4,71)
2.1	Route stembureau	8,33 (3,60)	7,19 (4,19)
2.2	Tussendeur 1	8,88 (2,93)	9,77 (1,45)
2.3	Tussendeur 2	10 (0)	10 (0)
3.1	Deur stemruimte	7,61 (3,64)	7,56 (4,07)
3.2	Verkeersruimte	9,00 (3,03)	9,69 (1,67)
3.3	Stoelen	10 (0)	6,20 (4,61)
3.4	Stemhokje	9,36 (1,47)	8,67 (2,11)
3.5	Stembus	10 (0)	9,74 (1,21)
3.6	Kandidatenlijst	9,04 (1,31)	8,46 (1,87)
3.7	Leesloep	9,92 (0,60)	8,29 (3,30)

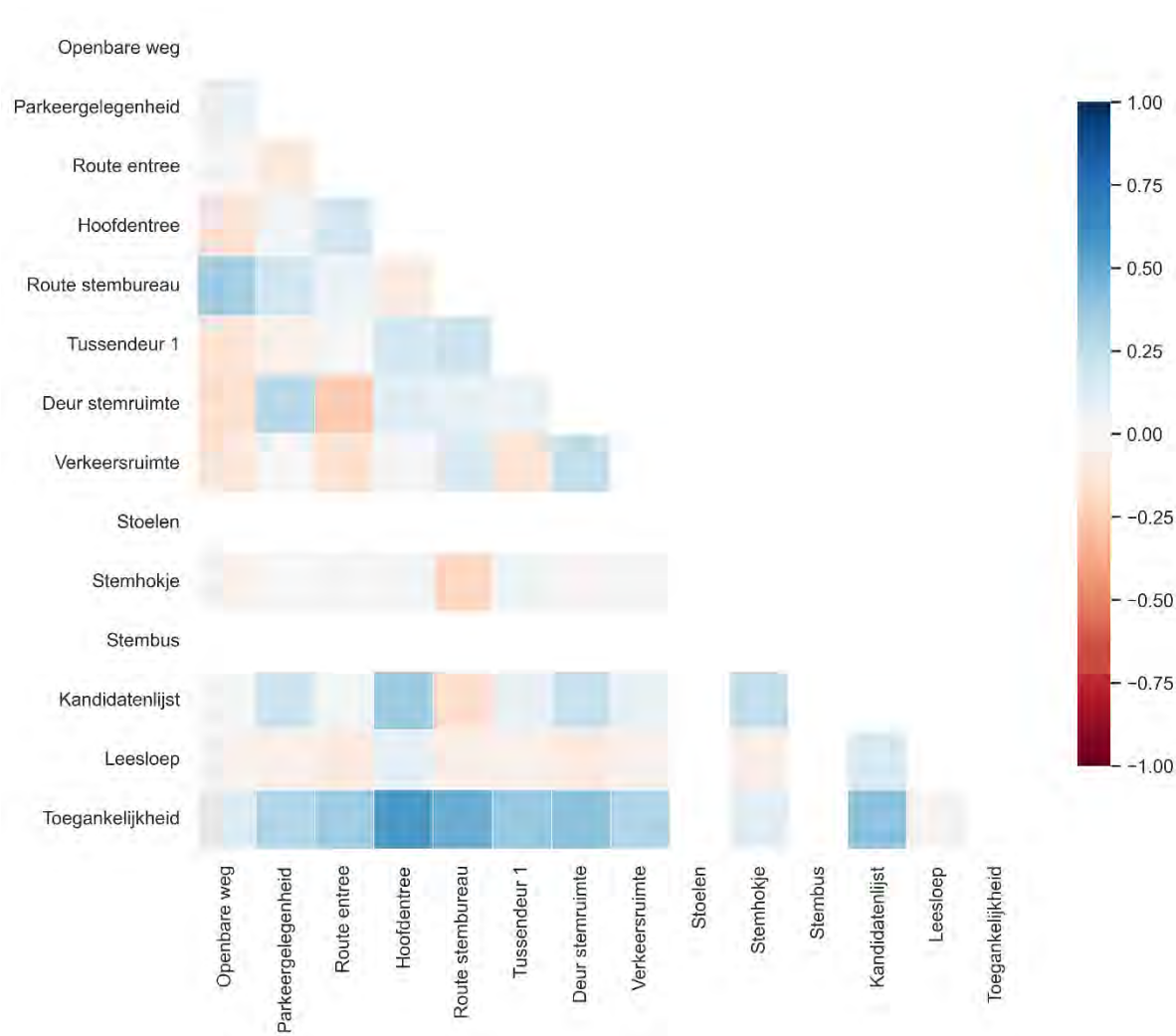
Opmerkelijk is dat sommige scores sterk zijn gedaald in 2023, zoals parkeergelegenheid en stoelen. Waar de twee onderdelen maximaal of bijna maximaal scoorden in 2022, daalde de gemiddelde score in 2023 tot bijna de helft. Bij stoelen gaat het om de beschikbaarheid van stoelen en het hebben van stoelen met armleuningen. In het steekproefonderzoek van 2022 werden de stoelen buiten beschouwing gelaten aangezien deze vanwege Covid maatregelen niet mochten worden gebruikt. Dit verklaart het verschil. Bij parkeergelegenheid gaat het om parkeervoorzieningen bij de entree en (bij voorkeur) uitgevoerd met een Gehandicapten Parkeerplaats Algemeen (GPA). De standaarddeviatie, een mate van spreiding, bij parkeergelegenheid is ook groter geworden, gestegen van 1,52 naar 4,73. Als laatste zien we dat onderdeel 3.7 'Leesloep' een veel groter standaarddeviatie heeft in 2023.

