

Vrije Universiteit Amsterdam

Faculteit der Bètawetenschappen (Faculty of Science)

Bachelor Aarde, Economie & Duurzaamheid

Tegen de Grenzen van het Stroomnet: Ruimtelijke drijfveren en de impact van netcongestie op recente uitbreidingen van grote logistieke centra in Nederland (2020- 2024)

Concept Bachelor Thesis

Aantal woorden: 7969

Auteur: Steyn Bosch

Studentnummer: 2822302

Email: s.j.bosch@student.vu.nl

Eerste Beoordelaar: Eric Koomen

Tweede Beoordelaar: Maureen Lankhuizen

Datum van Inlevering: 11-06-2026

Plaats: Amsterdam

Abstract

Dit onderzoek analyseert de ruimtelijke drijfveren achter de logistieke schaa sprong in Nederland gedurende de periode 2020-2024, waarbij expliciet onderscheid wordt gemaakt tussen middelgrote (L/XL) en megagrote (XXL) distributiecentra. Aan de hand van een ruimtelijke gridanalyse en een Multinomiaal Logit Model (MNL) is getoetst hoe bereikbaarheid, clustering, ruimtelijk beleid en fysieke beperkingen de locatiekeuze beïnvloeden. De resultaten tonen aan dat L/XL-panden integreren in bestaande stedelijke netwerken, terwijl XXL-panden door ruimtegebrek uitwaaiëren naar de periferie (logistics sprawl). Daarnaast blijkt actuele netcongestie te fungeren als een harde ruimtelijke barrière, die de vestigingskans van XXL-panden disproportioneel met 70% verlaagt ten opzichte van congestievrije gebieden. Hiermee overschaduwen overheidssturing en fysieke capaciteitsgrenzen in toenemende mate de klassieke transportgeografie.

Inhoud

Abstract.....	2
1. Inleiding.....	4
1.1 Aanleiding en Context	4
1.2 Probleemstelling & Research Gap	4
1.3 Onderzoeksvragen	5
1.4 Hypothesen	5
2. Theoretisch Kader	6
2.1 Locatietheorie en Logistiek.....	6
2.2 Heterogeniteit en Logistics Sprawl	6
2.3 De Rol van Ruimtelijk Beleid	7
2.4 Frictie tussen Beleid, Techniek en Markt: De Nieuwe Realiteit	8
3 Data en Onderzoeksmethode	9
3.1 Dataverzameling en Actualisering	9
3.2 Classificatie en Validatie	11
3.2.1 SQL-Statusselectie en de LISA-dominantieregels	11
3.2.2 Eigen Geautomatiseerde Algoritmes: Spookpanden- en Kassen-killer	12
3.2.3 De Erfenis-Audit en Risico-trechter	12

3.2.4	Handmatige Validatie via de GIS/Shiny Interface	13
3.2.5	De Definitieve Logistieke Voorraad (Stand eind 2024)	13
3.3	Verklaren Ruimtelijke Drijfveren	15
3.3.1	Selectie variabelen	16
3.3.2	Econometrische verantwoording en uitsluiting van variabelen	18
4.	De Logistieke Schaalsprong: Ruimtebeslag 2020-2024.....	19
4.1	Het Logistieke Ruimtebeslag	19
4.2	Geografische patronen en Logistics Sprawl (Verschillen tussen L en XXL)	21
5	Drijfveren van Ontwikkeling	22
5.1	Interpretatie van het Multinomiaal Logit-Model	23
5.2	Impact van Bereikbaarheid en Locatiefactoren	24
5.3	De Effectiviteit van Ruimtelijk Beleid en Omgevingsfactoren	25
5.4	Ruimtelijke Kansenallocatie en Modelvoorspellingen	26
6	Conclusie en Discussie.....	28
6.1	Conclusie	28
6.2	Theoretische Reflectie en Discussie.....	29
6.3	Methodologische Beperkingen	29
6.4	Beleidsaanbevelingen en Toekomstig Onderzoek	30
6.5	Gebruik van AI	31
	Literatuurlijst	32
	Bijlages.....	37
	Bijlage A: De Batch-scripts (OSGeo4W Shell & PowerShell)	37
	Bijlage B: Geautomatiseerde Workflow in R (scripts 1 t/m 3).....	38
	Bijlage C: Visuele Verificatie Logistiek Vastgoed	46
	Bijlage D: Logistics Building Review App - Setup Guide	48
	Bijlage E: Opzet MNL Model (Scripts 4 t/m 6)	55

1. Inleiding

1.1 Aanleiding en Context

De sterke groei van de logistieke sector in Nederland is een direct gevolg van de voortdurende fragmentatie van mondiale productieketens en versnelde adoptie van e-commerce (Tare et al., 2024). Sinds 2020 is de vraag naar logistiek vastgoed verder gegroeid, wat heeft geleid tot een intensief maatschappelijk debat over de ‘verdozing’. De logistieke groei veroorzaakt toenemende beleidsmatige frictie met betrekking tot het fysieke ruimtebeslag en de landschappelijke kwaliteit (College van Rijksadviseurs, 2019; Nefs & Daamen, 2022; Savills, 2025). Hoewel logistiek vastgoed essentieel is voor de nationale economie, waarschuwen adviesorganen voor de onomkeerbare effecten op de ruimtelijke kwaliteit en de toenemende claim op schaarse ruimte in Nederland (PBL, 2026; Rli, 2016).

1.2 Probleemstelling & Research Gap

Hoewel de logistieke sector in 2021 is gemonitord (Nefs, 2023), vertoont het wetenschappelijke inzicht in de recente ontwikkelingen nog hiaten op de volgende vlakken:

- **Actualiteit en dynamiek:** De meest recente geodataset (Nefs, 2023) is gebaseerd op data tot uiterlijk 2021. De veranderingen in keuze van locatie, duurzaamheid en schaalgrootte (Savills, 2025), in de periode 2021-2025 zijn wetenschappelijk nog niet in kaart gebracht.
- **Heterogeniteit in de verzadigde markt:** Eerdere studies behandelen logistiek vaak als een homogene entiteit (Tare et al., 2024). Er is een gebrek aan inzicht in hoe de locatievoorkeuren verschillen tussen middelgrote (L) en de sterk groeiende XXL-panden (>40.000 m²) in een markt waar in traditionele hotspots de opname is gedaald (Savills, 2025).
- **Fysieke restricties:** Huidige modellen houden nog geen rekening met relatief nieuwe, harde grenzen aan de groei, zoals de wijdverspreide netcongestie die sinds 2022 sturend is geworden in de ruimtelijke ordening en ontwikkeling van XXL-Distributiecentra (CE Delft, 2024).

De locatiekeuze die in dit onderzoek wordt geanalyseerd, is veranderd van een puur economische afweging naar de vraag of een project überhaupt nog wel gebouwd kan worden. Door de frictie tussen economische groei en fysieke, infrastructurele grenzen kwantitatief te analyseren sluit dit onderzoek aan bij het perspectief van Aarde, Economie en Duurzaamheid (AED). Met als specifieke uitbreiding dat de netcongestie is toegevoegd, hiermee brengt deze

scriptie de frictie in kaart tussen de groei van logistiek en harde capaciteitsgrenzen in Nederland.

1.3 Onderzoeksvragen

De hoofdvraag van dit onderzoek luidt: *In welke mate verschillen de ruimtelijke drijfveren voor locatiekeuze van nieuwe XXL-distributiecentra (>40.000 m²) van de drijfveren van middelgrote logistieke panden (L/XL: 10.000 – 40.000 m²) in de periode 2020-2024?*

De focus ligt hierbij op L/XL- en XXL-panden, aangezien juist deze grootschalige ontwikkelingen verantwoordelijk zijn voor de actuele ruimtelijke en infrastructurele capaciteitsproblemen.

Ter ondersteuning van de hoofdvraag worden de volgende deelvragen beantwoord:

1. Hoe heeft het logistieke ruimtebeslag in de categorieën L/XL en XXL zich ruimtelijk ontwikkeld tussen 2020 en 2024?
2. Wat is de relatieve invloed van bereikbaarheidsfactoren en locatiefactoren op de vestigingskansen van deze verschillende maatklassen?
3. Welke rol spelen ruimtelijk beleid en fysieke restricties, in het bijzonder netcongestie, bij de huidige spreiding van deze specifieke maatklassen?

1.4 Hypothesen

Op basis van de locatietheorie en empirische bevindingen uit eerdere studies worden de volgende hypothesen getoetst. Deze hypothesen zijn gebaseerd op de locatietheorieën en het concept van logistics sprawl, welke in het navolgende theoretisch kader (Hoofdstuk 2) uitgebreid worden toegelicht.

- **H1 Toegankelijkheid:** XXL-panden vertonen een grotere afhankelijkheid van multimodale knooppunten en snelwegnetwerken dan L/XL-panden, vanwege de enorme schaal van hun goederenstromen.
- **H2 Ruimtebeslag en Landwaarde:** XXL-ontwikkelingen zijn gevoeliger voor hoge grondprijzen. Zij wijken sneller uit naar goedkopere periferie-locaties dan L/XL-panden, wat bijdraagt aan 'logistics sprawl'.
- **H3 Beleid:** Restrictieve zonering remt de ontwikkeling van logistiek significant, waarbij XXL-panden door actieve weigering op ongewenste locaties vaker door provinciaal beleid worden beperkt dan L/XL-panden.

- **H4 Fysieke Restricties:** Acute netcongestie fungeert post-2022 als een harde ruimtelijke barrière voor nieuwe logistieke ontwikkelingen. XXL-panden worden door de elektrificatie van vrachtwagens en procesautomatisering harder geraakt dan L/XL panden, waardoor zij vaker gedwongen worden uit te wijken naar suboptimale locaties buiten overbelaste netwerkzones.

2. Theoretisch Kader

Dit hoofdstuk geeft een beknopt overzicht van de wetenschappelijke literatuur over de locatiekeuze voor logistiek vastgoed. Allereerst worden klassieke locatietheorieën rondom transport en agglomeratie besproken. Vervolgens wordt ingegaan op het fenomeen logistics sprawl, waarna tot slot het stimulerend en restrictief beleid worden belicht.

2.1 Locatietheorie en Logistiek

Klassiek gezien bepalen transportkosten, grondprijzen en de voordelen van clustering (agglomeratievoordelen) waar een distributiecentrum wordt gebouwd (Tare et al., 2024). Verschillende onderzoeken laten zien dat een goede bereikbaarheid via water, spoor en snelwegen historisch gezien doorslaggevend zijn (Bowen, 2008; Holl & Mariotti, 2018). Hiernaast is ook de afstand tot de consument erg belangrijk, dit weegt door snelle leveringen binnen e-commerce tegenwoordig extra zwaar mee (Nguyen & Sano, 2010; Tare et al., 2024).

Daarnaast spelen de voordelen van clustering een grote rol (van den Heuvel et al., 2013). Logistieke bedrijven gaan vaak bewust dicht bij elkaar zitten om hier samen van te profiteren. Denk hierbij aan de toegang tot een grote, gespecialiseerde groep werknemers en het delen van specifieke vakkennis (van den Heuvel et al., 2013). Wel laat praktijkonderzoek zien dat deze voordelen in de ruimte snel afzakken (spatial decay): hoe verder een bedrijf buiten de kern van het logistieke knooppunt zit, hoe minder het merkt van deze positieve effecten (Rosenthal & Strange, 2003).

2.2 Heterogeniteit en Logistics Sprawl

De logistieke sector moet niet als een homogene entiteit worden beschouwd (Sakai et al., 2020; Tare et al., 2024). Afhankelijk van de specifieke positie in de supply chain en de schaalgrootte van een faciliteit, verschillen de locatie-eisen en de landwaarde-elasticiteit drastisch. Middelgrote distributiecentra (L/XL) en logistiek vastgoed met een sterke retail- of e-fulfilmentfunctie zoeken veelal naar locaties nabij stedelijke kerngroengebieden om de 'last mile' distributie te optimaliseren. Omdat zij de nabijheid tot de eindconsument prioriteren, hebben zij een hogere betalingsbereidheid voor schaarse grond en accepteren zij de hogere hedonische grondprijzen die inherent zijn aan centrale locaties (Sakai et al., 2015).

Daar tegenover staat de categorie van XXL-panden, die wordt gekenmerkt door een exceptioneel groot, aaneengesloten ruimtebeslag en een sterke afhankelijkheid van grootschalige (internationale) macro-hotspots. Vanwege de wetten van de schaalvergroting en de noodzaak tot kapitaalintensieve mechanisatie zijn XXL-ontwikkelingen extreem gevoelig voor hoge landwaarden (Nefs & Daamen, 2022; Tare et al., 2024). Wanneer de grond in de traditionele kernregio's te duur of schaars wordt, wijken deze logistieke reuzen uit naar goedkopere, minder gecongesteerde gebieden in de periferie.

Deze uitwaaiing staat bekend als logistics sprawl (Dablanc & Ross, 2012; Heitz et al., 2020). Juist deze grootschalige deconcentratie intensificeert het conflict rondom de 'verdozing', omdat de panden opduiken in voorheen open landschappen (College van Rijksadviseurs, 2019). Onderzoek waarschuwt voor de negatieve maatschappelijke gevolgen hiervan, zoals extra vrachtwagenkilometers en landschapsdegradatie (Aljohani & Thompson, 2016).

2.3 De Rol van Ruimtelijk Beleid

Vanuit het perspectief van de duurzame ruimtelijke ordening bevindt de Nederlandse overheid zich in een complexe spagaat tussen economische belangen en ecologische randvoorwaarden. Decennialang werd de sector vanuit het Rijk actief gefaciliteerd via het Mainport- en Topsectorenbeleid, wat heeft geleid tot diepe ruimtelijke pad afhankelijkheid en zware concentraties in logistieke hotspots zoals Venlo, Midden-Brabant en de haven van Rotterdam (Tare et al., 2024). De cumulatieve druk op leefomgeving heeft echter geleid tot verschuiving in het ruimtelijk beleid richting restrictieve zonering en deconcentratiebeheersing.

Met de Ontwerp-Nota Ruimte (2026) kiest het Rijk voor hernieuwde, centrale regie waarbij de eis tot meervoudig ruimtegebruik en landschappelijke inpassing leidend is. Solitaire XXL-dozen op onontgonnen locaties (greenfields) worden beleidsmatig aan banden gelegd om de versnippering van het landschap tegen te gaan. De Commissie voor de milieueffectrapportage (2026) benadrukt echter dat de institutionalisering van dit restrictieve beleid vaak botst met de vigerende lokale bestemmingsplannen, wat leidt tot een frictie tussen nationale doelen en lokale uitvoering.

Op sub-nationaal niveau vertaalt dit zich in harde, dwingende maatregelen. De provincie Noord-Brabant heeft via een voorbereidingsbesluit de bouw van nieuwe XXL-logistiek buiten de grote, bovenregionale bedrijventerreinen tijdelijk stilgelegd (Provincie Noord-Brabant, 2023). Ook binnen Gelderland stellen de Regionale Programma's Werklocaties (RPW's) strenge grenzen en plafonds. Zowel de Groene Metropoolregio Arnhem-Nijmegen (2024) als de regio Kop van de Veluwe (2024) remmen grootschalige, solitaire distributiecentra af. In

plaats van nieuwe grond uit te geven, geven zij in hun plannen prioriteit aan het beter benutten en herstructureren van bestaande terreinen.

2.4 Frictie tussen Beleid, Techniek en Markt: De Nieuwe Realiteit

Het plan van de overheid om XXL-logistiek te concentreren op enkele strategische locaties, stuit in de praktijk op markttechnische barrières en leidt tot disrupties in de bestaande logistieke netwerken (Planbureau voor de Leefomgeving, 2026; Savills, 2025). Klassieke locatietheorieën (zoals de theorie van ruimtelijke clustering) voorspellen dat te veel concentratie op een plek uiteindelijk omslaat in negatieve agglomeratie-effecten, in de literatuur agglomeration diseconomies genoemd (Graham, 2007; Melo et al., 2009). Als zo'n kritisch punt wordt overschreden, wegen de voordelen van ruimtelijke clustering (agglomeratievoordelen) niet langer op kan tegen de nadelen. Denk hierbij aan grote files op de lokale wegen en een extreme krapte op de regionale arbeidsmarkt (Graham, 2007; Melo et al., 2009).

De gevolgen van deze theoretische omslag zijn inmiddels te zien op de Nederlandse vastgoedmarkt. Het aandeel van nieuwe logistieke panden in de traditionele hotspots (zoals Venlo en Midden-Brabant) daalt significant, terwijl de landelijke leegstand begin 2025 is gestegen tot 5,7% door een herstructurering van logistieke netwerken (Savills, 2025).

Daarnaast wordt de locatiekeuze voor grootschalige logistiek na 2024 niet meer alleen bepaald door bereikbaarheid of grondprijzen, maar in toenemende mate gedicteerd door harde fysieke infrastructuurgrenzen (CE Delft, 2024). Recente sectoranalyses tonen aan dat netcongestie inmiddels is uitgegroeid tot een sturende ruimtelijke beperking (CE Delft, 2024; Enexis Netbeheer, 2024).

Netbeheerders kampen met een structurele transportschaarste en aanzienlijk maakbaarheidsgat door een tekort aan technici, waardoor de uitbreiding van de Elektriciteits- en Hoogspanningsnetten (HS-netten) zwaar vertraagd is (Enexis Netbeheer, 2024). Omdat nieuwe XXL-distributiecentra door vloot-elektrificatie en automatiseringssystemen transformeren tot industriële grootverbruikers, dwingt het capaciteitstekort hen tot complexe, kostbare ingrepen (CE Delft, 2024). Bedrijven zijn in toenemende mate afhankelijk van mitigerende maatregelen, zoals lokale batterij-opslag of het vormen van gesloten energiehubbs (CE Delft, 2024). Dit maakt de locatiekeuze fundamenteel anders dan klassieke ruimtelijke modellen voorspellen, aangezien stroombeschikbaarheid de traditionele locatievoordelen inmiddels volledig overschaduwde. Deze actuele, empirische realiteit toont aan dat de bestaande academische locatiemodellen herzien moeten worden.

In recente studies, zoals die van Tare et al. (2024) en Heitz et al. (2020), wordt de logistieke expansie nog vrijwel uitsluitend verklaard vanuit traditionele transportkosten en agglomeratievoordelen. De fysieke remmende werking van energie-infrastructuur (netcongestie) op locatiekeuze is in deze kwantitatieve modellen tot op heden nog nauwelijks onderzocht. Dit zorgt voor een belangrijk gat in onze kennis. Dit probleem wordt versterkt door de stikstofcrisis en Europese Natura 2000-wetgeving (Planbureau voor de Leefomgeving, 2026). Strengere juridische restricties op onontgonnen locaties (greenfields) veranderen de locatiekeuze van een puur economische afweging in een vraag of een project planologisch überhaupt haalbaar is.

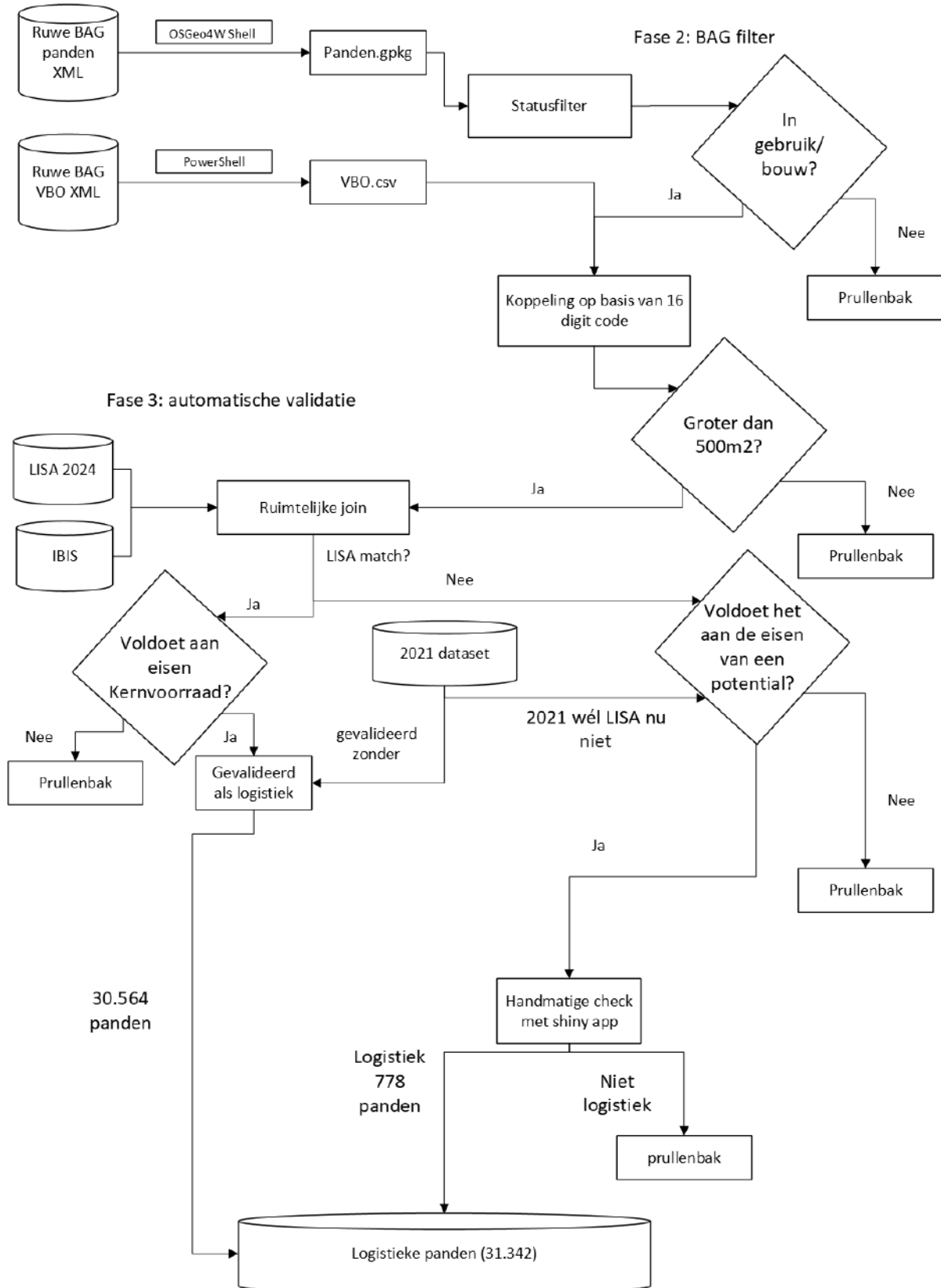
3 Data en Onderzoeksmethode

3.1 Dataverzameling en Actualisering

Voor dit onderzoek is de Monitor Logistiek Vastgoed (Nefs, 2023) geactualiseerd. Waar de eerdere dataset gebruik maakte van data tot 2021, is in dit onderzoek gebruik gemaakt van een nieuwe landelijke dataset die de voorraad tot 2024 in kaart brengt. Hiervoor zijn de geometrische contouren en bouwjaren uit de meest recente Basisregistratie Adressen en Gebouwen (BAG) gekoppeld aan de werkgelegenheidsdata van LISA (2024) en het formele zoneringsregister IBIS (bedrijventerreinen). Een gedetailleerde weergave van de data-preparatie is terug te vinden in Bijlage A en B.