Om de verbanden tussen de toetsingsonderdelen inzichtelijk te maken, wordt een correlatieanalyse uitgevoerd. Correlaties geven inzicht in de statistische samenhang tussen de beoordelingen waarbij de sterkte van de samenhang wordt uitgedrukt in een correlatiecoëfficiënt. De waarde daarvan kan variëren tussen de -1 en +1. Een coëfficiënt van 0 betekent absoluut geen samenhang en een coëfficiënt van +1 of -1 betekent een perfecte positieve of respectievelijk negatieve samenhang. In de algemene zin wordt de grens getrokken bij een waarde van +0,7 of -0,7, waarboven of -onder er wordt gesproken over een matige tot sterke correlatie. Een voorbeeld van een perfecte negatieve samenhang (met coëfficiënt -1) is tussen de datasets van leeftijden en geboortejaren van mensen. Daar waar de een omhoog gaat, gaat de ander altijd omlaag. Via correlatie wordt er inzicht verleend in de (on)afhankelijkheid van de beoordelingen. Correlatie is echter geen bewijs van causaliteit, al kan het daar wel een aanwijzing voor zijn.

De *Pearson*¹⁵ correlatie tussen de toetsingsonderdelen onderling en tussen de toetsingsonderdelen en de toegankelijkheidsscore (Variabele "Toegankelijkheid") wordt weergegeven in **Figuur 6** en **Figuur 7**.