De data-preparatie in R bouwt in de basis voort op de historische monitor van de logistieke voorraad (Nefs, 2023), en is schematisch weergegeven in Figuur 1.

Fase 1: Pre processing BAG buiten R



Figuur 1: Stroomschema van de datavoorbereiding en validatiestappen

Een cruciale stap binnen deze trechter is de 16-digit BAG-fix, uitgevoerd in de R-omgeving (Script 1). De BAG-database gebruikt identificatiecodes die vaak tekstuele toevoegingen bevatten (zoals 'NL.IMBAG.Pand.'). Wanneer datasets gekoppeld worden, veroorzaken deze letters foutmeldingen. Met behulp van een filterscript in R (een Regex-operatie) zijn deze letters verwijderd, zodat er per gebouw alleen de schone 16-cijferige getallencode overbleef. Hierdoor konden de gebouwen wiskundig foutloos gekoppeld worden aan externe datasets over het gebruiksdoel, zonder dat het programma vastliep of gebouwen dubbel telde. De omvang filter (grondoppervlakte $\geq 500\text{m}^2$) wordt vervolgens als eerste stap bij het inlezen in de R-omgeving toegepast.

3.2 Classificatie en Validatie

Om de ruwe data te zuiveren tot een betrouwbare logistieke voorraad, is een semi-geautomatiseerd classificatieprotocol toegepast.

3.2.1 SQL-Statusselectie en de LISA-dominantiereg

Om de BAG-data schoon in te lezen is via een SQL-query (Script 1, bijlage B) gefilterd op de actuele status "Pand in gebruik" of "Bouw gestart". Oude, administratieve aanpassingen zijn hierbij direct verwijderd.

De indeling van de logistieke sectorfuncties is in deze gebaseerd op de structuur van Nefs (2024, Table APP1.1) met een paar methodologische aanpassingen. Waar de geodata uit 2021 (Nefs, 2023) zich focuste op drie hoofdcategorieën, is deze dataset opgedeeld in vijf groepen: (Groot)handel/Import/Export (Groep 1), Transport/Opslag/Andere logistieke dienstverlening (Groep 2), XXL Retail (Groep 3), Tankopslag (Groep 4) en Bulkopslag (Groep 5) (zie Bijlage C).

Binnen deze indeling is e-commerce (SBI 47.91) expliciet toegewezen aan Groep 1 (Trade), terwijl de overige fysieke detailhandel (SBI 47) als XXL Retail (Groep 3) is geclassificeerd. Bulkopslag (Groep 5) is in deze monitor strikt gedefinieerd als de afval- en recyclingketen (SBI 38 en 39).

Een pand wordt vervolgens pas geclassificeerd als logistiek wanneer er meer dan 50% van de totale geregistreerde werkgelegenheid (LISA-banen) binnen een gebouw afkomstig is uit deze specifieke logistieke SBI-sectoren. Hiermee wordt gegarandeerd dat industriële productiefaciliteiten met een ondergeschikte opslagruimte van minder dan de helft van het personeelsbestand methodologisch worden uitgesloten.

3.2.2 Eigen Geautomatiseerde Algoritmes: Spookpanden- en Kassen-killer

Hoewel de administratieve geschiedenis is geschoond, bevat de landelijke kaart soms tekenfouten. Om deze te verwijderen, zijn binnen deze thesis twee ruimtelijke algoritmes geprogrammeerd (Script 2, Bijlage B):

- 1. De Spookpanden-killer (Geometrische ontdebelling):** Soms zijn er per ongeluk twee versies van hetzelfde gebouw over elkaar heen getekend. Het algoritme zoekt deze dubbelingen op en bewaart alleen het grootste landoppervlak. Controle in de data liet zien dat dit in alle 3.336 gevallen correct samenviel met het nieuwste bouwjaar, waardoor de overschatting van de nieuwbouw is voorkomen
- 2. De Kassen-killer (Agrarische filtering):** Grote kassencomplexen groter dan 10.000m² lijken qua oppervlakte soms op grote mega-distributiecentra en vervuilen de selectie. Om deze kassen automatisch te verwijderen, maakt het algoritme gebruik van locatie en verblijfsobjecten. Wanneer een gebouw buiten een bedrijventerrein ligt en geen industriële functie heeft, wordt het uit de dataset gefilterd.
- 3. De Tank-fix (Geometrische classificatie vloeibare bulk):** Fysieke opslagtanks missen in de LISA-database vaak een geregistreerde bedrijfsfunctie (brievenbusadres), wat in geautomatiseerde analyses tot onterechte uitval leidt. Om deze cruciale logistieke infrastructuur te behouden, is een geometrisch algoritme geïmplementeerd op basis van het Isoperimetrisch Quotiënt, in geografisch onderzoek vaak aangeduid als de compactness score (Maceachren, 1985). Binnen deze GIS-rekenregel scoort een perfecte cirkel 1,0 en een perfect vierkant 0,785. Door een drempelwaarde van $>0,82$ te hanteren, worden rechthoekige en vierkante gebouwen wiskundig uitgesloten en worden uitsluitend (semi-)ronde objecten geselecteerd. Wanneer deze objecten binnen de formele contouren van een IBIS-bedrijventerrein vallen, krijgen hierdoor volautomatisch de classificatie Tankopslag (Groep 4).

3.2.3 De Erfenis-Audit en Risico-trechter

Bij het vergelijken van de nieuwe data met de voorraad uit de Monitor 2021 (Nefs, 2023), bleek dat een groep panden die destijds als logistiek was gekwalificeerd in de actuele LISA-set van 2024 0 banen registreerde of onvindbaar was.

Dit verschil ontstaat doordat de economische data (LISA) als puntlocaties op basis van postadressen is geregistreerd, terwijl de BAG-data werkt met getekende gebouwcontouren. Om te voorkomen dat actieve distributiecentra hierdoor onterecht uit de dataset verdwijnen, zijn historische logistieke panden uit 2021 die géén actuele LISA-match

hebben, maar wél op een formeel IBIS-bedrijventerrein liggen, als veilige marge alsnog opgenomen in de kernvoorraad. Deze panden zijn in 2021 handmatig gevalideerd en daarvan wordt aangenomen dat deze onveranderd gebleven zijn. Resterende panden zonder LISA-match én historische panden die in 2021 wel een LISA match hadden maar in 2024 niet, stroomden door naar de potentials-werklijst voor een handmatige, visuele verificatie.

3.2.4 Handmatige Validatie via de GIS/Shiny Interface

Nadat panden met bewezen logistieke LISA-banen en de goedgekeurde erfenis-panden uit de lijst waren gefilterd, werden er nog wat extra filters toegepast om tot de werklijsten voor potentials te komen. Resterende panden ($\geq 2.500\text{m}^2$, gebouwd na 1990) stroomden door naar een werklijst van 2.522 panden voor handmatige validatie.

Om deze overgebleven panden betrouwbaar te valideren is gebruikgemaakt van een interactieve Shiny-applicatie in R (zie Bijlage D), beschikbaar gesteld door de Erasmus Universiteit Rotterdam. De applicatie laat een satellietbeeld en de contouren zien van het betreffende pand, waarna de checker aan de hand van een visuele verificatie checklist (zie bijlage C) kan beslissen of dit pand bij een van de 5 categorieën hoort. Dit leidde tot een totale goedkeuring van 778 logistieke objecten op basis van visuele macro-kenmerken en google maps.

3.2.5 De Definitieve Logistieke Voorraad (Stand eind 2024)

Na voltooiing van de algoritmische filters en de handmatige validatierondes is de Monitor Logistiek Vastgoed 2026 (met 2024 LISA data) in de GIS-omgeving vastgesteld op een totale omvang van 31.342 unieke, ruimtelijk gevalideerde panden. Dit betreft de geconsolideerde totale voorraad (stand tot eind 2024), en vormt het definitieve vertrekpunt voordat in Hoofdstuk 4 specifiek wordt ingezoomd op de ruimtelijke uitbreiding tussen 2020 en 2024.

De wiskundige opbouw van deze eindvoorraad is geconsolideerd vanuit twee hoofdwegen. Allereerst de geautomatiseerde kernvoorraad van 30.564 panden, deze bestaat uit LISA-matches én de historische erfenis-panden uit de monitor van 2021 (Nefs, 2023) die door de risico-trechter stroomden. De resterende panden zijn handmatig gecheckt via de Shiny-interface (zie bijlage D), wat 778 goedgekeurde potentials aan de database heeft toegevoegd. De exacte kwantitatieve verdeling van deze totale vastgoedvoorraad naar categorie en maatklasse is weergegeven in Tabel 1, en bijbehorende het ruimtebeslag (in hectaren) in Tabel 2.

Tabel 1 Overzicht van het aantal logistieke panden per categorie en grootteklasse in 2024 (Monitor 2026)

Logistieke Categorie	S (500 – 2.5k)	M (2.5k – 10k)	L (10k – 20k)	XL (20k – 40k)	XXL (>40k)	Totaal Panden
1. Groothandel / Trade	12.525	4.624	524	176	54	17.903
2. Transport & Logistiek	2.463	2.433	829	417	140	6.282
3. XXL Retail / E-commerce	3.364	2.715	232	53	8	6.372
4. Tankopslag / Vloeibaar	299	23	1	4	0	327
5. Bulkopslag	221	211	22	4	0	458
Totaal	18.872	10.006	1.608	654	202	31.342

Tabel 2 Overzicht van het ruimtebeslag (in hectaren) van logistieke panden per categorie en grootteklasse in 2024 (Monitor 2026)

Logistieke Categorie	S (500 – 2.5k)	M (2.5k – 10k)	L (10k – 20k)	XL (20k – 40k)	XXL (>40k)	Totaal m2
1. Groothandel / Trade	1.465,2	2.077,9	702,8	476,1	381,0	5.103,0
2. Transport & Logistiek	291,6	1.222,5	1.162,0	1.097,5	804,4	4,578,0
3. XXL Retail / E-commerce	398,8	1.147,8	309,4	143,6	46,4	2,046,0
4. Tankopslag / Vloeibaar	33,9	9,7	1,3	9,7	0,0	54,6
5. Bulkopslag	25,7	102,5	30,1	11,5	0,0	169,9
Totaal	2.215,3	4.560,4	2.205,6	1.738,3	1.231,8	11.951,4

Wanneer we naar deze tabellen kijken, is er een duidelijk verschil tussen het aantal panden en de daadwerkelijke ruimte die ze innemen. Kwantitatief wordt de voorraad gedomineerd door kleinschalige distributie: de maatklasse S (500 – 2.500 m²) vertegenwoordigt met ruim 18.000 objecten 60% van alle logistieke panden in Nederland. Dit zijn voornamelijk groothandels

Maar als we naar het werkelijke ruimtebeslag in vierkante hectaren (Tabel 2) kijken, wordt de invloed van de schaalvergroting zichtbaar. De selecte groep van 202 XXL-panden (>40.000m²) neemt in haar eentje een landoppervlak in van ruim 1.200 hectare. Binnen deze XXL-klasse is de sector Transport & Logistiek (Groep 2) veruit de grootste. Omdat de voetafdruk van deze gigantische panden fundamenteel verschilt van de reguliere groothandel, is het logisch om in het vervolg van dit onderzoek (het regressiemodel) duidelijk onderscheid te maken tussen deze verschillende maatklassen.

3.3 Verklaren Ruimtelijke Drijfveren

Om de locatiekeuze van logistiek te analyseren, wordt de methodologie van Tare et al. (2024) gebruikt en uitgebreid. De keuze om in deze analyse specifiek in te zoomen op de L/XL- en XXL-maatklassen (en panden onder 10.000m² uit te sluiten), is een bewuste afbakening. Het zijn namelijk primair deze grootschalige distributiecentra die verantwoordelijk zijn voor het enorme ruimtebeslag en daarmee de actuele maatschappelijke discussie rondom verdozing en netcongestie aanjagen.

Omdat de afhankelijke variabele in dit onderzoek bestaat uit discrete keuzes op gridniveau, is het Multinomiaal Logit Model (MNL) de aangewezen statistische benadering. De theoretische basis van het MNL-model in dit onderzoek is ingebed in het random utility framework. De log-kansen (log-odds) van elke mogelijke ruimtelijke uitkomst, berekend relatief ten opzichte van de referentiecategorie (“geen logistieke verandering”), volgen een lineair model (Tare et al., 2024):

$$\ln\left(\frac{P_{c,i}}{P_{c,j}}\right) = \alpha_i + \beta_i A_c + \delta_i L_c + \gamma_i S_c$$

Hierbij is:

- $P_{c,i}$: de waarschijnlijkheid op het optreden van logistieke ontwikkelingsklasse i (waarbij $i = 1$ voor L/XL-ontwikkeling en $i = 2$ voor XXL-ontwikkeling) in rastercel c .
- $P_{c,j}$: de waarschijnlijkheid op de referentie-uitkomst j (geen logistieke verandering in rastercel c).
- α_i : de constante factor (intercept) die de relatieve basisutiliteit per maatklasse i , weergeeft ten opzichte van ($j = 0$) bij een waarde van 0 voor de omgevingsfactoren. Dit wordt door het model geschat.
- A_c : een vector van onafhankelijke bereikbaarheidsfactoren (Accessibility) van cel c , waaronder de logaritmische afstand tot snelwegopritten van het hoofdwegenetwerk en intermodale terminals.
- L_c : een vector van sociaaleconomische locatiefactoren (Location) van cel c , operationeel gemaakt via hedonische grondprijzen en agglomeratie-indices op basis van LISA-banendichtheid.
- S_c : een vector van overheidsbeleid en fysieke restricties (Spatial policy & restrictions) van cel c . Naast binaire dummy-variabelen voor restrictieve provinciale en regionale zoning, zijn dit geactualiseerde model de regionale netcongestie-schaarste scores geïntegreerd als een nieuwe, dynamische restrictie-variabele.
- $\beta_i, \delta_i, \gamma_i$: de vector van te schatten regressiecoëfficiënten voor keuzemogelijkheid i , welke via de maximum likelihood methode statistisch worden gekwantificeerd.

3.3.1 Selectie variabelen

Om het MNL-model te kunnen draaien, zijn de ruimtelijk-economische concepten vertaald naar meetbare GIS-variabelen op gridniveau van 100x100 meter (exact 1 hectare per cel).

Voor de afhankelijke variabele is een ruimtelijke transitieanalyse uitgevoerd voor de periode 1 januari 2020 tot en met 31 december 2024. Uit de opgeschoonde pandendatabase zijn uitsluitend de gerealiseerde logistieke objecten met een bouwjaar binnen dit tijdvak geselecteerd. Vervolgens zijn deze vector-polygonen via ruimtelijke algebra getransformeerd naar een landelijk raster (grid) met een resolutie van 100 x 100 meter.

Wanneer een nieuwgebouwd pand de grenzen van meerdere rastercellen overschreed, kregen al deze geraakte cellen de waarde van de betreffende maatklasse toegekend. Bij overlap van verschillende logistieke klassen binnen een cel is er gekozen voor de dominante (grootste / XXL) klasse. Elke rastercel in Nederland kreeg zo een discrete waarde toegewezen: 1 voor L/XL nieuwbouw, 2 voor XXL-nieuwbouw, en 0 voor de referentiegroep (geen logistieke ontwikkelingen).

Een cruciaal methodologisch punt bij deze raster-benadering, zoals ook opgemerkt in de studie van Tare et al. (2024), is het omgaan met cellen die voor 2020 al deels of geheel bebouwd waren. Cellen met een hoge bestaande dichtheid hebben een theoretische logistieke uitbreidingskans van nul en zouden de referentiegroep (0) onterecht kunnen vertroebelen. Om deze ruis te elimineren, is de externe variabele voor beschikbare fysieke ruimte (*available_space_pct*) als uitsluitingsmasker ingezet. Dit masker voorkomt dat al volgebouwde gebieden het model verstoren.

Tabel 3 biedt een systematisch overzicht van de variabele in het model. Aanvullend toont Tabel 4 de beschrijvende statistieken van de continue onafhankelijke variabelen binnen de steekproef om inzicht te geven in de gehanteerde datareeksen.

Tabel 3 Overzicht en econometrische rol van variabelen in het MNL-model

Variabele in R	Type	Econometrische rol (H5)	Operationalisering & Beschrijving	Databron / Herkomst (Jaartal)
transitie_status	Categoriaal	Afhankelijke variabele ($\ln\left(\frac{P_{ci}}{P_{cj}}\right)$)	Discrete keuze per gridcel: • 0 = Geen transitie • 1 = L/XL-nieuwbouw (10.000 – 40.000 m2) • 2 = XXL-nieuwbouw (>40.000 m2)	Eigen synthese VU Amsterdam/ EUR (2026): BAG, LISA en Monitor Logistiek Vastgoed.
afstand_snelweg	Continu	Bereikbaarheidsvector (A_c)	Hemelsbrede afstand (in km) tot het nationale snelwegennetwerk (hoofdwegennet).	Nationaal Wegenbestand (NWB) / Rijkswaterstaat (2026).
afstand_op_afrit	Continu	Bereikbaarheidsvector (A_c)	Hemelsbrede afstand (in km) tot de dichtstbijzijnde formele snelwegafrit (fysieke netwerkknoop).	Nationaal Wegenbestand (NWB) / Rijkswaterstaat (2026).
afstand_terminal	Continu	Bereikbaarheidsvector (A_c)	Hemelsbrede afstand (in km) tot de dichtstbijzijnde multimodale binnenvaart-, spoor- of zeehaventerminal.	PDOK BRT Vervoer (OGC API-extractie, 2026).
Afstand_luchthaven	Continu	Bereikbaarheidsvector (A_c)	Hemelsbrede afstand (in km) tot de Top 5 BRT luchthavens (mainports).	Verkregen VU (ruimtebeslag)
rshdban_100	Continu	Locatiefactorenvector (L_c)	Agglomeratie-proxy: Totaal aantal geregistreerde banen binnen de rastercel (omgevingsdichtheid).	LISA data-extractie via VU Amsterdam (2025).
grondprijs	Continu	Locatiefactorenvector (L_c)	Geïndexeerde residuele/hedonische grondprijzen voor bedrijventerreinen op PC4-niveau (in €/m ²).	Verkregen via VU (?)
restrictie_randstad	Binair	Beleids- & Restrictievector (S_c)	1 = Locatie ligt binnen de provincies N-Holland, Z-Holland of Utrecht; 0 = Periferie. Meet institutionele druk/Randstad-vlucht.	Provinciale Omgevingsvisies (POVI, 2024).
beleid_stimulering	Binair	Beleids- & Restrictievector (S_c)	1 = Ligging binnen nationaal aangewezen logistieke topcorridor (o.a. speerpuntregio's op basis van PC2-hotspots); 0 = Nee.	Nationale Nota Ruimte / MIRT-programmering (2024).
netcongestie_dummy	Binair	Beleids- & Restrictievector (S_c)	1 = Aanwezigheid van acute structurele netcongestie (afname- of invoedingsbeperking niveau 3); 0 = Vrije netcapaciteit.	Capaciteitskaart Netbeheer Nederland via Vereniging Deltametropool (2025).
available_space_pct	Continu	Geëlimineerde variabele (Ex-ante)	Uitsluitingsmasker (Filters): Percentage onbebouwde fysieke ruimte binnen de gridcel. Uitgesloten wegens zero-variance (\$100\%\$ voor XXL).	CBS Vierkantstatistieken (100x100m grid, 2024).
restrictie_nnn	Binair	Geëlimineerde variabele (Ex-ante)	Uitsluitingsmasker (Filters): 1 = Ligging binnen het Natuurnetwerk Nederland (NNN). Uitgesloten wegens perfecte separatie (\$0\$ transities).	Provinciale Georegisters / Ministerie van LNV (GML-extractie, 2024).

Tabel 4 Beschrijvende statistieken van de continue variabelen (N = 6.011)

Variabele	Minimum	Gemiddelde (Mean)	Maximum
Afstand tot snelweg (km)	0,0	5,3	35,9
Afstand tot afrit (km)	0,0	3,5	28,1
Afstand tot terminal (km)	0,0	3,0	22,8
Afstand tot luchthaven (km)	0,1	46,3	118,1
Banendichtheid (per 1.000)	0,3	18,6	81,2
Grondprijs (€/m ²)	2,0	88,4	400,0

Noot. gebaseerd op de case-control sample output uit R

3.3.2 Econometrische verantwoording en uitsluiting van variabelen

Tijdens het bouwen van het statistische model is een aantal variabelen aangepast of uitgesloten om te zorgen dat het model wiskundig betrouwbaar kon rekenen.

Allereerst zijn twee variabelen geëlimineerd wegens extreme multicollineariteit. De bevolkingsdichtheid is weggelaten omdat deze vrijwel parallel liep aan de banendichtheid (Pearson's $r > 0,90$). Om dezelfde reden is de specifieke afstand tot zeehavens weggelaten, omdat deze al in het netwerk van multimodale terminals zat had deze automatisch veel overlap (Pearson's $r > 0,80$). Het gelijktijdig opnemen van deze variabelen zou het model in de war brengen.