¹⁵ <https://wikistatistiek.amc.nl/Correlatie>





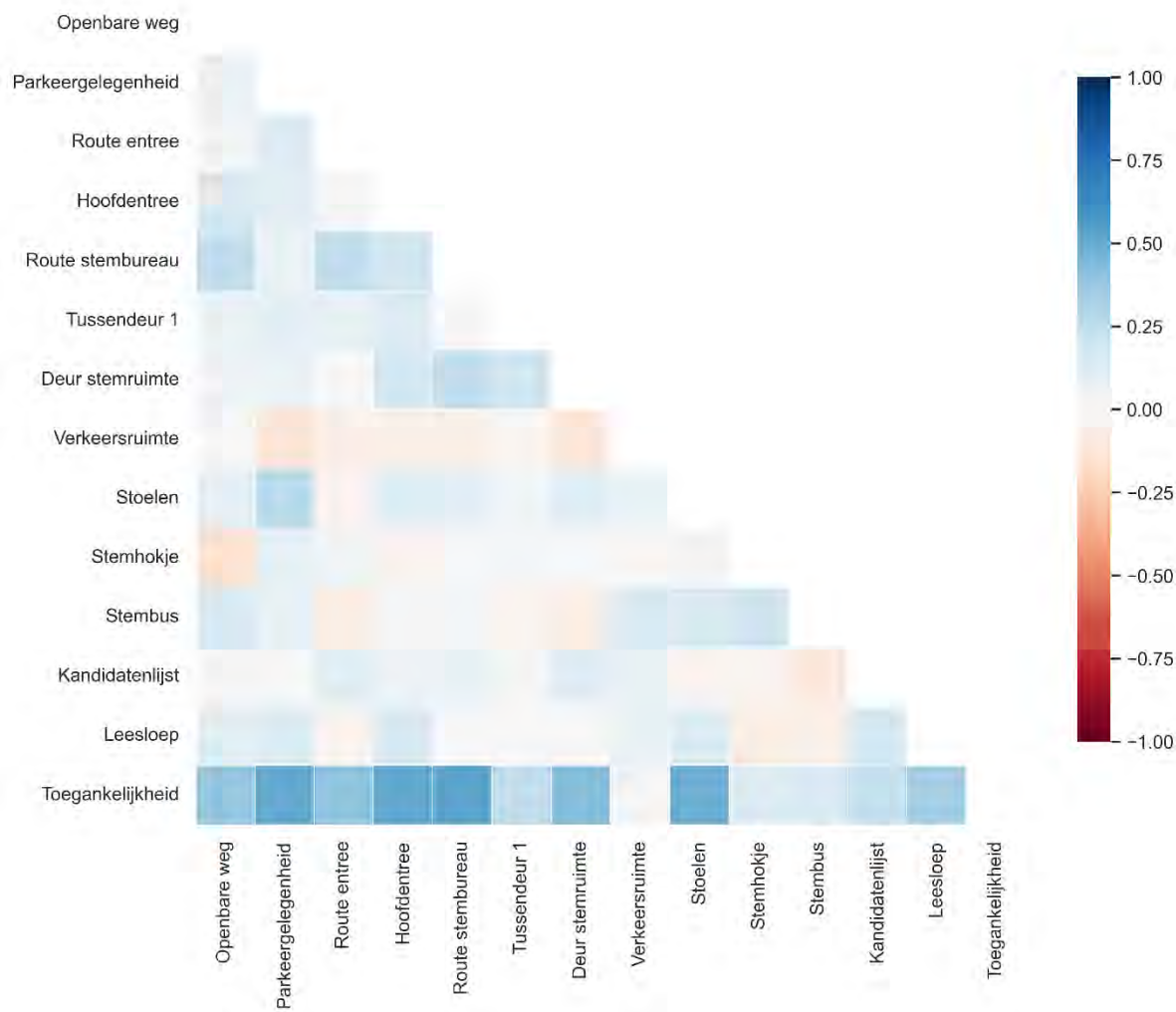
Figuur 6: Correlatieplot van toetsingsonderdelen en toegankelijkheidsscore in 2022.

In de figuren staat blauw (+1) voor maximaal positief gecorreleerd, rood (-1) voor maximaal negatief gecorreleerd, en wit (0) voor helemaal niet gecorreleerd.

CITS onderdeel 2.3 ("Tussendeur 2") is niet meegenomen in de correlatieanalyse omdat deze in beide jaren maximaal scoorden bij alle stemlokalen. In 2022 scoorden daarnaast onderdelen 3.3 ("Stoelen") en 3.5 ("Stembus") ook maximaal zoals ook weergegeven in **Tabel 4**.