Ten tweede was een wiskundige schaaltransformatie nodig voor de afstanden. Omdat afstanden in meters tot zeer grote getallen leiden (bijvoorbeeld bij luchthavens tot gemiddeld 46.000) en grondprijzen juist relatief kleine getallen zijn (gemiddeld 88,4) (zie Tabel 4), ontstaan er extreme schaalverschillen waardoor het computermodel (de maximum likelihood schatter) vastloopt. Door alle afstanden in het R script om te rekenen van meters naar kilometers, kon het rekenmodel stabiel lopen.

Tot slot is er een belangrijke keuze gemaakt rondom de variabele voor *available_space* (het percentage beschikbare, onbebouwde ruimte in een gridcel) en ligging in het Natuurnetwerk Nederland (NNN). Uit de data-analyse bleek dat nieuwe XXL-logistiek (>40.000m²) in de praktijk uitsluitend wordt gebouwd op kavels die nog 100% onbebouwd waren. Ook vonden er uiteraard geen veranderingen plaats in NNN-gebieden. Er zit op die locaties dus geen enkele variatie in de gerealiseerde data (nul variantie). Omdat een logit-model alleen kansen kan berekenen als er ergens een statistische afweging is, crashte het algoritme (*perfect quasi-complete separation*).

Het uitsluiten van beschikbare ruimte en natuur uit het model is daarom geen methodologisch falen, maar weerspiegelt de harde realiteit. 100% vrije ruimte en de afwezigheid van natuurbescherming zijn absolute, dwingende eisen voor een XXL-ontwikkelaar voordat er überhaupt gebouwd kan worden. Daarom zijn deze twee factoren

wiskundig niet als marginale locatiedrijvers in de regressie gestopt, maar in de eindfase van de data-analyse direct toegepast als binaire ruimtelijke uitsluitingsmaskers tijdens de voorspelfase.

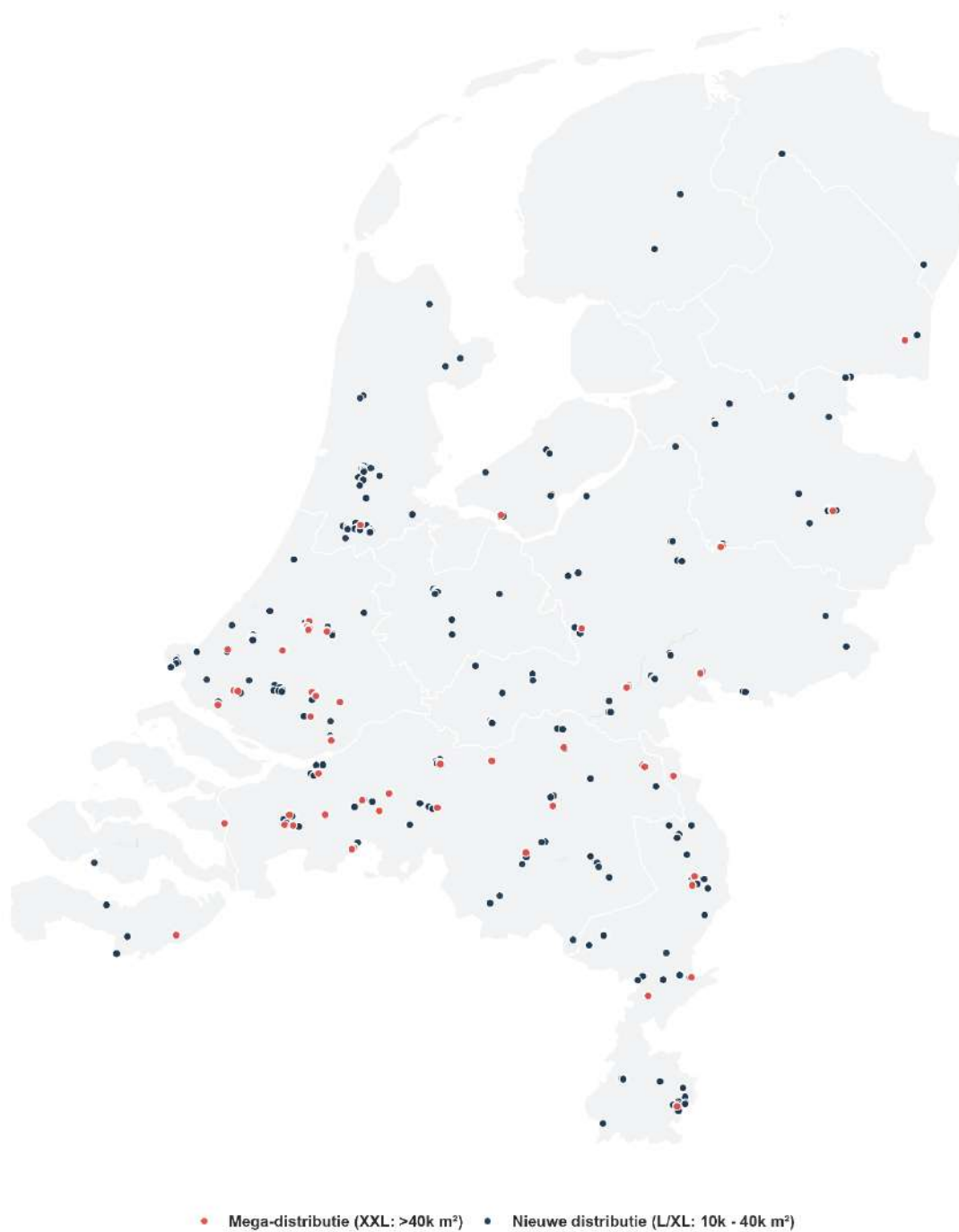
4. De Logistieke Schaalsprong: Ruimtebeslag 2020-2024

In dit hoofdstuk worden de eerste resultaten van de data-analyse gepresenteerd met betrekking tot absolute en relatieve omvang van de logistieke schaalsprong. Er wordt een kwantitatief en ruimtelijk overzicht gegeven van de ontwikkeling van het logistieke ruimtebeslag sinds 2020, waarbij expliciet de verschuivingen in het vastgoedlandschap en de groeiende dominantie van extreem grote (XXL) distributiecentra ten opzichte van middelgrote (L/XL) panden in kaart worden gebracht.

4.1 Het Logistieke Ruimtebeslag

Zoals in het voorgaande hoofdstuk (Tabel 1 en Tabel 2) is vastgesteld, kenmerkt de Nederlandse logistieke sector zich door een ruimtelijke tweedeling tussen het aantal fysieke objecten en het daadwerkelijke ruimtebeslag. Binnen de landelijke gemaakte dataset van 31.342 panden is de klasse S (<2.500m²) dominant in absolute aantallen (ruim 60%), maar claimt deze groep met 2.215 hectare relatief weinig landoppervlak. De daadwerkelijke logistieke schaalsprong wordt kwantitatief gedreven door de XXL-klasse (>40.000m²).

Om de recente expansiedrift ruimtelijk te maken, zijn de harde nieuwbouwtransities geanalyseerd op een gridresolutie van 100x100 meter (exact één hectare per gridcel) voor de periode tussen 1 januari 2020 en 31 december 2024 (Figuur 2). In deze vier jaar is er sprake van een totale fysieke expansie van 1.011 hectare landoppervlakte. Deze gerealiseerde transities vallen uiteen in twee actieve klassen: Klasse 1 (Nieuwe L/XL distributie, 10.000 – 40.000m²) verantwoordelijk voor 674 hectare en Klasse 2 (Nieuwe XXL distributie >40.000m²) verantwoordelijk voor 337 hectare aan nieuw ruimtebeslag.



Figuur 2 Locaties van de tussen 2020 en 2024 nieuw ontwikkelde L/XL-en XXL-logistieke panden. Weergegeven als dominante ontwikkeling per 100x100m gridcel. De kaart laat zien dat L/XL- en XXL-panden sterk geclusterd zijn rond grote snelwegen (zoals de A15 en A2) en logistieke knooppunten (Rotterdam, Amsterdam, Venlo). Deze locatiefactoren zijn als bereikbaarheidsvariabelen meegenomen in het regressiemodel.

4.2 Geografische patronen en Logistics Sprawl (Verschillen tussen L en XXL)

Wanneer de ruimtelijke allocatie van deze twee actieve transitieklassen geografisch wordt geprojecteerd (Figuur 2), wordt zichtbaar dat de recente logistieke expansie gepaard gaat met verschillende spreidingspatronen. Waar de toevoeging van nieuwe L/XL-distributie (Klasse 1) grotendeels fijnmazig geïntegreerd blijft binnen de bestaande stedelijke netwerken en traditionele bedrijventerreinen, vertoont de XXL-megadistributie (Klasse 2) een deconcentratietendens buiten de bestaande stedelijke kernen. Dit fenomeen, in de internationale literatuur gedefinieerd als logistics sprawl (Aljohani & Thompson, 2016; Cidell, 2010; Dabanc & Ross, 2012), kenmerkt zich door een selectieve uitwaaiing van megahallen naar perifere logistieke corridors.

Deze ruimtelijke expansie is echter geen homogene deconcentratie over het gehele landoppervlak. Zoals gekwantificeerd in Tabel 5 concentreert de dynamiek zich binnen een select aantal strategische corridors en specifieke macro-hotspots. De in Tabel getoonde Top-10 PC2-regio's is gezamenlijk verantwoordelijk voor maar liefst 44% van alle landelijke nieuwbouwtransities in deze periode, wat wijst op zware regionale clustering.

Tabel 5 overzicht van het aantal nieuwe logistieke transitie per PC2-regio en het procentuele aandeel binnen de maatklasse (2020-2024)

Rang	PC2-Regio	Grootschalige distributie (L/XL)	Mega-distributie (XXL)	Totaal aantal transitie
1	26 (Westland / Bleiswijk)	29 (4,3%)	47 (13,9%)	76 (7,5%)
2	59 (Venlo / Venray)	48 (7,1%)	13 (3,9%)	61 (6,0%)
3	47 (Roosendaal / Moerdijk)	34 (5,0%)	25 (7,4%)	59 (5,8%)
4	60 (Eindhoven / Tilburg)	35 (5,2%)	6 (1,8%)	41 (4,1%)
5	48 (Breda / Etten-Leur)	15 (2,2%)	24 (7,1%)	39 (3,9%)
6	54 (Oss / Veghel)	20 (3,0%)	19 (5,6%)	39 (3,9%)
7	76 (Zwolle / Kampen)	32 (4,7%)	5 (1,5%)	37 (3,7%)
8	66 (Arnhem / Nijmegen)	11 (1,6%)	22 (6,5%)	33 (3,3%)

9	29 (Rotterdam-Albrandswaard)	16 (2,4%)	15 (4,5%)	31 (3,1%)
10	31 (Utrecht / Nieuwegein)	22 (3,3%)	9 (2,7%)	31 (3,1%)
-	Top 10 Totaal	262 (38,9%)	185 (54,9%)	447 (44,2%)

In kwantitatief opzicht is deze sprawl-dynamiek het sterkst zichtbaar in West-Brabant en op de flanken van de primaire achterlandassen richting het Duitse achterland. De regio Breda/Etten-Leur (PC2-gebied 48) is illustratief voor deze pure XXL-schaalsprong: hier overstijgt het aantal XXL-transities (24 gridcellen) de reguliere L/XL-transities (15 gridcellen) met bijna een factor twee.

Vergelijkbare ontwikkelingen voltrekken langs de achterlandcorridors van de A15 en A12, specifiek binnen het knooppunt Arnhem-Nijmegen (PC2 66) met 33 transities (waarvan 22 XXL). Langs de maritieme achterlandcorridor rondom Roosendaal en Moerdijk (PC2 47) is er sprake van een sterke mix met in totaal 59 transities (34 L/XL tegenover 25 XXL).

Ter contrast laat de traditionele logistieke hotspot Venlo/Venray (PC2 59) zien dat zij haar groei voornamelijk handhaaft via fijnmaziger L/XL vastgoed, van de 59 transities zijn er hier slechts 13 XXL. Dit duidt op mogelijke agglomeratienadelen (agglomeration diseconomies), zoals arbeidsmarktkrapte, waardoor XXL-ontwikkelingen uitwijken naar opkomende corridors.

Een significante ruimtelijke uitzondering op het sprawl-fenomeen vormt de regio Westland/Bleiswijk/Den Hoorn (PC2 26). Met 76 gerealiseerde transities (29 L/XL en 47 XXL) fungeert dit gebied als de absolute koploper van de logistieke expansie tussen 2020 en 2024. De extreme ruimtelijke clustering van e-commerce en agro-logistiek ketens direct aan de rand van de Randstad trotseert in dit gebied de fysieke schaarste en hoge grondprijzen. Dit kwantitatieve bewijs toont aan dat voor deze specifieke stromen de dwingende ruimtelijke noodzaak tot nabijheid van de consumentenmarkt en mainports (Tare et al., 2024) zwaarder weegt dan de centrifugale krachten die logistics sprawl normaliter dicteren (Heitz et al., 2020).

5 Drijfveren van Ontwikkeling

Dit hoofdstuk bevat de kernresultaten van de ruimtelijke regressieanalyse. Waar hoofdstuk 4 het logistieke ruimtebeslag in kaart bracht, toetst dit hoofdstuk de onderliggende ruimtelijke-economische drijfveren. Via het Multinomiaal Logit Model (MNL) (de opzet daarvan staat in

Bijlage E) is geanalyseerd welke factoren de vestigingskans van middelgrote (L/XL) en megagrote (XXL) distributiecentra beïnvloeden, ten opzichte van gebieden zonder logistieke ontwikkeling. Eerst worden de modelparameters gepresenteerd (5.1), gevolgd door een interpretatie van bereikbaarheid en locatiefactoren (5.2) en ruimtelijk beleid en netcongestie (5.3). Het hoofdstuk sluit af met een visuele weergave van de ruimtelijke kansenallocation (5.4).

5.1 Interpretatie van het Multinomiaal Logit-Model

Tabel 6 toont de regressiecoëfficiënten (β) en standaardfouten (SE) van het MNL-model. Het model is stapsgewijs opgebouwd om de toegevoegde waarde van beleids- en restrictievariabelen te kwantificeren. Model 1 (Basismodel) bevat uitsluitend fundamentele geografische en sociaaleconomische locatiefactoren en behaalt een McFadden Pseudo R^2 van 0,123 in Model 2 (Integraal Beleidsmodel) zijn ruimtelijk beleid, mainports en netcongestie toegevoegd, wat resulteert in een aanzienlijke verbetering van de fit (McFadden Pseudo $R^2 = 0,182$). De interpretatie in de komende paragrafen is gebaseerd op dit definitieve, integrale model (Model 2), waarin de meeste geoperationaliseerde variabelen statistisch significant zijn.

Tabel 6 Schattingsresultaten van het MNL-model voor distributiecentra (weergegeven in β -coëfficiënten / log-odds

Variabele	Model 1: Basismodel		Model 2: Integraal Beleidsmodel	
	β (L/XL)	β (XXL)	β (L/XL)	β (XXL)
Constante (Intercept)	0,657*** (0,174)	0,054 (0,232)	0,970*** (0,192)	-0,647* (0,299)
Bereikbaarheid				
Afstand tot snelweg (km)	-0,044*** (0,013)	-0,295*** (0,029)	-0,018 (0,013)	-0,266*** (0,031)
Afstand tot afrit (km)	-0,055*** (0,011)	-0,060*** (0,015)	-0,040*** (0,011)	-0,048** (0,016)
Afstand tot intermodale terminal (km)	—	—	-0,287*** (0,030)	-0,150*** (0,040)
Afstand tot luchthaven (km)	—	—	-0,010*** (0,002)	-0,005 (0,003)
Locatiefactoren				
Grondprijs (€/m ²)	-0,002*** (0,001)	-0,001 (0,001)	-0,003*** (0,001)	-0,002* (0,001)
Banendichtheid (per 1.000)	-0,110*** (0,010)	-0,074*** (0,013)	-0,080*** (0,011)	-0,039** (0,015)
Beleid & Fysieke Restricties				

Beleid: Stimuleringscorridor (1=Ja)	—	—	0,754*** (0,110)	1,826*** (0,131)
Beleid: Restrictie Randstad (1=Ja)	—	—	0,024 (0,116)	0,596*** (0,174)
Netcongestie (1=Harde stop)	—	—	-0,549** (0,193)	-1,185*** (0,358)

Modelstatistieken

Log-Likelihood	-2953,5	-2754,8
McFadden Pseudo R ²	0,123	0,182
Aantal observaties (N)	6.011	6.011

*Noot. De gepresenteerde schattingsresultaten betreffen log-odds (β -coëfficiënten). De referentiecategorie bij de multinomiale logit-schatting is 'Geen transitie' ($j = 0$). Alle afstanden zijn getransformeerd naar kilometers. Significantieniveaus: * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$. Standaardfouten (SE) staan tussen haakjes.*

5.2 Impact van Bereikbaarheid en Locatiefactoren

De resultaten bevestigen de rol van transportkosten en multimodale netwerken bij logistieke locatiekeuze, maar tonen ook heterogeniteit tussen de maatklassen. De directe afstand tot een snelwegafrit heeft een significant negatief effect op de vestigingskans van zowel L/XL-panden ($\beta = -0,040$; $p < 0,001$) als XXL-panden ($\beta = -0,048$; $p < 0,01$). Een negatieve coëfficiënt voor afstand betekent hier hoe korter de afstand, hoe groter de kans op vestiging.

Maar wanneer we kijken naar hemelsbrede afstand tot het snelwegnetwerk ontstaat er een tweedeling. Voor XXL-panden is de nabijheid tot de snelweg een sterke en significante drijfveer ($\beta = -0,266$; $p < 0,001$), terwijl dit effect voor L/XL-panden een stuk kleiner is ($\beta = -0,154$; $p < 0,05$). Dit biedt ondersteuning voor Hypothese 1 (Toegankelijkheid): middelgrote logistiek oriënteert zich prima op de daadwerkelijke op- en afritten voor regionale bevoorrading, terwijl de omvangrijke (inter)nationale goederenstromen van XXL-ontwikkelingen correleren met een grotere afhankelijkheid van het hoofdwegennet.

Wat betreft intermodale bereikbaarheid vergroot een kortere afstand tot intermodale terminals de kans op logistieke ontwikkeling. Tegen verwachting in is deze afhankelijkheid sterker voor het L/XL-segment ($\beta = -0,287$; $p > 0,001$) dan voor het XXL-segment ($\beta = -0,150$; $p < 0,001$). Deze afwijking geldt ook voor de afstand tot belangrijke luchthavens. Nabijheid tot een luchthaven (zoals Schiphol) vergroot de vestigingskans van L/XL-panden significant ($\beta = -0,010$; $p > 0,001$), terwijl dit effect voor de XXL-logistiek afwezig is en statistisch insignificant is ($p = 0,145$). Dit nuanceert Hypothese 1 verder en suggereert dat L/XL-distributie sterker verweven is met time-critical e-commerce en high-value luchtvracht terwijl XXL-faciliteiten zich hier niet door laten sturen.

Agglomeratie en grondprijzen

De sociaaleconomische variabelen leveren kwantitatief bewijs voor het fenomeen logistics sprawl. In tegenstelling tot de theorie van agglomeratievoordelen, vertoont een hoge lokale banendichtheid een negatief effect op logistieke nieuwbouw. Dit deconcentratie-effect is sterker voor L/XL-panden ($\beta = -0,080$; $p > 0,001$) dan voor XXL panden ($\beta = -0,039$; $p < 0,01$).

Hetzelfde geldt voor grondprijzen. Hogere grondprijzen verlagen de vestigingskansen voor beide klassen, maar dit remmende effect is voornamelijk bepalend voor de L/XL-categorie ($\beta = -0,003$; $p < 0,001$). Bij XXL-ontwikkelingen is de negatieve impact van dure grond statistisch gezien zwakker en nauwelijks significant ($\beta = -0,002$; $p < 0,05$). Dit leidt tot een gedeeltelijke verwerping van Hypothese 2 (Ruimtebeslag en Landwaarde). De resultaten wijzen uit dat met name middelgrote distributiecentra (L/XL) gevoelig zijn voor hoge grondprijzen en uitwijken naar periferie, terwijl XXL-ontwikkelingen minder prijselastisch reageren, mits de locatie aan de overige infrastructurele voorwaarden voldoet.

5.3 De Effectiviteit van Ruimtelijk Beleid en Omgevingsfactoren

Aan de hand van de berekende parameters in Model 2 (Tabel 6) is getoetst of ruimtelijke beleidsmaatregelen en fysieke belemmeringen invloed hebben op de locatiekeuze van logistiek. Hierbij worden de ruwe β -coëfficiënten (log-odds) geïnterpreteerd, waarbij een positieve coëfficiënt (>0) wijst op een grotere waarschijnlijkheid op vestiging, en een negatieve coëfficiënt (<0) op een kleinere waarschijnlijkheid.

Ruimtelijke beleidsmaatregelen vertonen een sturend effect op de logistieke expansie. Provinciaal en nationaal stimulerend beleid langs de formele logistieke topcorridors correleert positief met nieuwe vestigingen. Dit beleid blijkt vooral effectief voor XXL-ontwikkelingen ($\beta = 1,826$; $p < 0,001$), en in mindere mate voor L/XL-panden ($\beta = 0,754$; $p < 0,001$). Dit ondersteunt Hypothese 3 (Beleid), de positieve log-odds tonen aan dat de overheden er daadwerkelijk in slagen om grootschalige logistiek te bundelen in de daarvoor aangewezen macro-hotspots

De dummy-variabele voor de Randstad (bedoeld om restrictief provinciaal beleid te meten) toont daarentegen een complexer beeld. Terwijl het geen significante invloed uitoefent op L/XL ontwikkelingen ($p = 0,836$), hangt het positief samen met XXL-ontwikkelingen ($\beta = 0,596$; $p < 0,001$). Deze variabele meet niet enkel beleid, maar representeert ook de enorme economische aantrekkingskracht van dit kerngebied. Dit wijst erop dat de nabijheid van mainports en de omvangrijke consumentenmarkt in de Randstad nog steeds buitengewoon

aantrekkelijk is voor specifieke deelsectoren. Voor deze spelers weegt de locatievoorkeur zwaarder dan de planologische restricties. Dit fenomeen remt logistics sprawl enigszins af. De noodzaak voor nabijheid voorkomt dat de hele sector naar de goedkopere periferie uitwaaiert.

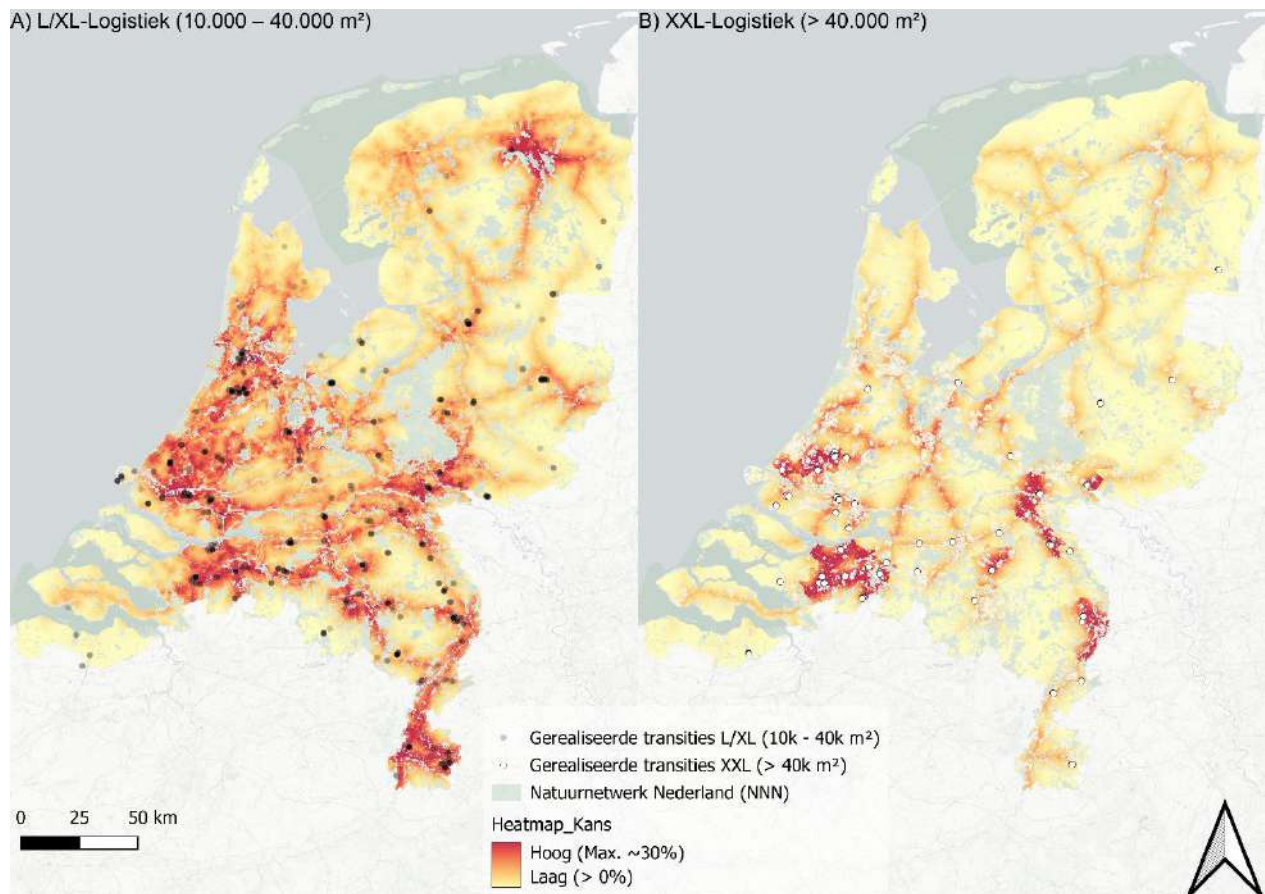
Netcongestie als ruimtelijke barrière

De analyse toont aan dat netcongestie een significante, belemmerende factor is voor nieuwe logistieke vestigingen. Gebieden met een geregistreeerde transport schaarste (niveau 3) op het elektriciteitsnet trekken aanzienlijk minder ontwikkelingen aan.

Dit manifesteert zich het sterkst in het XXL-segment ($\beta = -1,185$; $p < 0,001$) in ten opzichte van het L/XL segment ($\beta = -0,549$; $p < 0,01$). Om de daadwerkelijke impact beter te interpreteren is deze log-odds coëfficiënt in dit onderzoek wiskundig omgerekend naar een odds-ratio (e^β). Een odds-ratio kleiner dan 1 duidt op een afname van de vestigingskans. Uit deze berekening volgt een odds-ratio van 0,3 ($e^{-1,185}$). Dit betekent dat de relatieve vestigingskans van een XXL-distributiecentrum met 70% afneemt wanneer een gebied kampt met netcongestie, vergeleken met congestievrije gebieden. Voor L/XL-panden resulteert de berekening ($e^{-0,549}$) in een odds-ratio van 0,58, oftewel een afname van 42%. Deze bevindingen bieden sterk empirisch bewijs voor Hypothese 4 (Fysieke Restricties): voldoende stroomcapaciteit is een harde randvoorwaarde geworden voor locatiekeuze.

5.4 Ruimtelijke Kansenallocatie en Modelvoorspellingen

Om de wiskundige resultaten uit de vorige paragrafen ruimtelijk inzichtelijk te maken, zijn de berekende log-odds (de β 's uit Model 2) omgerekend naar ruimtelijke kansenverdelingen op gridniveau (zie Bijlage E). Hiervoor is de formule van het Integrale Beleidsmodel via ruimtelijke algebra geprojecteerd op de rasterkaart van Nederland. Het resultaat hiervan is gevisualiseerd in de voorspellingsmodellen (zie Figuur 3a en 3b).



Figuur 3 Ruimtelijk transitiepotentieel en gerealiseerde locaties voor L/XL-logistiek (A) versus XXL-logistiek (B) in Nederland (2020-2024). De heatmap toont de berekende kansen op basis van het Integrale Beleidsmodel (Model 2). De rode gebieden hebben de hoogste wiskundige waarschijnlijkheid op vestiging, rekening houdend met de vooraf uitgesloten maskers (natuurgebieden en beschikbare ruimte).

Deze geografische validatie illustreert de waargenomen asymmetrie in locatievoorkeuren. Het L/XL-segment (Figuur 3a) vertoont een relatief diffuse spreiding rondom stedelijke netwerken en inland terminals. Het XXL-segment (Figuur 3b) laat daarentegen een geconcentreerde marktdruk zien in specifieke macro-hotspots zoals Midden- en West-Brabant en Noord-Limburg. Tot slot visualiseren de kaarten de invloed van de ex-ante uitsluitingsmaskers (zoals verantwoord in Hoofdstuk 3). De voorspelde economische vestigingskansen wordt op diverse strategische locaties direct begrensd door het Natuurnetwerk Nederland (NNN) of een gebrek aan aaneengesloten, onbebouwde ruimte. Dit benadrukt dat moderne logistieke locatiekeuze niet enkel een economische afweging is, maar een optimalisatie op het snijvlak van bereikbaarheid, institutionele ruimte en fysieke (net)capaciteit.

6 Conclusie en Discussie

Dit afsluitende hoofdstuk beantwoordt de hoofd- en deelvragen op basis van de empirische bevindingen. Vervolgens worden deze resultaten en de implicaties voor ruimtelijk beleid in een bredere maatschappelijke discussie geplaatst vanuit het perspectief van Aarde, Economie en Duurzaamheid (AED). Tot slot wordt gereflecteerd op de limitaties van het onderzoek en worden suggesties voor toekomstig onderzoek aangedragen.

6.1 Conclusie

Dit onderzoek heeft de ruimtelijke drijfveren achter de Nederlandse logistieke schaa sprong tussen 2020 en 2024 geanalyseerd. De hoofdvraag richtte zich op de mate waarin de locatiekeuze van nieuwe XXL-distributiecentra (>40.000 m²) verschilt van middelgrote L/XL panden (10.000 – 40.000 m²). Uit de resultaten van het Multinomiale Logit Model (MNL) blijkt dat de logistieke sector geen homogene entiteit is, de geanalyseerde maatklassen vertonen hierin asymmetrische locatievoorkeuren.

Allereerst integreren middelgrote L/XL-distributiecentra zich veelal fijnmazig in bestaande netwerken. Zij blijken significant gevoeliger voor agglomeratienadelen en hoge grondprijzen. XXL-megadistributiecentra ontvluchten deze lokale stedelijke druk en waaieren selectief uit naar perifere logistieke corridors (logistics sprawl). Dit doen zij echter primair gedreven door de strikte noodzaak voor 100% onbebouwde grond (greenfields), aangezien zij minder prijselastisch reageren op grondwaarden dan L/XL-panden. Een uitzondering hierop vormt de distributie rondom het Westland, waar de marktnoodzaak om dicht bij de consument te zitten zwaarder weegt dan de drang om de stad te ontvluchten.

Daarnaast is de invloed van infrastructuur asymmetrisch. Waar L/XL-panden een sterke afhankelijkheid tonen van luchthavens en intermodale inland terminals, vereisen de omvangrijke goederenstromen van XXL-faciliteiten voornamelijk directe ontsluiting op het nationale wegennetwerk. Tot slot toont het onderzoek aan dat actuele fysieke restricties de traditionele marktfactoren overschaduwden. Netcongestie fungeert tegenwoordig als een harde ruimtelijke barrière. Een tekort aan stroomcapaciteit (Niveau 3) verlaagt de vestigingskans voor XXL-panden met 70%, wat aanzienlijk zwaarder is dan L/XL-panden (42%). Hiermee kan worden geconcludeerd dat de locatiekeuze van moderne logistiek is getransformeerd tot een ruimtelijke optimalisatieopgave, waarbij fysieke netcapaciteit en de beschikbaarheid van blanco bouwruimte sturender zijn dan de klassieke transportgeografie.

6.2 Theoretische Reflectie en Discussie

De bevindingen uit dit onderzoek vragen om een nieuwe blik op bestaande logistieke locatiemodellen. Klassieke locatietheorieën (waaronder die van Tare et al., 2024) verklaren de expansie van distributiecentra primair vanuit een rationele afweging tussen transportkosten, landwaarde en agglomeratievoordelen. Dit onderzoek toont aan dat dit paradigma verschuift.

Ten eerste verrijkt dit onderzoek de theorie rondom logistics sprawl (o.a. Dablanc & Ross, 2012). De literatuur veronderstelt doorgaans dat de schaalvergroting van XXL-logistiek de primaire oorzaak is van uitwaaiering naar de periferie wegens hoge grondprijzen. De modelresultaten nuanceren dit: het zijn juist de middelgrote distributiecentra (L/XL) die statistisch gezien de sterkste negatieve gevoeligheid vertonen voor hoge landwaarden en banendichtheid. XXL-ontwikkelaars lijken dankzij schaalvoordelen kapitaalkrachtiger en accepteren hogere grondprijzen, zolang de locatie perfect multimodaal is ontsloten en binnen een nationaal stimuleringsgebied valt.

Ten tweede introduceert de factor netcongestie een nieuwe dimensie in de ruimtelijke economie. Waar eerdere literatuur zich focuste op infrastructuur voor mobiliteit (wegen en spoor), wijst dit onderzoek uit dat de energie-infrastructuur een limiterende factor is geworden. Omdat distributiecentra transformeren tot industriële grootverbruikers, weegt de beschikbaarheid van stroom in de huidige praktijk zwaarder dan de traditionele voordelen van ruimtelijke clustering. Dit verklaart ook de geobserveerde negatieve agglomeratie-effecten: traditionele logistieke hotspots naderen hun fysieke en infrastructurele draagkracht.

6.3 Methodologische Beperkingen

Een fundamentele beperking van het gebruik van BAG-data is dat deze database panden registreert op basis van bouwtechnische eenheden (muren en funderingen) en niet op basis van commercieel gebruik of eigendom (Kadaster, z.d.). Dit resulteert in specifieke situaties waarin een aaneengesloten, gigantisch distributiecentrum ten onrechte door de aanwezigheid van interne brandscheidingsmuren of gefaseerde bouwstromen is opgeknipt in meerdere uitzonderlijke polygonen binnen de BAG-database. Hierdoor kunnen deze sub-eenheden ten onrechte in de L/XL-klasse vallen, in plaats van een samenhangend XXL-vlak te vormen. Hoewel deze methodologische frictie in breder geografisch onderzoek naar logistieke monitoring vaak kwalitatief wordt erkend (Nefs, 2023), blijft het een inherente beperking van puur geautomatiseerde ruimtelijke joins. Dit zorgt ervoor dat er een lichte onderschatting van het absolute aantal XXL-objecten is in de puur kwantitatieve analyses.

Hoewel het MNL-model een robuuste verklaring biedt voor de recente logistieke ruimtedynamiek, kent het onderzoek methodologische beperkingen. Ten eerste is de variabele voor beschikbare fysieke ruimte (*available_space_pct*) ex-ante uit het regressiemodel gesloten wegens zero-variance. XXL-distributiecentra claimen een dusdanig groot aaneengesloten oppervlak dat zij in de data uitsluitend landen op kavels die bij aanvang 100% onbebouwd zijn. Hierdoor kon deze factor wiskundig niet als marginale locatiedrijver worden gekwantificeerd, maar diende het alleen als een binaire uitsluitingsfactor (masker).

Ten tweede bracht het isoleren van specifieke infrastructuur-effecten wiskundige uitdagingen met zich mee. Zo ontbrak de mogelijkheid om specifieke afstand tot de haven van Rotterdam als unieke variabele te kwantificeren, omdat deze nagenoeg perfect ruimtelijk correleert met het algehele netwerk van multimodale terminals. Het macro-bereikbaarheidseffect van deze zeehaven wordt hierdoor noodgedwongen deels geabsorbeerd door de algemene terminal-variabele.

Ten derde weerspiegelen de variabelen voor inwoners- en banendichtheid slechts de kwantitatieve marktomvang. Aangezien logistiek sterk afhankelijk is van een specifieke groep arbeiders, wordt de kwalitatieve, lokale arbeidskrapte in de huidige proxy onvoldoende gemeten. Tot slot introduceert de hoge gridresolutie (100x100 meter) het statistische risico op ruimtelijke autocorrelatie. Het toegepaste logit-model veronderstelt onafhankelijke foutsommen, terwijl logistieke transities in de realiteit de waarschijnlijkheid van ontwikkelingen in direct naburige rastercellen beïnvloeden. Het corrigeren van deze afhankelijkheid via ruimtelijke (SAR) modellen is een vereiste voor toekomstig onderzoek op gridniveau.

6.4 Beleidsaanbevelingen en Toekomstig Onderzoek

Gezien vanuit het perspectief van Aarde, Economie en Duurzaamheid (AED) leggen de resultaten een complexe frictie bloot tussen beleidsambities en marktrealiteit. Met de Ontwerp-Nota Ruimte (2026) zet de Rijksoverheid in op strikte clustering van XXL-logistiek ter bescherming van het open landschap. Hoewel het model aantoont dat nationaal corridor-beleid effectief is, zorgt ruimtelijke schaarste ervoor dat de markt gelijktijdig de periferie in wordt gedreven. Voor middelgrote logistiek (L/XL) gebeurt dit primair vanwege een vlucht voor hoge stedelijke grondprijzen, terwijl XXL-logistiek naar de periferie wordt gedwongen wegens een tekort aan beschikbare netcapaciteit en grote onbebouwde kavels.

De belangrijkste aanbeveling voor beleidsmakers is dan ook om ruimtelijke ordening en energie-planologie niet langer gescheiden te behandelen. Het concentreren van logistiek op macro-hotspots is ineffectief als de lokale energie-infrastructuur deze clustering niet kan faciliteren. Regionale overheden dienen de ontwikkeling van bedrijventerreinen direct te koppelen aan de uitbreidingsplannen van netbeheerders, of het faciliteren van gesloten energiehubs ('smart grids') als harde vergunningseis te stellen.

Voor toekomstig onderzoek wordt sterk aanbevolen om kwalitatieve arbeidsmarktdata te integreren om de impact van personeelsschaarste op de locatiekeuze te isoleren.

Daarnaast kan het methodologische probleem van ruimtelijke autocorrelatie op gridniveau worden verholpen door de inzet van Spatial Autoregressive (SAR) modellen, waarmee ruimtelijke afhankelijkheid tussen aangrenzende locaties wiskundig kan worden gecorrigeerd.

6.5 Gebruik van AI

Tijdens de uitvoering van dit onderzoek is Generatieve AI, (zoals Gemini en NotebookLM) als ondersteunend hulpmiddel ingezet. De modellen zijn specifiek gebruikt voor hulp bij de data-preparatie, het oplossen van complexe R-code en het corrigeren van zinnen en APA-richtlijnen. Het gebruik van deze middelen heeft alleen gediend als technologisch en methodologisch hulpmiddel. Het denkwerk, de formulering van hypothesen en de uiteindelijke interpretatie van de resultaten zijn volledig zelfstandig uitgevoerd. Dit is in lijn met de academische richtlijnen van de VU. Een uitgebreide en gedetailleerde verantwoording van het exacte AI-gebruik is opgenomen in het bijgevoegde onderzoeksportfolio.

Literatuurlijst

- Aljohani, K., & Thompson, R. G. (2016). Impacts of logistics sprawl on the urban environment and logistics: Taxonomy and review of literature. *Journal of Transport Geography*, 57(C), 255-263.
- Bowen, J. T. (2008). Moving places: The geography of warehousing in the US. *Journal of Transport Geography, Growing Public Transport Patronage*, 16(6), 379-387.
<https://doi.org/10.1016/j.jtrangeo.2008.03.001>
- CE Delft. (2024). *Netcongestie: Welke mitigerende maatregelen kunnen logistieke bedrijven nemen?* (No. 230142). CE Delft. https://ce.nl/wp-content/uploads/2024/07/CE_Delft_230142_Vervolgonderzoek-mitigerende-maatregelen_def.pdf
- Cidell, J. (2010). Concentration and decentralization: The new geography of freight distribution in US metropolitan areas. *Journal of Transport Geography, Tourism and climate change*, 18(3), 363-371. <https://doi.org/10.1016/j.jtrangeo.2009.06.017>
- College van Rijksadviseurs. (2019). *(X)XL verdozing—Minder, compacter, geconcentreerder, multifunctioneler* | College van Rijksadviseurs [Nieuwsbericht]. Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. <https://doi.org/10/29/advies-xxl-verdozing>
- Commissie voor de milieueffectrapportage. (2026). *Advies over het bereik en het detailniveau van de milieueffectrapportage voor de Nota Ruimte* (No. 3794). <https://www.commissiemer.nl/advies/nota-ruimte/>

- Dablanc, L., & Ross, C. (2012). Atlanta: A mega logistics center in the Piedmont Atlantic Megaregion (PAM). *Journal of Transport Geography, Special Section on Theoretical Perspectives on Climate Change Mitigation in Transport*, 24, 432-442.
<https://doi.org/10.1016/j.jtrangeo.2012.05.001>
- Enexis Netbeheer. (2024). *Het investeringsplan van Enexis Netbeheer* | Enexis Netbeheer. Enexis. <https://www.enexis.nl/over-ons/wie-zijn-wij/ons-investeringsplan>
- Graham, D. J. (2007). Variable returns to agglomeration and the effect of road traffic congestion. *Journal of Urban Economics*, 62(1), 103-120.
<https://doi.org/10.1016/j.jue.2006.10.001>
- Groene Metropoolregio Arnhem-Nijmegen. (2024). *Regionaal Programma Werklocaties (RPW)*. <https://www.gmr.nl/regionaal-programma-werklocaties-rpw-aanbod-in-balans-brengen/>
- Heitz, A., Dablanc, L., Olsson, J., Sanchez-Diaz, I., & Woxenius, J. (2020). Spatial patterns of logistics facilities in Gothenburg, Sweden. *Journal of Transport Geography*, 88, 102191. <https://doi.org/10.1016/j.jtrangeo.2018.03.005>
- Holl, A., & Mariotti, I. (2018). *The Geography of Logistics Firm Location: The Role of Accessibility*. 337-361. <https://doi.org/10.13140/RG.2.2.10113.04961>
- Kadaster. (z.d.). *Alles over de BAG*. Kadaster Zakelijk. Geraadpleegd 28 mei 2026, van <https://www.kadaster.nl/zakelijk/registraties/basisregistraties/bag>
- Maceachren, A. M. (1985). Compactness of Geographic Shape: Comparison and Evaluation of Measures. *Geografiska Annaler: Series B, Human Geography*, 67(1), 53-67.
<https://doi.org/10.1080/04353684.1985.11879515>

- Melo, P. C., Graham, D. J., & Noland, R. B. (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics*, 39(3), 332-342.
- Nefs, M. (2023). *Dutch Distribution Centres 2021 Geodata (Versie 2)* (Versie 2) [Geopackage, with copies in Shapefile, CSV and Geojson. Method in Pdf.]. 4TU.ResearchData. <https://doi.org/10.4121/19361018.V2>
- Nefs, M. (2024). *Landscapes of Trade: Towards sustainable spatial planning for the logistics complex in the Netherlands* [Dissertation (TU Delft), A+BE | Architecture and the Built Environment]. <https://doi.org/10.71690/abe.2024.10>
- Nefs, M., & Daamen, T. A. (2022). Behind the Big Box: Understanding the planning-development dialectic of large distribution centres in Europe. *European Planning Studies*, 31, 1-22. <https://doi.org/10.1080/09654313.2022.2057792>
- Nguyen, Y., & Sano, K. (2010). Location Choice Model for Logistic Firms with Consideration of Spatial Effects. *Transportation Research Record: Journal of the Transportation Research Board*, 2168, 17-23. <https://doi.org/10.3141/2168-03>
- Planbureau voor de Leefomgeving. (2026). *Naar een Duurzame Ruimtelijke Inrichting van Nederland*. PBL. <https://www.pbl.nl/system/files/document/2026-01/pbl-2026-reflectie-ontwerp-nota-ruimte-5977b.pdf>
- Provincie Noord-Brabant. (2023). *Vorbereidingsbesluit grootschalige logistiek* (No. 2509). <https://zoek.officielebekendmakingen.nl/prb-2023-2509.pdf>
- Regio Kop van de Veluwe. (2024). *Regionaal Programma Werklocaties KvdV 2024-2028: Definitieve versie programmeringsafspraken*.