Het is te zien aan de kleuren in de figuren dat er geen sterke verbanden zijn tussen de toetsingsonderdelen en de toegankelijkheidsscore (zie onderste rij). De toegankelijkheidsscore bij het steekproefonderzoek van 2022 hangt het sterkst af van de beoordeling "Hoofdentree". Dit heeft een correlatiecoëfficiënt van 0,57. In 2023 hangt de toegankelijkheidsscore het sterkst af van "Route stembureau" met correlatiecoëfficiënt van 0,54.¹⁶ Dit zijn geen sterke correlaties.

¹⁶ Een correlatie van 0 staat voor helemaal niet gecorreleerd, dus geen samenhang, en een correlatie van 1 staat voor volledig positief gecorreleerd, dus maximaal samenhang. De grens voor sterke correlaties wordt vaak gelegd op een coëfficiënt van 0,7.



Figuur 7: Correlatieplot van toetsingsonderdelen en toegankelijkheidsscore in 2023.

Ook is te zien dat de correlaties tussen de toetsingsonderdelen onderling niet sterk zijn. De sterkste correlatie is tussen "Kandidatenlijst" en "Hoofdentree" in 2022 met een coëfficiënt van 0,36. In 2023 is dit tussen "Stoelen" en "Parkeergelegenheid" met een coëfficiënt van 0,27. Dit zijn zwakke correlaties.

Er zijn geen sterke verbanden gevonden tussen de CITS toetsingsonderdelen en eindbeoordeling. Er zijn ook geen sterke verbanden gevonden tussen de toetsingsonderdelen onderling.





4 Steekproeftrekking

Het onderzoek naar de toegankelijkheid van stemlokalen wordt gedaan op een steekproef van gemeenten. Sinds 2023 worden voor de steekproeftrekking de adviezen van prof. Wit gevolgd.³ Dit betekent dat er voor het selecteren van de stemlokalen verhoudingsgewijs geselecteerd wordt, lettend op de inwonersaantallen binnen provincies en gemeenten. Hiermee wordt er rekening gehouden met een goede afspiegeling van de landelijke situatie.

Centerdata is door BZK gevraagd om de steekproeftrekking voor de komende verkiezingen te doen op de wijze die is omschreven in het rapport van prof. Wit. Er zijn echter twee beperkingen opgelegd. De eerste is dat de gemeenten die voor de steekproef van de afgelopen verkiezingen waren geselecteerd uitgesloten moeten worden van herselectie. Dit om de last te minimaliseren voor gemeenten wat betreft het onderzoek naar toegankelijkheid. De tweede beperking is dat de gemeenten per provincie zo mogelijk dicht in de buurt van elkaar moeten liggen, zodat het voor controleurs haalbaar is om alle gemeenten op de dag van de verkiezingen te bezoeken.

Na overleg met BZK is er een aanpassing gedaan aan een van deze beperkingen. Bij het uitsluiten van de gemeenten die in de vorige steekproef zaten is besloten dat de grote gemeenten (met een inwonersaantal van meer dan 150.000) niet zullen worden uitgesloten van een hertrekking. Dit omdat het aantal grote gemeenten in Nederland relatief laag is. Door het uitsluiten van een relatief hoog percentage van deze gemeenten kan er een steekproefbias optreden.

Er is een algoritme geschreven dat de steekproeftrekking automatisch doet. Het algoritme neemt de genoemde wensen en beperkingen in beschouwing.