- https://www.nunspeet.nl/fileadmin/Nunspeet_documenten/Producten/Over_Nunspeet/Regionaal_Programma_Werklocaties_KvdV_2024-2028.pdf
- Rli. (2016). *Mainports voorbij* (No. 2016/02). Raad voor de leefomgeving en infrastructuur.
- <https://www.rli.nl/publicaties/2016/advies/mainports-voorbij>
- Rosenthal, S. S., & Strange, W. C. (2003). GEOGRAPHY, INDUSTRIAL ORGANIZATION, AND AGGLOMERATION. *Review of Economics & Statistics*, 85(2), 377-393. (9987309).
- <https://doi.org/10.1162/003465303765299882>
- Sakai, T., Beziat, A., & Heitz, A. (2020). Location factors for logistics facilities: Location choice modeling considering activity categories. *Journal of Transport Geography*, 85, 102710. <https://doi.org/10.1016/j.jtrangeo.2020.102710>
- Sakai, T., Kawamura, K., & Hyodo, T. (2015). Locational dynamics of logistics facilities: Evidence from Tokyo. *Journal of Transport Geography*, 46, 10-19.
- <https://doi.org/10.1016/j.jtrangeo.2015.05.003>
- Savills. (2025). *Herorientatie in logistiek vastgoed: Schokken en stabiliteit in de Nederlandse markt*. <https://pdf.euro.savills.co.uk/the-netherlands/whitepaper-savills-nederland-herorientatie-in-logistiek-vastgoed.pdf>
- Tare, A., Nefs, M., Koomen, E., & Verhoef, E. (2024). Spatial drivers of logistics development in the Netherlands. *Journal of Transport Geography*, 121, 104047.
- <https://doi.org/10.1016/j.jtrangeo.2024.104047>
- van den Heuvel, F., Langen, P., van Donselaar, K., & Fransoo, J. (2013). Spatial Concentration and Location Dynamics in Logistics: The Case of a Dutch Province.

International Journal of Computational Geometry and Applications - IJCGA, 28.

<https://doi.org/10.1016/j.jtrangeo.2012.10.001>

Bijlages

Bijlage A: De Batch-scripts (OSGeo4W Shell & PowerShell)

```
@echo off
echo Batch-Script Panden
SET output_gpkg=bag_panden_totaal_NL.gpkg
SET log_file=import_log.txt

:: Verwijder oude log als die bestaat
if exist %log_file% del %log_file%

echo [1/3] Database initialiseren...
:: Zoek het eerste XML bestand automatisch op
for /f %a in ('dir /b *.xml ^| findstr /n . ^|
findstr "^1:"') do set "first_file=%a"
set "first_file=%first_file:*=%"

ogr2ogr -f GPKG %output_gpkg% %first_file% pand
-nln pand -overwrite

echo [2/3] Bulk-import gestart. Dit gaat lang
duren...
echo Starttijd: %time% > %log_file%

for %i in (*.xml) do (
    echo Verwerken: %i
    ogr2ogr -update -append -skipfailures
    %output_gpkg% %i pand -nln pand
    if errorlevel 1 echo FOUT bij bestand: %i
    >> %log_file%
)

echo [3/3] Indexeren en optimaliseren...
ogrinfo %output_gpkg% -sql "CREATE INDEX
idx_pand_id ON pand (identificatie)"
ogrinfo %output_gpkg% -sql "SELECT
CreateSpatialIndex('pand', 'geom')"

echo Eindtijd: %time% >> %log_file%
echo KLAAR. Check %log_file% voor eventuele
fouten.
pause
```

```
@echo off
echo Batch-Script VBO
setlocal enabledelayedexpansion

set OUTPUT_FILE=vbo_master_extract.csv

echo
=====
echo BAG VBO EXTRACTOR - BRUTE FORCE REGEX EDITION
echo
=====
echo vbo_id;pand_id;gebruiksdoel > %OUTPUT_FILE%

:: We gebruiken PowerShell om de tekst als rauwe
data te lezen.
:: Dit omzeilt alle XML-fouten en Namespace-
ellende.
powershell -Command ^
    "$files = Get-ChildItem *.xml;" ^
    "foreach ($f in $files) {" ^
    "    Write-Host 'Ploegen door: ' $f.Name;" ^
    "    $txt = Get-Content $f.FullName -Raw;" ^
    "    $matches = [regex]::Matches($txt,
    '(?s)<Objecten:Verblijfsobject>.*</Objecten:Verblijfsobject>');" ^
    "    foreach ($m in $matches) {" ^
    "        $vbo = [regex]::Match($m.Value,
    'identificatie[^\>]*>(\d+)' ).Groups[1].Value;" ^
    "        $pand = [regex]::Match($m.Value,
    'PandRef[^\>]*>(\d+)' ).Groups[1].Value;" ^
    "        $doel = [regex]::Match($m.Value,
    'gebruiksdoel>(.*)<' ).Groups[1].Value;" ^
    "        if ($vbo -and $pand) {" ^
    "            $vbo + ';' + $pand + ';' + $doel >>
    %OUTPUT_FILE%;" ^
    "        }" ^
    "    }" ^
    "}"

echo
=====
echo KLAAR! Als de CSV nu nog leeg is, bestaat de
data niet.
echo
=====
pause
```

Bijlage B: Geautomatiseerde Workflow in R (scripts 1 t/m 3)

```
#
=====
# SCRIPT 1: DATA-PREPARATIE BAG
# DOEL: Schone basiskaart maken van alle NL
panden >= 500m2.
# LOGICA: VBO-data (gebruiksdoelen) wel koppelen
voor vrijkijking, maar nog NIET
# op filteren (LISA/Script 2 is later
leidend).
#
=====

library(sf)
library(data.table)
library(dplyr)
library(here)
library(stringr)

cat("🚀 START: Script 1 - BAG Preparatie & VBO
Koppeling\n")
cat("=====
=====\\n")

# --- 1. BRONNEN & SQL SELECTION ---
vbo_path <- here("BAG", "VBO",
"vbo_master_extract.csv")
gpkg_path <- here("BAG", "PND",
"bag_panden_totaal_NL.gpkg")

# We filteren direct bij het inlezen op actieve
panden om geheugen te besparen.
query_pnd <- "SELECT * FROM pand
WHERE status IN ('Pand in
gebruik', 'Pand in gebruik (niet ingemeten)',
'Bouw gestart')
AND eindGeldigheid IS NULL
AND eindRegistratie IS NULL"

# --- 2. VBO DATA INLEZEN & AGGREGEREN ---
cat("📂 [1/4] VBO data laden en
aggregeren...\\n")

# VBO's (Verblijfsobjecten = de functies in een
pand) inlezen.
# colClasses als character behoudt de
belangrijke '0' vooraan de 16-cijferige ID's.
vbo_dt <- fread(vbo_path, colClasses =
"character", sep = ";", encoding = "UTF-8")
setnames(vbo_dt, old = names(vbo_dt), new =
c("vbo_id", "pand_id", "gebruiksdoel"))

# Rol de VBO's op per gebouw. We leggen vast of
er 'industrie' in zit
# en we plakken alle gebruiksdoelen aan elkaar
in één string.
vbo_agg <- vbo_dt[, .(
aantal_vbo = .N,
has_industrie = any(grepl("industrie",
gebruiksdoel, ignore.case = TRUE)),
gebruiksdoelen_list =
paste(unique(str_trim(gebruiksdoel)), collapse =
"|")
), by = .(pand_id)]

# Zorgt voor een zuivere 16-cijferige ID-sleutel

vbo_agg[, pand_id_clean :=
str_extract(as.character(pand_id), "\\d{16}$")]

# --- 3. METRIEKEN BEREKENEN & RUIMTELIJKE JOIN
---
cat("📊 [2/4] BAG Panden verwerken...\\n")
panden_raw <- st_read(dsn = gpkg_path, query =
query_pnd, quiet = TRUE)

panden_final <- panden_raw %>%
st_zm(drop = TRUE) %>% # Verwijder
hoogte (3D) dimensies, dit voorkomt errors in
GIS
st_make_valid() %>% # Repareert
tekenfouten (bijv. polygonen die zichzelf
snijden)
mutate(
identificatie_clean =
str_extract(as.character(identificatie),
"\\d{16}$"),
bouwjaar =
as.numeric(oorspronkelijkBouwjaar),
oppervlakte_calc =
as.numeric(st_area(geom)),
omtrek =
as.numeric(st_length(st_boundary(geom))),

# Formule voor Compactness: Scoort van 0
(extreem langwerpig) tot 1 (perfecte
cirkel/tank).
compactness = (4 * pi * oppervlakte_calc) /
(omtrek^2)
) %>%
# Filter direct op oppervlakte (Logistiek is
altijd >= 500m2).
# Compactness > 0.01 vlakt corrupte BAG-
tekenfouten uit (rechte lijntjes).
filter(oppervlakte_calc >= 500, compactness >
0.01) %>%

# Voeg de VBO data toe
left_join(vbo_agg %>% select(pand_id_clean,
has_industrie, aantal_vbo, gebruiksdoelen_list),
by = c("identificatie_clean" =
"pand_id_clean")) %>%
mutate(
has_industrie = ifelse(is.na(has_industrie),
FALSE, has_industrie),
aantal_vbo = ifelse(is.na(aantal_vbo), 0,
aantal_vbo),

# Classificeer in T-shirt maten volgens
methodologie Nefs (2021)
size_class = case when(
oppervlakte_calc < 2500 ~ "S (500-
2.5k)",
oppervlakte_calc < 10000 ~ "M (2.5k-
10k)",
oppervlakte_calc < 20000 ~ "L (10k-20k)",
oppervlakte_calc < 40000 ~ "XL (20k-
40k)",
oppervlakte_calc >= 40000 ~ "XXL (>40k)"
)
)

# Voor de audit leggen we vast hoeveel unieke
rijen we HIER hebben, vóór het ontdebellen.
aantal_voor_ontdubbeling <- nrow(panden_final)

# --- 4. ADMINISTRATIEVE ONTDUBBELING ---
```

```

cat("🚧 [3/4] Dubbele pand-ID's eruit
slicen...\n")

# Soms zitten er 2 versies van hetzelfde pand-ID
in de BAG (bijv. een mutatie).
# We pakken de meest actuele (grootste
footprint) eruit en verwijderen de dubbele.
panden_final <- panden_final %>%
  group_by(identificatie_clean) %>%
  slice_max(oppervlakte_calc, n = 1, with_ties =
FALSE) %>%
  ungroup()

# --- 5. AUDIT & VALIDATIE ---
cat("\n🇳🇱 START: Audit\n")

# Controle of onze slice_max regel waterdicht
is.
validatie_check <- panden_raw %>%
  mutate(identificatie_clean =
str_extract(as.character(identificatie),
"\d{16}$")) %>%
  group_by(identificatie_clean) %>%
  filter(n() > 1) %>% # Check alleen de panden
met dubbelingen
  mutate(
    oppervlakte_calc =
as.numeric(st_area(geom)),
    bouwjaar =
as.numeric(oorspronkelijkBouwjaar)
  ) %>%
  summarise(
    kleiner_en_nieuwer = any(oppervlakte_calc <
max(oppervlakte_calc) & bouwjaar ==
max(bouwjaar, na.rm=TRUE)),
    .groups = "drop"
  )

if("TRUE" %in%
names(table(validatie_check$kleiner_en_nieuwer))
) {
  cat("⚠️ LET OP: Soms is een nieuwer pand
administratief kleiner geregistreerd.\n")
} else {
  cat("✅ Validatie geslaagd: Grootste
oppervlakte is representatief.\n")
}

# Tellen van het aantal verwijderde dubbelingen
aantal_na_ontdubbeling <- nrow(panden_final)
cat("🔥 AANTAL DUBBELE REGELS VERWIJDERD:",
(aantal_voor_ontdubbeling -
aantal_na_ontdubbeling), "\n")

# --- 6. EXPORT ---
cat("📁 [4/4] Exporteren naar RDS...\n")
saveRDS(panden_final, here("R",
"panden_basis_2026_vFinal.rds"))
cat("\n✅ SCRIPT 1 COMPLETELY DONE.\n")

```

```

#
=====
# SCRIPT 2: LISA INTEGRATIE, ERFENIS RESCUE &
CLASSIFICATIE
# DOEL: Verrijken van de ruimtelijke panden met
economische data (LISA 2024 banen).
# Het script splitst de dataset in een
bewezen 'Kernvoorraad' en een
# werkljst met 'Potentials' die handmatig
gecontroleerd moet worden.
#
=====

library(sf)
library(dplyr)
library(data.table)
library(here)
library(stringr)
library(tidyr)

cat("🔵 START: Script 2 - LISA Integratie &
Herclassificatie\n")
cat("=====
=====\\n")

# --- 1. DEFINITIES: FUNCTIONELE CATEGORIEËN
(SBI-CODES) ---
clean_sbi <- function(x) str_replace_all(x,
"[\\s ]", "")

# We groeperen de SBI-codes van de Kamer van
Koophandel in 5 logistieke hoofdklassen.
# Deze klassen weerspiegelen de aard van de
distributie (Trade, Transport, Retail, Tank,
Bulk).
## HIER IS MET AI WEL IETS MISGEGAAN DENK IK,
KLOPT NIET MET DE TABEL, HEB AANGEPAST
sbi_cat1_trade <- clean_sbi(c("45.11.1",
"45.19.1", "45.31.1", "45.31.2", "45.40.1",
"46.21.1", "46.21.2", "46.21.3", "46.21.4",
"46.21.5", "46.21.6",
"46.21.7",
"46.21.8", "46.21.9", "46.22", "46.23.1",
"46.23.2", "46.24.1", "46.24.2", "46.31.1",
"46.31.2", "46.32",
"46.33.1",
"46.33.2", "46.34", "46.35", "46.36", "46.37",
"46.38.1", "46.38.2", "46.38.3", "46.38.4",
"46.38.9", "46.39",
"46.41.1",
"46.41.2", "46.42.1", "46.42.2", "46.42.3",
"46.42.4", "46.42.5", "46.42.9", "46.43.1",
"46.43.2", "46.43.3",
"46.43.4",
"46.43.5", "46.43.6", "46.44.1", "46.44.2",
"46.46.1", "46.46.2", "46.47.1", "46.47.2",
"46.47.3", "46.48",
"46.49.1",
"46.49.2", "46.49.3", "46.49.4", "46.49.5",
"46.49.6", "46.49.7", "46.49.8", "46.49.9",
"46.51", "46.52", "46.61",
"46.62",
"46.63", "46.64", "46.65", "46.66", "46.68.1",
"46.68.2", "46.69.1", "46.69.2", "46.69.3",
"46.69.4", "46.69.5",
"46.69.6",
"46.69.7", "46.69.9", "46.71.1", "46.71.2",
"46.71.3", "46.72.1", "46.72.2", "46.72.3",
"46.73.1", "46.73.2",
"46.73.3",
"46.73.4", "46.73.5", "46.73.6", "46.73.7",
"46.73.8", "46.73.9", "46.74.1", "46.74.2",
"46.75.1", "46.75.2",
"46.76.1",
"46.76.2", "46.76.9", "46.77.1", "46.77.2",
"46.77.9", "46.90.1", "46.90.2", "47.91.1",
"47.91.2", "47.91.3",
"47.91.4",
"47.91.5", "47.91.6", "47.91.8", "47.91.9",
"73.20"))
sbi_cat2_transport <- clean_sbi(c("49.20",
"49.41", "50.20.1", "50.20.2", "50.40.1",
"50.40.2", "50.40.3", "51.21", "52.10.2",
"52.10.9", "52.21", "52.22",
"52.23",
"52.24.1", "52.24.2", "52.29.1", "52.29.2",
"53.10", "53.20.1", "53.20.2"))
# sbi_cat3_retail <- clean_sbi(c("47.1",
"47.11", "47.2", "47.21", "47.22", "47.22.1",
"47.22.2", "47.23", "47.24", "47.24.1",
"47.24.2", "47.25", "47.26", "47.29", "47.29.1",
"47.29.2", "47.29.3", "47.29.9", "47.3",
"47.30", "47.4", "47.41", "47.42", "47.43",
"47.43.1", "47.43.2", "47.5", "47.51",
"47.51.1", "47.51.2", "47.51.3", "47.52",
"47.52.1", "47.52.2", "47.52.3", "47.52.4",
"47.52.5", "47.52.6", "47.52.7", "47.53",
"47.54", "47.54.1", "47.54.2", "47.54.3",
"47.54.4", "47.59", "47.59.1", "47.59.2",
"47.59.3", "47.59.4", "47.59.5", "47.59.6",
"47.59.7", "47.6", "47.61", "47.62", "47.63",
"47.64", "47.64.1", "47.64.2", "47.64.3",
"47.64.4", "47.65", "47.7", "47.71", "47.71.1",
"47.71.2", "47.71.3", "47.71.4", "47.71.5",
"47.71.6", "47.71.7", "47.71.8", "47.72",
"47.72.1", "47.72.2", "47.73", "47.74",
"47.74.1", "47.74.2", "47.75", "47.76",
"47.76.1", "47.76.2", "47.76.3", "47.77",
"47.78", "47.78.1", "47.78.2", "47.78.3",
"47.78.9", "47.79", "47.79.1", "47.79.2",
"47.79.3", "47.8", "47.81", "47.81.1",
"47.81.9", "47.82", "47.89", "47.89.1",
"47.89.2", "47.89.9", "47.9", "47.99",
"47.99.1", "47.99.2", "47.99.9"))
sbi_cat4_tank <- clean_sbi(c("52.10.1"))
#HIER ZAT EEN FOUT TANKOPSLAG IS 52101 IPV 52102
sbi_cat5_bulk <-
clean_sbi(c("38.1", "38.2", "38.3", "39.0"))
#afvalketen

# --- 2. DATA LADEN ---
cat("📂 [1/4] Gevalideerde BAG-data, LISA 2024,
IBIS en Monitor 2021 inlezen...\n")
panden_basis <- readRDS(here("R",
"panden_basis_2026_vFinal.rds"))

# LISA is geolocatie-gebaseerd (punten). We
projecteren deze naar EPSG:28992
(Rijksdriehoeksmeting).
lisa_sf <- readRDS(here("LISA",
"LISA_2024_clean.rds")) %>%
st_as_sf(coords = c("XCOORD", "YCOORD"), crs =
28992)

ibis <- st_read(here("IBIS",
"ibis_openbaar.shp"), quiet = TRUE) %>%
st_transform(28992)
nefs_2021 <- st_read(here("landscapes of trade",
"distributioncenters2021NL.gpkg"), layer =
"buildings2021NL", quiet = TRUE)

# --- 3. LISA RUIMTELIJKE JOIN & AGGREGATIE ---
cat("📂 [2/4] Ruimtelijke join (Point-in-
Polygon): Banen optellen per gebouw...\n")

# Koppel LISA bedrijfsvestigingen aan de
gebouwen waar ze binnen vallen

```

```

panden_lisa_join <- st_join(panden_basis,
  lisa_sf, join = st_intersects)
panden_dt <-
as.data.table(st_drop_geometry(panden_lisa_join)
)
panden_dt[, sbi_clean := clean_sbi(SBI08)]

# Stap A: Wijs elk LISA-record toe aan één
categorie op basis van de Nefs-regels
match_cat1 <- paste(sbi_cat1_trade, collapse =
"|")
match_cat2 <- paste(sbi_cat2_transport, collapse
= "|")
match_cat4 <- paste(sbi_cat4_tank, collapse =
"|")

panden_dt[, assigned_cat := case_when(
  str_starts(sbi_clean, match_cat1) ~ 1,
  str_starts(sbi_clean, match_cat2) ~ 2,
  str_starts(sbi_clean, match_cat4) ~ 4,
  grepl("^47", sbi_clean) ~ 3, # Heel
SBI 47 is Retail, online retail valt al onder
Cat 1
  grepl("^38|^39", sbi_clean) ~ 5, # Bulk
(Afvalsector)
  TRUE ~ NA_real_ #
Geen logistiek
)]

# Stap B: Tel het aantal banen op per pand en
bereken de som van alle logistieke banen
panden_stats <- panden_dt[, .(
  totaal_banen = sum(as.numeric(BANEN), na.rm =
TRUE),
  banen_cat1 =
sum(as.numeric(BANEN)[assigned_cat == 1], na.rm
= TRUE),
  banen_cat2 =
sum(as.numeric(BANEN)[assigned_cat == 2], na.rm
= TRUE),
  banen_cat3 =
sum(as.numeric(BANEN)[assigned_cat == 3], na.rm
= TRUE),
  banen_cat4 =
sum(as.numeric(BANEN)[assigned_cat == 4], na.rm
= TRUE),
  banen_cat5 =
sum(as.numeric(BANEN)[assigned_cat == 5], na.rm
= TRUE),
  hoofd_huurder = first(NAAM[order(-
as.numeric(BANEN)])),
  sbi_omschr = first(OMSCHR[order(-
as.numeric(BANEN)]))
), by = identificatie_clean]

panden_stats[, totaal_logistieke_banen :=
banen_cat1 + banen_cat2 + banen_cat3 +
banen_cat4 + banen_cat5]

# Stap C: Pand krijgt een definitieve LISA-
categorie als >50% van de banen logistiek is (de
grootste sector wint)
panden_stats[, lisa_cat_num := case_when(
  totaal_banen > 0 & (totaal_logistieke_banen /
totaal_banen) > 0.5 ~ case_when(
    banen_cat1 >= pmax(banen_cat2, banen_cat3,
banen_cat4, banen_cat5) & banen_cat1 > 0 ~ 1,
    banen_cat2 >= pmax(banen_cat1, banen_cat3,
banen_cat4, banen_cat5) & banen_cat2 > 0 ~ 2,
    banen_cat3 >= pmax(banen_cat1, banen_cat2,
banen_cat4, banen_cat5) & banen_cat3 > 0 ~ 3,
    banen_cat4 >= pmax(banen_cat1, banen_cat2,
banen_cat3, banen_cat5) & banen_cat4 > 0 ~ 4,
    banen_cat5 >= pmax(banen_cat1, banen_cat2,
banen_cat3, banen_cat4) & banen_cat5 > 0 ~ 5,
    TRUE ~ NA_real_
  )
  )

```

```

),
  TRUE ~ NA_real_
)]

# --- 4. HISTORISCHE AFSTEMMING (Erfenis 2021) ---
cat("🚧 [3/4] Integreren van Monitor 2021
(Erfenis-check)...\n")

# Bereid historische data voor (strip de zware
geometrie voor snelle verwerking)
nefs_historisch <- nefs_2021 %>%
  st_set_geometry(NULL) %>%
  mutate(
    BAG_id_clean =
str_extract(as.character(BAG_id), "\\d{16}$"),
    historische_banen = as.numeric(Jobs_tot),
    had_lisa_match_2021 = historische_banen > 0,
    cat_2021_num = case_when(
      Function == "Trade, Import, Export" ~ 1,
      Function == "Transport and logistic
services" ~ 2,
      Function == "XXL retail center" ~ 3
    )
  )

# Koppel alle data aan elkaar. Actuele LISA-data
is leidend, anders 2021 als fallback.
panden_verrijkt <- panden_basis %>%
  left_join(panden_stats, by =
"identificatie_clean") %>%
  left_join(nefs_historisch %>%
select(BAG_id_clean, cat_2021_num,
historische_banen, had_lisa_match_2021),
    by = c("identificatie_clean" =
"BAG_id_clean")) %>%
  mutate(
    final_cat_num = case_when(
      !is.na(lisa_cat_num) ~ lisa_cat_num, #
2024 data is leidend
      is.na(lisa_cat_num) & !is.na(cat_2021_num)
~ cat_2021_num, # Fallback
      TRUE ~ NA_real_
    )
  )

# --- 5. RUIMTELIJKE FILTERS (Kassen-killer &
Tank-fix) ---
cat("🏢 [4/4] Ruimtelijke restricties berekenen
(IBIS & Vormefficiëntie)...\n")

panden_verrijkt <- panden_verrijkt %>%
  mutate(
    op_ibis = lengths(st_intersects(., ibis)) >
0,

    # ROTTERDAMSE HAVEN FIX: Zeer ronde objecten
(compactness > 0.82) krijgen geforceerd
Tankopslag (Cat 4)
    final_cat_num = case_when(
      compactness > 0.82 &
(clean_sbi(sbi_omschr) %in% sbi_cat4_tank |
is.na(final_cat_num)) ~ 4,
      TRUE ~ final_cat_num
    )
  ) %>%

  # KASSEN-KILLER: Verwijdt grote (10k+)
agrarische kassen (extreem langwerpig/grillig:
compactness < 0.15)
  filter(!(has_industrie == FALSE &
oppervlakte_calc >= 10000 & compactness < 0.15 &
op_ibis == FALSE))

# --- 6. SPLITSING IN WERKSTROMEN (Kernvoorraad
vs Potentials) ---

```

```

cat("🔍 [5/5] Splitsen van werkstromen en
ruimtelijke differentiatie...\n")

# A. KERNVOORRAAD (Bewezen logistieke panden)
kernvoorraad <- panden_verrijkt %>%
  filter(!is.na(final_cat_num)) %>%
  filter(
    # Scenario A: Actuele LISA-match. (Retail
    altijd > 500m2, andere sectoren op IBIS of
    >2500m2)
    (!is.na(lisa_cat_num) &
     (
       (lisa_cat_num == 3 & ((op_ibis == TRUE
       & oppervlakte_calc >= 500) | oppervlakte_calc >=
       2500)) |
       (lisa_cat_num != 3 & (op_ibis == TRUE
       | oppervlakte_calc >= 2500))
     )
    ) |
    # Scenario B: Reddingsboei - Historisch
    logistiek pand waarvan het LISA-puntje net
    buiten de muren valt, maar ligt wel op IBIS
    (is.na(lisa_cat_num) &
     !is.na(cat_2021_num) & op_ibis == TRUE) |
    # Scenario C: Historisch bewezen
    leegstaand vastgoed (vanaf 500m2)
    (is.na(lisa_cat_num) &
     !is.na(cat_2021_num) & (is.na(historische_banen)
     | historische_banen == 0) & oppervlakte_calc >=
     500) |
    # Scenario D: Tank-Fix - Vangt de ronde
    opslagtanks op IBIS op die geen LISA-adres
    hebben
    (is.na(lisa_cat_num) & is.na(cat_2021_num)
     & final_cat_num == 4 & op_ibis == TRUE)
  )

# B. POTENTIALS (Restlijst voor handmatige
check)
potentials_bruto <- panden_verrijkt %>%
  filter(!identificatie_clean %in%
kernvoorraad$identificatie_clean) %>%
  filter(is.na(totaal_banen) | totaal_banen ==
0) %>% # Geen (andere) actieve economische
functie
  filter(is.na(sbi_omschr)) %>%
  filter(
    (had_lisa_match_2021 == TRUE &
    oppervlakte_calc >= 2500) |
    (op_ibis == TRUE & oppervlakte_calc >=
    2500 & bouwjaar >= 1990) |
    (op_ibis == FALSE & oppervlakte_calc >=
    10000 & bouwjaar >= 1990) # Vangnet voor
    gigantische stand-alone centra
  ) %>%
  distinct(identificatie_clean, .keep_all =
TRUE)

# Splits werklijst op basis van
bedrijventerrein-ligging
potentials_ibis <- potentials_bruto %>%
  filter(op_ibis == TRUE)
potentials_buiten <- potentials_bruto %>%
  filter(op_ibis == FALSE)

# --- 7. EXPORT & AUDIT MATRIX SCRIPT 2 ---
cat("\n=====
=====
\n")

cat("📄 ROUTERING AUDIT MONITOR (SCRIPT 2)\n")
cat("=====
=====
\n")

audit_summary <- data.frame(
  Onderdeel = c("Gevalideerde Kernvoorraad",
    "Openstaande Potentials (BINNEN
IBIS)",

```

```

"Openstaande Potentials (BUITEN
IBIS / Standalone")),
  Aantal = c(nrow(kernvoorraad),
nrow(potentials_ibis), nrow(potentials_buiten))
)
print(audit_summary)

cat("\n📄 Exporteren van gesplitste werklijsten
naar schijf...\n")
saveRDS(kernvoorraad, here("R",
"kernvoorraad_valide_final.rds"))

saveRDS(potentials_ibis, here("R",
"potentials_ibis_werklijst.rds"))
saveRDS(potentials_buiten, here("R",
"potentials_buiten_werklijst.rds"))
st_write(potentials_ibis, here("R",
"potentials_werklijst_BINNEN_IBIS.gpkg"),
delete_dsn = TRUE, quiet = TRUE)
st_write(potentials_buiten, here("R",
"potentials_werklijst_BUITEN_IBIS.gpkg"),
delete_dsn = TRUE, quiet = TRUE)

cat("\n✅ SCRIPT 2 VOLTOOID. Bestanden staan
klaar.\n")

```

```
#
=====
# SCRIPT 3: INTEGRALE RUIMTELIJKE HARMONISATIE &
# FINALE DATA-CLEANUP (vFinal)
# DOEL: Bouwen van het 100% dekkende
# eindproduct. Integreert 5 team GPKGs,
# automatiseert IBIS-potentials, en
# isoleert de standalone buitengebied-wachtrij.
#
=====

library(sf)
library(dplyr)
library(data.table)
library(here)
library(stringr)
library(tidyr)

cat("🚀 START: Script 3 - Finale Integrale
Harmonisatie & Oplevering\n")
cat("=====
=====\\n")

# --- 1. SCHONE BASELINES INLEZEN ---
cat("📂 [1/4] Inlezen van de drie gesplitste
stromen uit Script 2...\n")
pad_kernvoorraad <- here("R",
"kernvoorraad_valide_final.rds")
pad_potentials_ibis <- here("R",
"potentials_ibis_werklijst.rds")
pad_potentials_buiten <- here("R",
"potentials_buiten_werklijst.rds")

if(!file.exists(pad_kernvoorraad) ||
!file.exists(pad_potentials_ibis) ||
!file.exists(pad_potentials_buiten)) {
  stop("🛑 STOP: De output van Script 2
ontbreekt. Run eerst Script 2!")
}

schone_kernvoorraad <-
readRDS(pad_kernvoorraad) %>%
mutate(herkomst_basis = "Kernvoorraad")
schone_potentials_in <-
readRDS(pad_potentials_ibis) %>%
mutate(herkomst_basis = "Potential (Binnen
IBIS)")
schone_potentials_uit <-
readRDS(pad_potentials_buiten) %>%
mutate(herkomst_basis = "Potential (Buiten
IBIS)")

# Samenvoegen tot één basismatrix voor
ruimtelijke matching
totale_voorraad_basis <-
bind_rows(schone_kernvoorraad,
schone_potentials_in, schone_potentials_uit)

# --- 2. VERWERKEN VAN DE 5 TEAM-CONTROLES
# (QGIS) ---
cat("🔧 [2/4] Verwerken van de 5 handmatige
QGIS-checks van het team...\n")
pad_gpkg_1 <- here("final potential check",
"Maryam - FINAL.gpkg")
pad_gpkg_2 <- here("final potential check",
"logistics gebouwen taibi manualcheck.gpkg")
pad_gpkg_3 <- here("final potential check",
"Merten_FINAL_part1.gpkg")
pad_gpkg_4 <- here("final potential check",
"Merten_FINAL_part2.gpkg")
pad_gpkg_5 <- here("R", "potentials",
"potentials_gecheckt_deel4.gpkg")
```

```
read_team_sf <- function(pad) {
  # Veiligheidscheck: bestaat het bestand wel?
  if(!file.exists(pad)) return(NULL)

  df <- st_read(pad, quiet = TRUE)

  # Check of de essentiële beoordelingskolom
  aanwezig is
  if(!"is_logistiek" %in% names(df))
  return(NULL)

  # Flexibele check voor de ID-kolomnaam (soms
  valt '_clean' weg na QGIS export)
  id_col <- intersect(names(df),
c("identificatie_clean", "identificatie"))[1]

  df %>%
    mutate(
      identificatie_clean = if(!is.na(id_col))
as.character(.data[[id_col]]) else
NA_character_,
      # Forceer numeriek (1 t/m 5), onderdruk R-
      warnings als er per ongeluk tekst in staat
      is_logistiek_num =
suppressWarnings(as.numeric(as.character(is_logi
stiek)))
    ) %>%
    # Gooi rijen weg die door het team leeg zijn
    gelaten
    filter(!is.na(is_logistiek_num)) %>%
    select(identificatie_clean,
is_logistiek_num, geom)
  }

  # Laad alle 5 de bestanden in en plak ze onder
  elkaar, projecteer naar exact hetzelfde CRS als
  de basis
  suppressWarnings({
    team_master_sf <- bind_rows(
      read_team_sf(pad_gpkg_1),
      read_team_sf(pad_gpkg_2),
      read_team_sf(pad_gpkg_3),
      read_team_sf(pad_gpkg_4),
      read_team_sf(pad_gpkg_5)
    ) %>%
    st_transform(st_crs(totale_voorraad_basis))
  })

  # --- STAP 1: Harde tekstuele ID-koppeling ---
  cat("🔗 Stap 1: Harde tekstuele ID-koppeling
(QGIS -> Basis)...\n")
  team_id_tabel <- team_master_sf %>%
    st_drop_geometry() %>%
    filter(!is.na(identificatie_clean)) %>%
    group_by(identificatie_clean) %>%
    slice_max(is_logistiek_num, n = 1, with_ties =
FALSE) %>%
    ungroup()

  voorraad_stap1 <- totale_voorraad_basis %>%
    left_join(team_id_tabel, by =
"identificatie_clean")

  # --- STAP 2: Ruimtelijk vangnet ---
  cat("📍 Stap 2: Ruimtelijk vangnet toepassen
voor verschoven QGIS-panden...\n")
  ruimtelijke_checks <- team_master_sf %>%
    filter(!identificatie_clean %in%
voorraad_stap1$identificatie_clean[!is.na(voorra
ad_stap1$is_logistiek_num)]) %>%
    filter(is_logistiek_num %in% c(1, 2, 3, 4, 5))
    %>%
    st_centroid()

  panden_voor_spatial_join <- voorraad_stap1 %>%
  filter(is.na(is_logistiek_num))
```



```

    categorie_label = case_when(
      categorie_num == 1 ~ "1. Groothandel / Trade",
      categorie_num == 2 ~ "2. Transport & Logistiek",
      categorie_num == 3 ~ "3. XXL Retail / E-commerce",
      categorie_num == 4 ~ "4. Tankopslag / Vloeibaar",
      categorie_num == 5 ~ "5. Bulkopslag",
      TRUE ~ "Onbekend"
    ),
    bouwjaar = ifelse(bouwjaar == 9999, 2025, bouwjaar)
  ) %>%
  select(identificatie_clean, bouwjaar, oppervlakte_calc, size_class, compactness, totaal_banen, hoofd_huurder, sbi_omschr, categorie_num, categorie_label, op_ibis, herkomst_eindproduct = status_controle, herkomst_basis, geom)

# --- 5. FINALE INTEGRALE AUDIT MATRIX ---
cat("\n=====
=====\\n")
cat("\n🇳🇱 FINALE AUDIT MATRIX MONITOR LOGISTIEK 2026\\n")
cat("=====
=====\\n")

audit_matrix <- eindproduct_overdracht %>%
  st_drop_geometry() %>%
  group_by(categorie_label, size_class) %>%
  summarise(n = n(), .groups = "drop") %>%
  pivot_wider(names_from = size_class, values_from = n, values_fill = 0)

vereiste_kolommen <- c("S (500-2.5k)", "M (2.5k-10k)", "L (10k-20k)", "XL (20k-40k)", "XXL (>40k)")
for(col in vereiste_kolommen) {
  if(!col %in% names(audit_matrix))
    audit_matrix[[col]] <- 0
}

audit_matrix <- audit_matrix %>%
  select(categorie_label,
    all_of(vereiste_kolommen)) %>%
  mutate(Totaal_Panden =
    rowSums(across(all_of(vereiste_kolommen))))

print(audit_matrix)
cat("=====
=====\\n")

cat("\n=====
=====\\n")
cat("\n🇳🇱 RUIMTEBESLAG MATRIX (Oppervlakte in m2)\\n")
cat("=====
=====\\n")

ruimtebeslag_matrix <- eindproduct_overdracht %>%

```

```

  st_drop_geometry() %>%
  group_by(categorie_label, size_class) %>%
  summarise(totale_opp_m2 =
    sum(oppervlakte_calc, na.rm = TRUE), .groups = "drop") %>%
  pivot_wider(names_from = size_class, values_from = totale_opp_m2, values_fill = 0)

for(col in vereiste_kolommen) {
  if(!col %in% names(ruimtebeslag_matrix))
    ruimtebeslag_matrix[[col]] <- 0
}

ruimtebeslag_matrix <- ruimtebeslag_matrix %>%
  select(categorie_label,
    all_of(vereiste_kolommen)) %>%
  mutate(Totaal_m2 =
    rowSums(across(all_of(vereiste_kolommen))))

print(ruimtebeslag_matrix)
cat("=====
=====\\n")

cat("\n=====
=====\\n")
cat("\n🇳🇱 RUIMTELIJKE HERKOMST MATRIX (IBIS vs Standalone)\\n")
cat("=====
=====\\n")

herkomst_matrix <- eindproduct_overdracht %>%
  st_drop_geometry() %>%
  group_by(categorie_label, op_ibis) %>%
  summarise(n = n(), .groups = "drop") %>%
  mutate(Locatie = ifelse(op_ibis == TRUE, "Binnen IBIS Bedrijventerrein", "Buiten IBIS (Standalone)")) %>%
  select(-op_ibis) %>%
  pivot_wider(names_from = Locatie, values_from = n, values_fill = 0)

print(herkomst_matrix)
cat("=====
=====\\n")

# --- 6. EXPORTEREN FINALE DATASET ---
cat("\n🇳🇱 Definitieve monitor wegschrijven naar schijf...\\n")
final_gpkg_path <- here("eindset", "MONITOR_LOGISTIEK_2026_DEFINITIEF.gpkg")
final_rds_path <- here("eindset", "MONITOR_LOGISTIEK_2026_DEFINITIEF.rds")

if (file.exists(final_gpkg_path))
  file.remove(final_gpkg_path)

st_write(eindproduct_overdracht,
  final_gpkg_path, driver = "GPKG", delete_dsn = TRUE, dataset_options = "OVERWRITE=YES", quiet = TRUE)
saveRDS(eindproduct_overdracht, final_rds_path)

cat("\n✅ SCRIPT 3 VOLTTOOID. Het eindproduct is succesvol gegenereerd en weggeschreven!\\n")

```

Bijlage C: Visuele Verificatie Logistiek Vastgoed

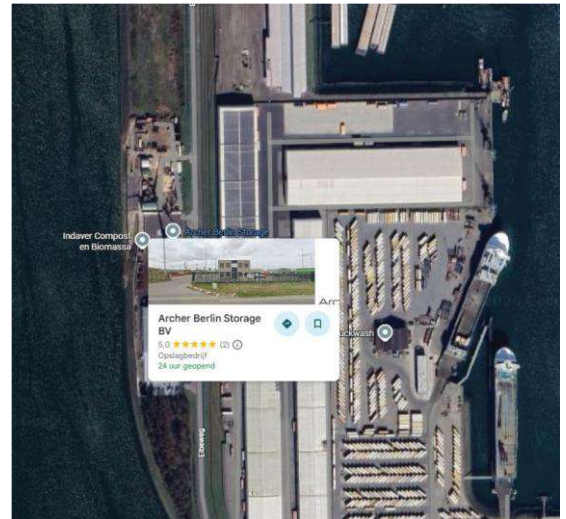
Categorieën

1. (Groot)handel, import, export
2. Transport, opslag en andere logistieke dienstverlening
3. XXL Retail
4. Tankopslag
5. Bulkopslag

Stap 1. Identificatie van bedrijf via Google Maps

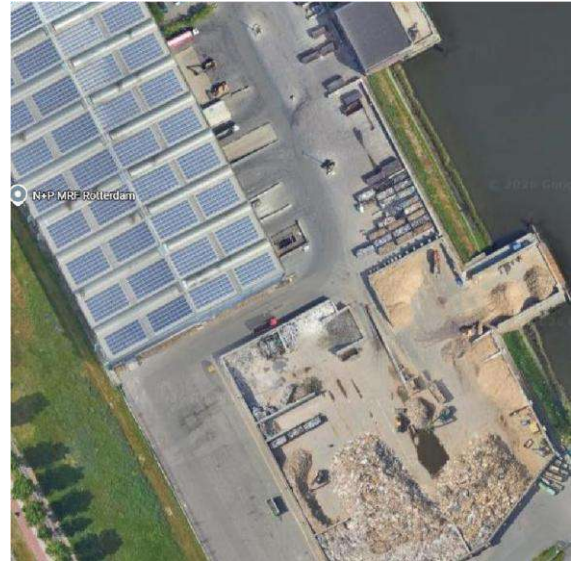
(Naam en link) of streetview (naam of logo op gevel)

- Naam van bedrijf in groothandel, e-commerce, import of export -> categorie 1
- Naam van bedrijf in transport, opslag, distributie of logistieke dienstverlening -> categorie 2
- Naam van winkel met zelfbedieningsmagazijn (Ikea, Gamma etc.) -> categorie 3
- Ronde tank met evt. logo van oliemaatschappij of tankopslag (Vopak) -> categorie 4
- Naam van bedrijf niet actief in bovenstaande activiteiten -> categorie 0 (te verwijderen) • Geen bedrijfsnaam -> ga naar Stap 2



Stap 2. Identificatie van activiteit via fysieke context

- Pand met opstelruimte voor trucks en weinig loading docks -> categorie 1
- Pand met opstelruimte voor trucks en veel loading docks -> categorie 2
- Pand met opslag van bulkmaterialen op terrein (Bv. voor recycling, geen bouw) -> categorie 5
- Pand zonder loading docks of bulkopslag -> categorie 0 (te verwijderen)
- Twijfel -> categorie 9 (2nd opinion)



Bijlage D: Logistics Building Review App - Setup Guide

This document explains how to set up and run the Shiny review interface that lets you classify building polygons one at a time on top of an interactive Google Map. Follow the four parts below in order. Anywhere you see add `xxxx` here in the code, replace it with your own value.



Part 1 - Set up the Google Maps API

The app embeds Google Maps in your browser, so you need a Google Cloud project with the Maps JavaScript API enabled and an API key. This takes about 10 minutes the first time.

1.1 Create a Google Cloud project

- Go to <https://console.cloud.google.com> and sign in with a Google account.
- In the top bar, click the project dropdown and choose New Project.
- Give it a name (e.g. logistics-review) and click Create.
- Make sure the new project is selected in the top bar before continuing.

1.2 Enable billing

Google requires a billing account even for low-volume use to make sure that it is a real account, and each account gets a recurring monthly free credit for Maps. You will not be charged unless you exceed it (which never happens according to my friend).

1.3 Enable the Maps JavaScript API

- Left sidebar: APIs & Services - Library.
- Search for **Maps JavaScript API** and click **Enable**.

1.4 Create an API key

- APIs & Services - Credentials - + Create Credentials - API key.
- A dialog shows your key (starts with AlzaSy...). Copy it now.
- Click **Edit API key** on the new key.

- Under **Application restrictions**, choose **None** for local use. (If you later host the app publicly, switch to HTTP referrers.)
- Under **API restrictions**, choose **Restrict key** and tick only **Maps JavaScript API**. This way, if the key ever leaks, it cannot be used for other Google services.
- Click Save.

1.5 Store the API key on your machine

Never paste API keys into shared scripts. R can read keys from a small config file at startup instead.