4.1 Het algoritme

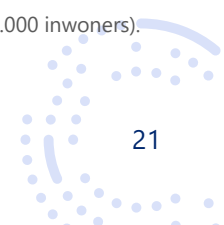
Een relatief simpel en herbruikbaar algoritme om de steekproeftrekking te doen is geschreven met de programmeertaal Python. Als input vraagt het algoritme het gewenste aantal stemlokalen aan te geven. In dit geval gaat het om 120 stemlokalen, maar dit aantal kan worden aangepast. De vervolgstap is dat het algoritme het aangegeven aantal stemlokalen verdeelt onder de 12 provincies naar rato van de populatie van de provincies.¹⁷

Nadat de aantallen per provincie bekend zijn, kan het algoritme verder met het selecteren van de gemeenten. Eerst wordt er een bestand ingelezen om eerder beoordeelde gemeenten uit te sluiten van de steekproeftrekking. Dit bestand kan naar wens worden aangepast. De grote gemeenten in de lijst worden echter niet uitgesloten van deelname, omdat anders het aantal grote gemeenten dan drastisch daalt. Voor de huidige steekproeftrekking zijn in totaal 40 gemeenten uitgesloten.

Het algoritme selecteert om te beginnen de gemeentegrootte categorieën¹⁸ op basis van een waarschijnlijkheid die is gebaseerd op inwonersaantallen van gemeenten. Wanneer een

¹⁷ De inwonersaantallen van 2023 zijn als basis genomen, maar kunnen in de toekomst aangepast worden door het veranderen van de ingelezen Excel-file met nieuwe inwonersaantallen.

¹⁸ Er zijn drie gemeentegroottes, klein (< 40.000 inwoners), middel (40.000 – 150.000 inwoners), en groot (> 150.000 inwoners).





gemeentegrootte is geselecteerd, wordt uit de subset van mogelijke gemeenten (dus gegeven de gemeentegrootte, provincie, en degenen die niet uitgesloten zijn) een willekeurig gemeente gekozen. Per kleine gemeente worden er twee, per middelgrote gemeente worden er vier, en per grote gemeente worden er 8 stemlokalen aangewezen. Deze stappen volgen de beschreven wijze van het rapport van prof. Wit.

Een nuance is dat bij het selecteren van de gemeentegroottes er volgens een waarschijnlijkheid wordt gewerkt in plaats van een afronding. Het rapport van prof. Wit rondt de aantallen stemlokalen af op hele getallen. Volgens de methode van prof. Wit betekent dit dat uit de provincies Groningen en Flevoland nooit een grote gemeente geselecteerd kan worden, omdat de kans daarvoor te klein is. Het aantal wordt dus afgerond naar nul. In het algoritme is er wel een kans voor, ook al is die klein.

4.2 De steekproeftrekking

De steekproeftrekking resulteert in 35 gemeenten. 18 hiervan zijn klein, 12 middelgroot, en 5 groot. De gedetailleerde verdeling van de gemeenten is weergegeven in **Tabel 5**.

Tabel 5: De verdeling van de steekproeftrekking.

Provincie	Aantal stemlokalen	Klein (2x)	Middel (4x)	Groot (8x)	Provincie	Aantal stemlokalen	Klein (2x)	Middel (4x)	Groot (8x)
ZH	26	1	2	2	LB	8	4	0	0
NH	20	2	2	1	FR	4	0	1	0
NB	18	1	4	0	GR	4	0	1	0
GD	14	1	1	1	DR	3	2	0	0
UT	9	1	0	1	FL	3	2	0	0
OV	8	2	1	0	ZL	3	2	0	0

De tweede opgelegde beperking van BZK was dat de gemeenten per provincie zo mogelijk dicht in de buurt van elkaar moeten liggen, zodat het voor controleurs haalbaar is om alle geselecteerde gemeenten op de dag van de verkiezingen te bezoeken. Omdat per provincie het aantal niet-uitgesloten gemeenten beperkt is, is er weinig speling voor deze extra beperking. Algoritmisch gezien is er geen geprogrammeerde limitatie opgelegd qua afstand tussen de geselecteerde gemeenten.

De lijst met gemeenten van de steekproeftrekking voor de aankomende verkiezingen is als dataset opgeleverd op 17 juli 2023.