In R, run once:

```
file.edit("~/Renvirom")
```

Add one line, replacing the placeholder with your key, and save the file:

```
GOOGLE_MAPS_API_KEY=add_your_api_key_here
```

Then fully restart R (env vars are only read at startup). To verify, run:

```
nzchar(Sys.getenv("GOOGLE_MAPS_API_KEY")) # should print TRUE
```

Part 2 - Install R packages

Only needs doing once per machine. The sf package depends on system libraries (GDAL, PROJ); on Windows and macOS the binary install handles this automatically, on Linux you may need libgdal-dev and libproj-dev first.

```
install.packages(c("shiny", "sf", "DT", "bslib", "htmltools", "dplyr"))
```

Part 3 - Save the R script and fill in your paths

Copy the entire R code from Part 5 into a new file called review_app.R and save it anywhere on your machine. Before running, open the file and edit the three placeholders near the top:

- **INPUT_GPKG** - the full path to your input GeoPackage, e.g.
C:/Users/your_name/Downloads/your_file.gpkg
- **INPUT_LAYER** - the name of the layer inside the GeoPackage (often the same as the file name).
- **OUTPUT_GPKG** - the full path where review progress should be saved. Pick a new filename so the input is not overwritten.

On Windows, use forward slashes (/) in the paths even though Windows shows backslashes - R prefers them.

Part 4 - Run the app

In R or RStudio, run:

```
setwd("path_to_folder_with_review_app.R")
shiny::runApp("review_app.R", launch.browser = TRUE)
```

The app opens in your default browser. First launch will:

- Read your input GeoPackage.
- Create the output GeoPackage as a copy (review progress will be saved here).
- Build a queue of all rows where is_logistiek is NA and show the first one on the map.

Every classification writes immediately to the output GeoPackage. If the app crashes or you close the browser, your progress is safe. Re-launching the app loads from the output file if it exists, so you continue where you left off. To start over, delete the output file before launching. Important: do not keep the output GeoPackage open in QGIS while reviewing - on Windows, QGIS holds a file lock that will block the next save. If you want to peek mid-review, copy the file aside in File Explorer first.

Troubleshooting

- **Map area is blank:** press F12 in the browser to open DevTools and check the Console tab. The most common cause is the Maps JavaScript API not being enabled on the project, or an API-restriction blocking the key.

- **R error: GOOGLE_MAPS_API_KEY env var not set:** the ~/.Renviron file was not saved, or R was not restarted after editing it.
- **Error reading the gpkg:** double-check the INPUT_LAYER name. In R you can list layers with `sf::st_layers("path/to/file.gpkg")`.

Part 5 - The R script

Copy everything below into a file called `review_app.R`. Replace the three `add_XXXX_here` placeholders near the top with your own paths and layer name.

```
# Manual review GUI for logistics building
classification

# install.packages(c("shiny", "sf", "DT",
"bslib", "htmltools", "dplyr"))

suppressPackageStartupMessages({
  library(shiny)
  library(sf)
  library(DT)
  library(bslib)
  library(htmltools)
  library(dplyr)
})

# Input file (the already-classified
GeoPackage)
INPUT_GPKG <- "add_path_to_input_gpkg_here"
INPUT_LAYER <- "add_layer_name_here"

# Output file (created/overwritten - review
progress is saved here)
OUTPUT_GPKG <- "add_path_to_output_gpkg_here"

# Google Maps JavaScript API key.
# Set the env var GOOGLE_MAPS_API_KEY in
~/.Renviron, then restart R.
GMAPS_JS_KEY <-
Sys.getenv("GOOGLE_MAPS_API_KEY")
if (!nzchar(GMAPS_JS_KEY)) {
  stop("GOOGLE_MAPS_API_KEY env var not set. ",
       "Add it to ~/.Renviron and restart R.")
}

# Categories -> labels shown in the dropdown and
the keyboard hint bar.
CATEGORIES <- list(
  "0" = "0 - Not logistics (remove)",
  "1" = "1 - Wholesale / import / export",
  "2" = "2 - Transport, storage, logistics
services",
  "3" = "3 - XXL retail",
  "4" = "4 - Tank storage",
  "5" = "5 - Bulk storage",
  "9" = "9 - Doubt (2nd opinion)"
)

# load and prepare data

# If the manual-check file already exists from a
previous session, resume
# from it. Otherwise start fresh from the
classified file.
load_path <- if (file.exists(OUTPUT_GPKG))
OUTPUT_GPKG else INPUT_GPKG
message("Loading: ", load_path)

buildings <- st_read(load_path, layer =
INPUT_LAYER, quiet = TRUE)
```

```
# Reproject to WGS84 once. We keep the original
CRS for saving.
orig_crs <- st_crs(buildings)
buildings_wgs <- st_transform(buildings, 4326)

# Centroid lat/lon for map-centring.
centroids <-
st_centroid(st_geometry(buildings_wgs))
coords <- st_coordinates(centroids)
buildings$lon <- coords[, "X"]
buildings$lat <- coords[, "Y"]

# Polygon outline as a GeoJSON-friendly list of
lat/lon pairs.
ring_to_latlng <- function(ring_matrix) {
  lapply(seq_len(nrow(ring_matrix)), function(j)
  {
    list(lat = unname(ring_matrix[j, 2]),
         lng = unname(ring_matrix[j, 1]))
  })
}

poly_coords <-
lapply(seq_len(nrow(buildings_wgs)), function(i)
{
  geom <- st_geometry(buildings_wgs)[[i]]
  if (inherits(geom, "MULTIPOLYGON")) {
    lapply(geom, function(poly)
    ring_to_latlng(poly[[1]]))
  } else if (inherits(geom, "POLYGON")) {
    list(ring_to_latlng(geom[[1]]))
  } else {
    list()
  }
})

message("Loaded ", nrow(buildings), " features.
",
       sum(is.na(buildings$sis_logistiek)), "
need review.")

# On first launch, create the manual-check file
as a copy of the input.
if (!file.exists(OUTPUT_GPKG)) {
  message("Creating manual-check file: ",
OUTPUT_GPKG)
  init_out <- buildings
  init_out$lat <- NULL
  init_out$lon <- NULL
  init_out <- st_transform(init_out, orig_crs)
  st_write(init_out, OUTPUT_GPKG, layer =
INPUT_LAYER,
          delete_dsn = TRUE, quiet = TRUE)
  rm(init_out)
}

#user interface

app_css <- HTML("
  body { background: #f5f5f7; }
  #map { width: 100%; height: calc(100vh -
20px); border-radius: 8px; }
```

```

    .side-panel { padding: 12px; background:
white; border-radius: 8px;
    height: calc(100vh - 20px);
overflow-y: auto;
    box-shadow: 0 1px 3px
rgba(0,0,0,0.06); }
    .progress-bar-wrap { background:#e9e9eb;
border-radius:6px; height:8px;
    overflow:hidden; margin-
bottom:8px; }
    .progress-bar-fill { background:#0a84ff;
height:100%; transition:width .2s; }
    .attr-row { display:flex; justify-
content:space-between; padding:4px 0;
    border-bottom: 1px solid #f0f0f2;
font-size: 13px; }
    .attr-row .k { color:#6e6e73; }
    .attr-row .v { color:#1d1d1f; font-weight:500;
max-width:60%;
    text-align:right; word-
break:break-word; }
    .kbd { display:inline-block; padding:2px 6px;
border:1px solid #d2d2d7;
    border-radius:4px; font-
family:monospace; font-size:11px;
    background:#fafafa; margin: 0 2px; }
    .keyhint { font-size: 12px; color: #6e6e73;
margin-top: 8px;
    line-height: 1.6; }
    .cat-buttons { display: grid; grid-template-
columns: 1fr 1fr;
    gap: 6px; margin-top: 10px; }
    .cat-buttons .btn { font-size: 12px; padding:
8px; }
    h4 { margin-top: 0; }
")

ui <- fluidPage(
  theme = bs_theme(version = 5, bootswatch =
"flatly"),
  tags$head(
    tags$style(app_css),
    tags$script(HTML("
    let gmap = null;
    let overlays = [];
    let pendingPayload = null;
    let mapsReady = false;

    function clearOverlays() {
      overlays.forEach(o => o.setMap(null));
      overlays = [];
    }

    function renderMap(payload) {
      if (!mapsReady) {
        pendingPayload = payload;
        return;
      }
      const center = { lat: payload.lat, lng:
payload.lon };
      if (!gmap) {
        gmap = new
google.maps.Map(document.getElementById('map'),
{
          center: center,
          zoom: 19,
          mapTypeId: 'hybrid',
          tilt: 0,
          streetViewControl: true,
          mapTypeControl: true,
          fullscreenControl: true,
          rotateControl: true
        });
      } else {
        gmap.panTo(center);
        gmap.setZoom(19);

```

```

    }

    clearOverlays();

    let rings = payload.rings;
    if (!Array.isArray(rings)) rings =
[rings];
    rings.forEach(ring => {
      if (!Array.isArray(ring)) ring =
[ring];

      const poly = new
google.maps.Polygon({
        paths: ring,
        strokeColor: '#FF3B30',
        strokeOpacity: 1.0,
        strokeWeight: 3,
        fillColor: '#FF3B30',
        fillOpacity: 0.10,
        clickable: false
      });
      poly.setMap(gmap);
      overlays.push(poly);
    });

    const marker = new google.maps.Marker({
      position: center,
      map: gmap,
      opacity: 0.6
    });
    overlays.push(marker);
  }

  window.initGmaps = function() {
    mapsReady = true;
    if (pendingPayload) {
      renderMap(pendingPayload);
      pendingPayload = null;
    }
  };

  window.gm_authFailure = function() {
    const el =
document.getElementById('map');
    if (el) {
      el.innerHTML =
        '<div
style=\\"padding:20px;color:#b00;font-
family:sans-serif\\">' +
        '<b>Google Maps failed to
load.</b><br>' +
        'Check that the Maps JavaScript API
is enabled on this project ' +
        'and that the API key is not blocked
by referrer/IP restrictions.' +
        '</div>';
    }
    console.error('Google Maps auth
failure.');
```

```

    document.addEventListener('keydown',
function(e) {
    const tag = (e.target.tagName ||
    '').toLowerCase();
    if (tag === 'input' || tag ===
    'textarea' || tag === 'select') return;

    const key = e.key.toLowerCase();
    if
    ([ '0', '1', '2', '3', '4', '5', '9' ].includes(key)) {
        Shiny.setInputValue('kb_assign', {
cat: key, ts: Date.now() });
        e.preventDefault();
    } else if (key === 'n' || key ===
    'arrowright') {
        Shiny.setInputValue('kb_skip',
Date.now());
        e.preventDefault();
    } else if (key === 'b' || key ===
    'arrowleft') {
        Shiny.setInputValue('kb_back',
Date.now());
        e.preventDefault();
    } else if (key === 'u') {
        Shiny.setInputValue('kb_undo',
Date.now());
        e.preventDefault();
    }
    });
    "));
  ),

  fluidRow(
    column(
      width = 9,
      div(id = "map")
    ),
    column(
      width = 3,
      div(
        class = "side-panel",
        h4("Review queue"),
        div(class = "progress-bar-wrap",
          div(class = "progress-bar-fill",
            style = "width: 0%; ",
            id = "progressbar")),
        textOutput("progress_text"),
        tags$hr(),

        h5("Building"),
        uiOutput("attr_panel"),

        tags$hr(),
        h5("Assign category"),
        selectInput("category", NULL,
          choices = c("- select -" =
"", CATEGORIES),
          width = "100%"),
        next",
          class = "btn btn-primary w-
100"),

        div(class = "cat-buttons",
          actionButton("q_0", "0
Remove", class = "btn btn-outline-secondary"),
          actionButton("q_1", "1
Trade", class = "btn btn-outline-primary"),
          actionButton("q_2", "2 Logist.",
class = "btn btn-outline-primary"),
          actionButton("q_3", "3 XXL ret.",
class = "btn btn-outline-primary"),
          actionButton("q_4", "4
Tank", class = "btn btn-outline-primary"),
          actionButton("q_5", "5
Bulk", class = "btn btn-outline-primary"),

```

```

          actionButton("q_9", "9
Doubt", class = "btn btn-outline-warning")
        ),

        tags$hr(),
        div(style = "display: flex; gap: 6px; ",
          actionButton("back", "Back", class
= "btn btn-outline-secondary w-50"),
          actionButton("skip", "Skip", class
= "btn btn-outline-secondary w-50")
        ),
        actionButton("undo", "Undo last save",
          class = "btn btn-outline-
danger w-100",
          style = "margin-top:
6px;"),

        div(class = "keyhint",
          tags$b("Shortcuts: "), br(),
          span(class = "kbd", "0"),
          span(class = "kbd", "1"), span(class = "kbd", "2"),
          span(class = "kbd", "3"),
          span(class = "kbd", "4"), span(class = "kbd", "5"),
          span(class = "kbd", "9"), " assign ",
          br(),
          span(class = "kbd", "N"), " skip - ",
          span(class = "kbd", "B"), " back - ",
          span(class = "kbd", "U"), " undo"
        )
      )
    )
  )
)

# create server

server <- function(input, output, session) {

  state <- reactiveValues(
    data = buildings,
    queue =
  which(is.na(buildings$sis_logistiek)),
    queue_pos = 1L,
    last_saved = NULL
  )

  save_to_disk <- function() {
    out <- state$data
    out$.lat <- NULL
    out$.lon <- NULL
    out_proj <- st_transform(out, orig_crs)
    tmp <- paste0(OUTPUT_GPKG, ".tmp.gpkg")
    st_write(out_proj, tmp, layer =
INPUT_LAYER,
      delete_dsn = TRUE, quiet = TRUE)
    file.rename(tmp, OUTPUT_GPKG)
  }

  current_idx <- reactive({
    pos <- state$queue_pos
    if (pos < 1 || pos > length(state$queue))
    return(NA_integer_)
    state$queue[pos]
  })

  observe({
    idx <- current_idx()
    if (is.na(idx)) return()
    payload <- list(
      lat = buildings$.lat[idx],
      lon = buildings$.lon[idx],
      rings = poly_coords[[idx]]
    )
    session$sendCustomMessage("renderMap",
    payload)
  })
}

```

```

    cur_val <- state$data$sis_logistiek[idx]
    updateSelectInput(session, "category",
                      selected = if
(is.na(cur_val)) "" else as.character(cur_val))
  })

  output$progress_text <- renderText({
    total <- length(state$queue)
    done <-
sum(!is.na(state$data$sis_logistiek[state$queue])
)
    sprintf("Row %d of %d in queue - %d / %d
reviewed",
            state$queue_pos, total, done,
total)
  })

  observe({
    total <- length(state$queue)
    done <-
sum(!is.na(state$data$sis_logistiek[state$queue])
)
    pct <- if (total == 0) 0 else round(100 *
done / total)
    session$sendCustomMessage(
      type = "shiny-update-progress",
      message = list(pct = pct)
    )
  })

  output$attr_panel <- renderUI({
    idx <- current_idx()
    if (is.na(idx)) return(div("Queue
complete."))
    row <- state$data[idx, ]
    fmt <- function(x) if (is.null(x)) ||
is.na(x)) "-" else as.character(x)
    fmtnum <- function(x) if (is.na(x)) "-" else
formatC(x, big.mark=",", format="d")

    tagList(
      div(class="attr-row",
        span(class="k", "FID"), span(class="v",
fmt(row$fid))),
      div(class="attr-row",
        span(class="k", "Pand-ID"),
span(class="v", fmt(row$pand_id))),
      div(class="attr-row",
        span(class="k", "Plan"),
span(class="v", fmt(row$PLAN_NAAM))),
      div(class="attr-row",
        span(class="k", "Gebruiksdoel"),
span(class="v", fmt(row$gebruiksdoelen_list))),
      div(class="attr-row",
        span(class="k", "Oppervlakte (m2)"),
span(class="v", fmtnum(row$oppervlakte_calc))),
      div(class="attr-row",
        span(class="k", "Bouwjaar"),
span(class="v", fmt(row$bouwjaar))),
      div(class="attr-row",
        span(class="k", "Banen"),
span(class="v", fmt(row$BANEN))),
      div(class="attr-row",
        span(class="k", "Aantal vestigingen"),
span(class="v", fmt(row$aantal_vestigingen))),
      div(class="attr-row",
        span(class="k", "Size class"),
span(class="v", fmt(row$size_class))),
      div(class="attr-row",
        span(class="k", "Potential type"),
span(class="v", fmt(row$potential_type)))
    )
  })

  #Navigation actions

```

```

    go_next <- function() {
      if (state$queue_pos < length(state$queue))
{
        state$queue_pos <- state$queue_pos + 1L
      } else {
        showNotification("Reached end of queue.",
type = "message")
      }
    }

    go_back <- function() {
      if (state$queue_pos > 1) state$queue_pos <-
state$queue_pos - 1L
    }

    assign_and_advance <- function(cat_val) {
      idx <- current_idx()
      if (is.na(idx)) return()
      prev <- state$data$sis_logistiek[idx]
      state$data$sis_logistiek[idx] <-
as.integer(cat_val)
      state$last_saved <- list(idx = idx, prev =
prev)
      save_to_disk()
      go_next()
    }

    observeEvent(input$save_next, {
      val <- input$category
      if (!nzchar(val)) {
        showNotification("Pick a category first
(or press Skip).",
                        type = "warning",
duration = 2)
        return()
      }
      assign_and_advance(val)
    })

    observeEvent(input$skip, { go_next() })
    observeEvent(input$back, { go_back() })

    observeEvent(input$undo, {
      ls <- state$last_saved
      if (is.null(ls)) {
        showNotification("Nothing to undo.", type
= "warning", duration = 2)
        return()
      }
      state$data$sis_logistiek[ls$idx] <- ls$prev
      save_to_disk()
      state$last_saved <- NULL
      showNotification("Undone.", type =
"message", duration = 2)
    })

    for (cat in names(CATEGORIES)) {
      local({
        this_cat <- cat
        observeEvent(input[[paste0("q_",
this_cat)]], {
          assign_and_advance(this_cat)
        })
      })
    }

    observeEvent(input$kb_assign, {
      if (!is.null(input$kb_assign$cat)) {
        assign_and_advance(input$kb_assign$cat)
      }
    })

    observeEvent(input$kb_skip, { go_next() })
    observeEvent(input$kb_back, { go_back() })
    observeEvent(input$kb_undo, {
      ls <- state$last_saved
      if (is.null(ls)) return()
      state$data$sis_logistiek[ls$idx] <- ls$prev
    })

```

```
    save_to_disk()
    state$last_saved <- NULL
    showNotification("Undone.", type =
"message", duration = 2)
  })
```

```
  session$onSessionEnded(function() {
```

```
    isolate(save_to_disk())
  })
}
```

```
#shiny
shinyApp(ui, server)
```

Bijlage E: Opzet MNL Model (Scripts 4 t/m 6)

```
#
=====
# SCRIPT 4: OPBOUW REGRESSIEMATRIX & GRIDDING
(100X100M) FOR MNL-MODEL
# ONDERZOEK: MONITOR LOGISTIEK VASTGOED 2026 -
VU AMSTERDAM
#
=====

# 1. INITIALISATIE & PACKAGES LOADING
cat("📦 Laden van essentiële packages...\n")
library(sf) # Vector data handling
(GPKG, SHP, GML)
library(terra) # High-performance Raster
data engine (GeoTIFF)
library(tidyverse) # Data wrangling, piping &
pivoting
library(tidyterra) # Visuele controle van
rasters via ggplot
library(data.table) # Supersnelle tabel-
aggregatie voor PC6 data
library(here)

RD_NEW <- "EPSG:28992"

#
=====
# 2. DATA IMPORT & PROJECTIE CONTROL
#
=====

cat("📂 Inlezen van raster- en
vectordatabestanden...\n")

if(!file.exists("data_input/rsopafr_100.tif")) {
  stop("❌ Kritieke fout: rsopafr_100.tif is nog
niet gedownload of aanwezig!")
}

master_grid <-
rast("data_input/rsopafr_100.tif")
crs(master_grid) <- RD_NEW

banen_raster <-
rast("data_input/rshdban_100.tif")
inwoners_raster <-
rast("data_input/rshdinw_100.tif")
crs(banen_raster) <- RD_NEW
crs(inwoners_raster) <- RD_NEW

panden_sf <-
st_read("data_input/eindproduct_overdracht.gpkg")
)%>%
  st_transform(crs = RD_NEW)

wegvakken_raw <-
st_read("data_input/Wegvakken.gpkg")
terminals_sf <-
st_read("data_input/terminals.gpkg") %>%
  st_transform(RD_NEW)
nnn_sf <- st_read("data_input/inspire-pv-
ps.nlps-nnn.gml") %>% st_transform(RD_NEW)

# TYPE-VEILIGE AGGREGATIE VAN DE NETCONGESTIE
SHAPEFILE
cat("⚡ Start slimme aggregatie van de numerieke
PC6 netcongestiekaart...\n")
netcongestie_raw <-
st_read("data_input/netcongestie.shp") %>%

  st_transform(RD_NEW) %>%
  st_centroid()

coords_nc <- st_coordinates(netcongestie_raw)
netcongestie_dt <-
as.data.table(st_drop_geometry(netcongestie_raw)
)
netcongestie_dt[, `:=`(x = coords_nc[, "X"], y =
coords_nc[, "Y"])]

netcongestie_dt[, `:=`(
  grid_x = floor(x / 100) * 100 + 50,
  grid_y = floor(y / 100) * 100 + 50
)]

netcongestie_grid_dt <- netcongestie_dt[, .(
  netcongestie_dummy = as.numeric(any(afname ==
3 | invoeding == 3, na.rm = TRUE))
), by = .(grid_x, grid_y)]
netcongestie_grid_dt[is.na(netcongestie_dummy),
netcongestie_dummy := 0]

netcongestie_raster_raw <-
rast(netcongestie_grid_dt, type = "xyz", crs =
RD_NEW)
netcongestie_raster <-
resample(netcongestie_raster_raw, master_grid,
method = "near")
netcongestie_raster <-
classify(netcongestie_raster, matrix(c(NA, 0),
ncol = 2))
names(netcongestie_raster) <-
"netcongestie_dummy"

#
=====
# 3. GEOPTIMALISEERDE SNELE INFRASTRUCTUUR
FILTERING & RASTER-AFSTAND
#
=====

cat("🚗 Wegennet en Intermodale Terminals
verwerken...\n")

snelwegen_sf <- wegvakken_raw %>%
  st_transform(crs = RD_NEW) %>%
  filter(WEGTYPE %in% c("HT", "HW"))

snelwegen_raster_presence <-
rasterize(vect(snelwegen_sf), master_grid, field
= 1, background = NA)
afstand_snelweg_raster <-
distance(snelwegen_raster_presence)
names(afstand_snelweg_raster) <-
"afstand_snelweg"

terminals_sf <- terminals_sf %>%
  st_make_valid()
terminal_coords <- terminals_sf %>%
  st_centroid() %>% st_coordinates() %>%
  as.data.frame()
terminals_clean_sf <- st_as_sf(terminal_coords,
coords = c("X", "Y"), crs = RD_NEW)

terminals_raster_presence <-
rasterize(vect(terminals_clean_sf), master_grid,
field = 1, background = NA)
afstand_terminal_raster <-
distance(terminals_raster_presence)
names(afstand_terminal_raster) <-
"afstand_terminal"
```

```

# --- NIEUW: EXTRA BEREIKBAARHEIDS- & MARKT-
VARIABLEN INLEZEN ---
cat("🚢 Inlezen nieuwe externe variabelen
(Zeehavens, Airports, Reistijd)...\n")
dist_zeehaven <-
rast("data_input/DistTo_Zeehaven_2019.tif") %>%
resample(master_grid, method = "bilinear")
dist_airport <-
rast("data_input/DistTo_Airports_2019_Top5_BRT.tif") %>%
resample(master_grid, method =
"bilinear")
reistijd_100k <-
rast("data_input/Reistijd100kInw.tif") %>%
resample(master_grid, method = "bilinear")

names(dist_zeehaven) <- "afstand_zeehaven"
names(dist_airport) <- "afstand_luchthaven"
names(reistijd_100k) <- "reistijd_100k"

#
=====
# 4. AFHANKELIJKE VARIABELE, ACADEMISCHE CBS-
DATA & NIEUWE BELEIDSKAARTEN
#
=====
cat("🗺️ Rasteriseren van afhankelijke variabele
en beleidslagen...\n")

transitie_panden_sf <- panden_sf %>%
  filter(bouwjaar >= 2020 & bouwjaar <= 2024)
%>%
  mutate(Y_klasse = case_when(
    size_class %in% c("L (10k-20k)", "XL (20k-
40k)") ~ 1,
    size_class == "XXL (>40k)"
~ 2,
    TRUE
~ NA_real_
)) %>%
  filter(!is.na(Y_klasse))

y_raster <- rasterize(vect(transitie_panden_sf),
master_grid, field = "Y_klasse", fun = "max")
%>%
  classify(rcl = matrix(c(NA, 0), ncol = 2))
names(y_raster) <- "transitie_status"

op_ibis_raster <- rasterize(vect(panden_sf),
master_grid, field = "op_ibis", fun = "max") %>%
  classify(rcl = matrix(c(NA, 0), ncol = 2))
names(op_ibis_raster) <- "op_ibis"

nnn_raster <- rasterize(vect(nnn_sf),
master_grid, fun = "count") %>%
  classify(matrix(c(NA, 0), ncol = 2)) %>%
  classify(matrix(c(1, Inf, 1), ncol = 3))
names(nnn_raster) <- "restrictie_nnn"

cbs_space_sf <-
st_read("data_input/cbs_vk100_2024_v1.gpkg") %>%
  st_transform(RD_NEW) %>%
  mutate(
    aantal_woningen_clean =
ifelse(aantal_woningen < 0 | aantal_woningen >
500, 0, aantal_woningen),
    available_space_pct = 100 -
pmin((aantal_woningen_clean * 2), 100),
    available_space_pct =
replace_na(available_space_pct, 100)
  )

available_space_raster <-
rasterize(vect(cbs_space_sf), master_grid, field
= "available_space_pct")

```

```

names(available_space_raster) <-
"available_space_pct"

# --- NIEUW: DE OFFICIËLE BELEIDSKAARTEN INLEZEN
---
cat("📄 Officiële Beleidskaarten (Stimuli &
Restricties) inlezen en uitlijnen...\n")
# Zorg dat de namen van de TIF bestanden hier
exact kloppen met jouw hardeschijf!
beleid_stim <-
rast("data_input/Stimuli_Werken_Hard.tif") %>%
resample(master_grid, method = "near")
beleid_rest <-
rast("data_input/Restricties_Werken_ZeerHard.tif")
%>% resample(master_grid, method = "near")

# Vervang NA (geen beleid op die plek) door 0
beleid_stim <- classify(beleid_stim, cbind(NA,
0))
beleid_rest <- classify(beleid_rest, cbind(NA,
0))

names(beleid_stim) <- "beleid_stimulering"
names(beleid_rest) <- "beleid_restrictie"

#
=====
# 5. CONSOLIDATIE NAAR DE ULTIEME
REGRESSIEMATRIX
#
=====
cat("📊 Rasters stapelen en converteren naar
statistische matrix...\n")

regressie_stack <- c(
  y_raster, master_grid, banen_raster,
  inwoners_raster,
  afstand_snelweg_raster,
  afstand_terminal_raster, dist_zeehaven,
  dist_airport, reistijd_100k,
  op_ibis_raster, nnn_raster, beleid_stim,
  beleid_rest,
  available_space_raster, netcongestie_raster
)

regressie_matrix <-
as.data.frame(regressie_stack, xy = TRUE, na.rm
= FALSE) %>%
  as_tibble() %>%
  rename(afstand_op_afrit = rsopafr_100)

#
=====
# 6 & 7. ECONOMISCHE CONTEXT INBAKKEN (GEEN PC2-
RUIS MEER!)
#
=====
cat("🔪 Opschonen en grondprijzen
inbakken...\n")

regressie_matrix_clean <- regressie_matrix %>%
  filter(!is.na(afstand_op_afrit) |
!is.na(rshdban_100)) %>%
  mutate(
    rshdban_100 = replace_na(rshdban_100, 0),
    rshdinw_100 = replace_na(rshdinw_100, 0),
    available_space_pct =
replace_na(available_space_pct, 100),
    # Kleine data-imputatie voor nieuwe rasters
als ze aan de landsgrenzen wegvallen

```

```

    afstand_zeehaven =
    replace_na(afstand_zeehaven,
    max(afstand_zeehaven, na.rm=T)),
    afstand_luchthaven =
    replace_na(afstand_luchthaven,
    max(afstand_luchthaven, na.rm=T)),
    reistijd_100k = replace_na(reistijd_100k,
    mean(reistijd_100k, na.rm=T))
  )

matrix_punten_sf <- regressie_matrix_clean %>%
  st_as_sf(coords = c("x", "y"), crs = RD_NEW,
  remove = FALSE)

grondprij_s sf <-
st_read("data_input/grondprij_s.shp") %>%
  st_set_crs(RD_NEW) %>%
  st_transform(crs = RD_NEW) %>%
  st_make_valid()
matrix_met_grondprij_s <-
st_join(matrix_punten_sf, grondprij_s sf, join =
st_intersects)

regressie_matrix_final <- matrix_met_grondprij_s
%>%
  st_drop_geometry() %>%
  as_tibble() %>%
  mutate(
    transitie_status =
as.factor(transitie_status),

    # FIX: Vervang grondprijzen van 0 of NA door
het landelijk logistiek gemiddelde (85)
    grondprij_s = ifelse(is.na(grondprij_s) |
grondprij_s <= 0, 85, grondprij_s),
    # FIX: Cap extreme uitschieters (anti-Zuidas
ruis)
    grondprij_s = ifelse(grondprij_s > 400, 400,
grondprij_s)
  ) %>%
  filter(!is.na(afstand_op_afrit))

#
=====
# 8. CASE-CONTROL SAMPLING & HARDER REPLICATIE-
OPSLAG
#
=====
cat("🚧 Case-control sampling uitvoeren...\n")

set.seed(123)
cases_l_xl <- regressie_matrix_final %>%
filter(transitie_status == 1)
cases_xxl <- regressie_matrix_final %>%
filter(transitie_status == 2)
controls <- regressie_matrix_final %>%
filter(transitie_status == 0)

controls_sample <- controls %>% sample_n(5000)

model_data <- bind_rows(cases_l_xl, cases_xxl,
controls_sample) %>%
  mutate(transitie_status =
factor(transitie_status, levels = c("0", "1",
"2")))

write_csv(model_data,
"data_input/model_data_sampled.csv")
write_csv(regressie_matrix_final,
"data_input/transities_kaart_data.csv")
cat("🚧 Bestanden succesvol weggeschreven! Fase
4 volledig afgerond.\n")

```

```
#
=====
# SCRIPT 05: LOGISTIC HOTSPOT MAP & PC2 REGIO
ANALYSIS (HOOFDSTUK 4)
# ONDERZOEK: MONITOR LOGISTIEK VASTGOED 2026 -
VU AMSTERDAM
#
=====

# 1. INITIALISATIE & PACKAGES LOADING
cat("📦 Laden van basisscripts en
pakketten...\n")
library(tidyverse)
library(sf)
library(here)

RD_NEW <- "EPSG:28992"

#
=====
# 2. DATA IMPORT & HOTSPOT AGGREGATIE
#
=====
cat("📄 Inlezen van de transitiedata...\n")
if(!file.exists("data_input/transities_kaart_data.csv")) {
  stop("❌ Kritieke fout:
transities_kaart_data.csv is niet gevonden!")
}

kaart_data_raw <-
read_csv("data_input/transities_kaart_data.csv",
show_col_types = FALSE) %>%
# FIX: Bereken PC2 razendsnel ter plekke,
speciaal voor de Top-15 tabel!
mutate(PC2 =
as.numeric(substr(as.character(PC4), 1, 2)))

# Filter en formatteer de datapunten voor de
kaart
transities_clean <- kaart_data_raw %>%
  filter(transitie_status %in% c(1, 2)) %>%
  mutate(
    Transitie_Label = case_when(
      transitie_status == 1 ~ "Nieuwe
distributie (L/XL: 10k - 40k m²)",
      transitie_status == 2 ~ "Mega-distributie
(XXL: >40k m²)"
    ),
    Transitie_Label = factor(Transitie_Label,
levels = c(
  "Mega-distributie (XXL: >40k m²)",
  "Nieuwe distributie (L/XL: 10k - 40k m²)"
))
  )

# 🇳🇱 Genereren van de Top 15 Logistieke Regio's
op PC2-niveau voor Hoofdstuk 4
cat("🇳🇱 Berekenen van de PC2 Hotspot Tabel voor
de tekst...\n")
hotspot_tabel <- kaart_data_raw %>%
  filter(transitie_status %in% c(1, 2)) %>%
  group_by(PC2) %>%
  summarise(
    `Nieuwe distributie (L/XL)` =
sum(transitie_status == 1),
    `Mega-distributie (XXL)` =
sum(transitie_status == 2),
    Totaal_Transities = n(),
    .groups = "drop"
  ) %>%
```

```
arrange(desc(Totaal_Transities))

cat("\n=====
\n")
cat("🇳🇱 TOP 10 LOGISTIEKE HOTSPOTS VAN
NEDERLAND (PC2) 2020-2024 🇳🇱\n")
cat("=====
\n")
print(as.data.frame(head(hotspot_tabel, 10)),
row.names = FALSE)
cat("=====
\n\n")

#
=====
# 3. BASEMAP FIX: OFFICIËLE CBS PROVINCIEGRENZEN
VIA PDOK
#
=====
cat("🌲 Officiële CBS provincievlakken ophalen
via PDOK WFS-service...\n")

pdok_url <-
"https://service.pdok.nl/cbs/gebiedsindelingen/2
023/wfs/vl_0?request=GetFeature&service=WFS&vers
ion=2.0.0&typeName=provincie_gegeneraliseerd"

nederland_base <- st_read(pdok_url, quiet =
TRUE) %>%
  st_transform(crs = RD_NEW)

#
=====
# 4. HIGH-CONTRAST ACADEMIC RENDERING (THE FINAL
DESIGN)
#
=====
cat("🌲 De definitieve kaart renderen met egale
achtergrond...\n")

# Maak van de platte tabel een ruimtelijk
puntenbestand op basis van de x en y kolommen
transities_sf <- transities_clean %>%
  st_as_sf(coords = c("x", "y"), crs = RD_NEW)
%>%
  arrange(desc(Transitie_Label)) # Zorgt dat de
XXL-stippen bovenop liggen

# Harde RD-coördinaten van de landsgrenzen
(negeert de Noordzee-vlakken van het CBS)
xlim_fixed <- c(15000, 280000)
ylim_fixed <- c(305000, 620000)

ggplot() +
  geom_sf(data = nederland_base, fill =
"#f1f3f5", color = "#ffffff", size = 0.4) +
  geom_sf(
    data = transities_sf,
    aes(fill = Transitie_Label),
    color = "#ffffff",
    stroke = 0.4,
    alpha = 0.95,
    size = 1.8,
    shape = 21
  ) +
  scale_fill_manual(
    values = c(
      "Mega-distributie (XXL: >40k m²)" =
"#d9534f",
      "Nieuwe distributie (L/XL: 10k - 40k m²)"
= "#1f3a52"
```

```

    )
  ) +
  theme_minimal() +
  theme(
    text = element_text(family = "sans"),
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    axis.text = element_blank(),
    axis.title = element_blank(),
    axis.ticks = element_blank(),

    # Gecorrigeerde legenda: onder elkaar
    (vertical) in plaats van naast elkaar
    legend.position = "bottom",
    legend.direction = "vertical",
    legend.title = element_blank(),
    legend.text = element_text(size = 9, color =
"#2b2b2b", face = "bold"),
    legend.margin = margin(t = 15, b = 10),

    plot.background = element_rect(fill =
"#ffffff", color = NA),
    panel.background = element_rect(fill =
"#ffffff", color = NA),
    plot.margin = margin(20, 20, 20, 20)
  ) +
  coord_sf(xlim = xlim_fixed, ylim = ylim_fixed,
expand = FALSE, datum = NA)

#
=====
=====
# 5. HIGH-DPI EXPORT
#
=====
=====
ggsave(

"data_input/kaart_hoofdstuk4_schaalsprong_PRO.pn
g",
  width = 9,
  height = 10.5,
  dpi = 400,
  bg = "white"
)

cat("🚩 DOEL BEREIKT! De kaart staat nu
definitief op je schijf.\n")

```

```

#
=====
# SCRIPT 07: SCHATTEN VAN MULTINOMIAAL LOGIT
MODEL (MNL) & HEATMAPS
# ONDERZOEK: MONITOR LOGISTIEK VASTGOED 2026 -
VU AMSTERDAM
#
=====

library(mlogit)
library(tidyverse)
library(terra)
library(here)

cat("📁 Modeldata inlezen en
herformatteren...\n")
if(!file.exists("data_input/model_data_sampled.csv")) {
  stop("❌ Kritieke fout: model_data_sampled.csv
is niet gevonden!")
}

model_data_raw <-
read_csv("data_input/model_data_sampled.csv",
show_col_types = FALSE)

# Opschonen, SCHALEN en de beproefde PC2-dummies
herstellen!
model_data <- model_data_raw %>%
  mutate(
    transitie_status = factor(transitie_status,
levels = c("0", "1", "2")),

    # Afstanden naar kilometers
    afstand_snelweg_km = afstand_snelweg /
1000,
    afstand_op_afrit_km = afstand_op_afrit /
1000,
    afstand_terminal_km = afstand_terminal /
1000,
    afstand_luchthaven_km = afstand_luchthaven /
1000,

    # Banen naar 1000-tallen
    banen_1k = rshdban_100 / 1000,

    # DE VEILIGE BELEIDS-DUMMIES (Werkt 100x
beter dan de defecte rasterkaarten)
    PC2 = as.numeric(substr(as.character(PC4),
1, 2)),
    restrictie_randstad = ifelse(PC2 %in%
c(20:33, 35:42), 1, 0),
    beleid_stimulering = ifelse(PC2 %in% c(26,
59, 47, 48, 54, 66, 69), 1, 0)
  ) %>%
  as.data.frame() # Tibble decompositie-fix voor
mlogit

cat("📊 Data transformeren naar mlogit index
matrix...\n")
model_data_idx <- dfidx(
  model_data,
  choice = "transitie_status",
  shape = "wide",
  alt.levels = c("0", "1", "2")
)

#
=====
# 1. SCHATTEN VAN MODEL 1: HET ECONOMISCH-
GEOGRAFISCHE BASELINE MODEL

```

```

#
=====
cat("🚀 Model 1 schatten (Baseline)...\n")

model_1 <- mlogit(
  transitie_status ~ 0 | afstand_snelweg_km +
afstand_op_afrit_km + banen_1k + grondprijis,
  data = model_data_idx,
  reflevel = "0"
)

#
=====
# 2. SCHATTEN VAN MODEL 2 (INTEGRAAL MET
LUCHTHAVENS & BEPROEFT BELEID)
#
=====
cat("🚀 Model 2 schatten (Integraal
model)...\n")

model_2 <- mlogit(
  transitie_status ~ 0 | afstand_snelweg_km +
afstand_op_afrit_km + afstand_terminal_km +
afstand_luchthaven_km + banen_1k +
grondprijis + restrictie_randstad +
beleid_stimulering + netcongestie_dummy,
  data = model_data_idx,
  reflevel = "0"
)

#
=====
# 3. OUTPUT GENEREREN
#
=====
cat("\n=====
\n")

cat("📊 SUMMARY BASELINE MODEL
(MODEL 1)\n")
cat("=====
\n")
print(summary(model_1))

cat("\n=====
\n")

cat("🚀 SUMMARY INTEGRAL POLICY
MODEL (MODEL 2)\n")
cat("=====
\n")
print(summary(model_2))
cat("=====
\n\n")

# Pseudo R2 berekenen
pseudo_r2_m1 <- 1 - (as.numeric(logLik(model_1))
/ as.numeric(model_1$logLik0))
pseudo_r2_m2 <- 1 - (as.numeric(logLik(model_2))
/ as.numeric(model_2$logLik0))

cat("📊 McFadden Pseudo R² van Model 1
(Baseline):", round(pseudo_r2_m1, 4), "\n")
cat("🚀 McFadden Pseudo R² van Model 2
(Compleet):", round(pseudo_r2_m2, 4), "\n")

#
=====
# 4. HEATMAPS GENEREREN (VOLAUTOMATISCH VIA
MODEL OUTPUT)

```

```

#
=====
cat("\n📊 START: Genereren van de ruimtelijke
voorspellingskaarten (Heatmaps)...\n")

cat("📂 Inlezen van de landelijke grid-
data...\n")
full_grid_data <-
read_csv("data_input/transities_kaart_data.csv",
show_col_types = FALSE)

# Dezelfde schaling én PC2 dummies toepassen op
de hele landkaart
full_grid_scaled <- full_grid_data %>%
mutate(
  snelweg_km = afstand_snelweg / 1000,
  afrit_km = afstand_op_afrit / 1000,
  terminal_km = afstand_terminal / 1000,
  luchthaven_km = afstand_luchthaven / 1000,
  banen_1k = rshdban_100 / 1000,

  # Bereken de dummies voor heel Nederland
  PC2 = as.numeric(substr(as.character(PC4),
1, 2)),
  restrictie_randstad = ifelse(PC2 %in%
c(20:33, 35:42), 1, 0),
  beleid_stimulering = ifelse(PC2 %in% c(26,
59, 47, 48, 54, 66, 69), 1, 0)
)

cat("📊 Coëfficiënten automatisch uit Model 2
extraheren...\n")
cfs <- coef(model_2)

cat("📊 Kansberekening (Log-Odds) toepassen op
3,3 miljoen gridcellen...\n")
heatmap_data <- full_grid_scaled %>%
mutate(
  V1 = cfs["(Intercept):1"] +
(cfs["afstand_snelweg_km:1"] *
snelweg_km) +
(cfs["afstand_op_afrit_km:1"] *
afrit_km) +
(cfs["afstand_terminal_km:1"] *
terminal_km) +
(cfs["afstand_luchthaven_km:1"] *
luchthaven_km) +
(cfs["banen_1k:1"] *
banen_1k) +
(cfs["grondprijs:1"] *
grondprijs) +
(cfs["restrictie_randstad:1"] *
restrictie_randstad) +
(cfs["beleid_stimulering:1"] *
beleid_stimulering) +
(cfs["netcongestie_dummy:1"] *
netcongestie_dummy),

  V2 = cfs["(Intercept):2"] +
(cfs["afstand_snelweg_km:2"] *
snelweg_km) +
(cfs["afstand_op_afrit_km:2"] *
afrit_km) +
(cfs["afstand_terminal_km:2"] *
terminal_km) +
(cfs["afstand_luchthaven_km:2"] *
luchthaven_km) +
(cfs["banen_1k:2"] *
banen_1k) +
(cfs["grondprijs:2"] *
grondprijs) +
(cfs["restrictie_randstad:2"] *
restrictie_randstad) +

```

```

(cfs["beleid_stimulering:2"] *
beleid_stimulering) +
(cfs["netcongestie_dummy:2"] *
netcongestie_dummy),

  noemer = 1 + exp(V1) + exp(V2),
  Prob_L_XL_raw = exp(V1) / noemer,
  Prob_XXL_raw = exp(V2) / noemer,

  # Ruimtelijke uitsluitingsmaskers toepassen
  Prob_L_XL_final = ifelse(restrictie_nnn ==
1, 0, Prob_L_XL_raw),
  Prob_XXL_final = ifelse(restrictie_nnn == 1
| available_space_pct < 100, 0, Prob_XXL_raw)
)

cat("📂 Omzetten naar GeoTIFF Rasters en
exporteren...\n")
RD_NEW <- "EPSG:28992"

raster_l_xl <- rast(heatmap_data %>% select(x,
y, Prob_L_XL_final), type = "xyz", crs = RD_NEW)
raster_xxl <- rast(heatmap_data %>% select(x,
y, Prob_XXL_final), type = "xyz", crs = RD_NEW)

writeRaster(raster_l_xl,
"data_input/Heatmap_Kans_L_XL.tif", overwrite =
TRUE)
writeRaster(raster_xxl,
"data_input/Heatmap_Kans_XXL.tif", overwrite =
TRUE)

cat("\n✅ SUCCES! De regressie is gedraaid en de
Heatmaps staan klaar!\n")

```

