



university of
 groningen

Faculteit Gedrags- en
 Maatschappijwetenschappen



RDW / Politie / VbV

VOERTUIGKATVANGERS

*Onderzoek naar de kenmerken van voertuigkatvangers en hun impact
 op de samenleving*

Roma Lamain
 S4104080
 Scriptieopzet
 Master Sociologie
 Criminaliteit en Veiligheid
 Begeleider: G. Huitsing
 Referent: F. Giardini

Voorwoord

Voor u ligt mijn masterscriptie ‘Voertuigkatvangers: Onderzoek naar de kenmerken van voertuigkatvangers en hun impact op de samenleving’, die is geschreven ter voltooiing van mijn masteropleiding Sociologie met als specialisatie Criminaliteit en Veiligheid aan de Rijksuniversiteit Groningen. Dit onderzoek is in opdracht uitgevoerd van de RDW en het Centraal Justitieel Incassobureau (CJIB). Er is tot op heden weinig onderzoek gedaan naar voertuigkatvangers waardoor er behoefte was aan meer informatie. Gedurende de periode van februari 2023 tot april 2023 heb ik mijn afstudeerstage mogen volbrengen bij de RDW op de afdeling van Het Landelijk Intelligence- en expertisecentrum Voertuigcriminaliteit (LIV).

Ik wil mijn stagebegeleider Jeroen de Graaf bedanken voor zijn begeleiding en zijn hulp bij het creëren van de datasets. Ook wil ik mijn half-stagebegeleider, Arend Eiting, bedanken voor zijn unieke sociologische kennis binnen de RDW en de hulp die hij mij daardoor heeft kunnen bieden. Verder wil ik mijn scriptiebegeleider, Gijs Huitsing, bedanken voor de heldere en snelle begeleiding gedurende het gehele proces. Ik wil ook graag mijn referent, Francesca Giardini, bedanken voor haar waardevolle feedback op het theoretisch kader, wat mij enorm heeft geholpen.

Tot slot wil ik mijn familie bedanken, die altijd voor mij klaarstaan en mij hebben gesteund gedurende mijn stage en het onderzoeksproces. In het bijzonder wil ik mijn moeder bedanken, die vanwege haar functie bij de RDW heeft meegedacht over verschillende aspecten en heeft bijgedragen aan het bedenken van oplossingen.

Roma Lamain

Veendam, 1 februari 2024

Abstract

Voertuigkatvangers spelen een rol in georganiseerde criminaliteit door voertuigen op hun naam te registreren voor de werkelijke houders, waardoor de ware identiteit van de houder verborgen blijft en de voertuigen kunnen worden gebruikt voor criminele doeleinden. Er is tot op heden beperkt onderzoek gedaan naar voertuigkatvangers en hun impact op de samenleving. De probleemstelling die centraal staat in dit exploratieve onderzoek luidt als volgt: "Wat kenmerkt katvangers en wat zijn de mogelijke effecten van hun faciliterende rol op de veiligheid in de samenleving?". Het onderzoek maakt gebruik van literatuuronderzoek en data-analyse van het kentekenregister van de RDW en data van het CBS. De eerste hoofdhypothese die wordt opgesteld is dat katvangers (financieel) kwetsbaarder zijn dan niet-katvangers. De resultaten van het onderzoek tonen aan dat katvangers vaak jonge Nederlandse mannen zijn met financiële moeilijkheden. De eerste hoofdhypothese wordt ondersteund. Gebaseerd op literatuur wordt de tweede hoofdhypothese geformuleerd: voertuigen op naam van katvangers hebben een negatieve impact op de verkeersveiligheid. De resultaten ondersteunen deze verwachting. De voertuigen zijn minder vaak APK-gekeurd en verzekerd en katvangers hebben meer boetes open naam staan. Het is aannemelijk dat de criminelen roekelozer rijden en dat de voertuigen vaker worden ingezet voor criminele doeleinden. De pakkans voor criminelen is klein dankzij de katvangers, die zij als dekmantel gebruiken. Om het katvangerprobleem aan te pakken, worden preventieve maatregelen voorgesteld die gericht zijn op het bemoeilijken van het op naam zetten van een voertuig zoals het instellen van strengere controles bij het verkrijgen van een BSN en meer wegcontroles gericht op voertuigen die mogelijk door katvangers worden gebruikt. Op deze manier zou de identiteit van de criminelen die in de voertuigen rijden sneller achterhaald kunnen worden. Het is belangrijk om te benadrukken dat bepaalde gegevens niet zijn opgenomen in de datasets vanwege privacyoverwegingen. Hierbij gaat het voornamelijk om specifieke woonadressen en kan er bijvoorbeeld niet worden achterhaald of het gaat om een bedrijf met meerdere voertuigen. Ook is het daardoor onduidelijk of katvangers vaak dezelfde adressen opgeven of gebruikmaken van nep-adressen. Dit is relevant voor vervolgonderzoek waar ook aanbevelingen voor worden gedaan.

Inhoudsopgave

Inleiding.....	4
Theoretisch kader.....	6
Wat zijn katvangers en hoe gaan zij te werk?	6
Financiële kwetsbaarheid van katvangers.....	7
Achtergrondkenmerken	9
(Verkeers)veiligheid.....	9
Dit onderzoek	12
Methode.....	14
Steekproef en design.....	14
Personen en voertuigen.....	14
Variabelen	15
Datasets – achtergrondkenmerken.....	15
Datasets – voertuigen.....	17
Analyseplan.....	17
Resultaten	18
Descriptieve statistieken persoonsniveau.....	18
Modelfit.....	21
Assumpties en outliers	23
Analyse – Logistische regressie achtergrondkenmerken	23
Descriptieve statistieken voertuigkenmerken	25
Modelfit.....	27
Assumpties en outliers	27
Analyse – Logistische regressie voertuigkenmerken.....	30
Conclusie en discussie.....	31
Katvangers als fiscale wanbetalers.....	31
Achtergrondkenmerken	32
Impact op de verkeersveiligheid.....	33
Preventie van katvangers	35
Beperkingen.....	38
Literatuurlijst.....	40
Bijlage 1 – operationalisaties variabelen.....	43
Bijlage 2 – analyses	58
Bijlage 3 – assumpties en outliers.....	71

Inleiding

Katvanger, geldezel, money mule, stroman... dit zijn termen voor een persoon die dient als een dekmantel voor een crimineel. Ze worden ingezet voor verschillende soorten georganiseerde criminaliteit, zoals witwassen, identiteitsfraude en drugscriminaliteit (Versprille, 2022). Dit doen ze door hun identiteit uit te lenen aan een ander persoon, zodat die in de luwte criminele activiteiten kan uitvoeren. In deze scriptie wordt onderzoek gedaan naar voertuigkatvangers, die in ruil voor een kleine vergoeding een voertuig op hun naam registreren, waardoor de werkelijke houder onbekend blijft. De daadwerkelijke houder gebruikt de voertuigen vervolgens zelf voor criminele doeleinden. Zo worden de voertuigen bijvoorbeeld gebruikt voor overvallen of liquidaties (RDW, z.d.-a). Katvangers fungeren als facilitators van georganiseerde criminaliteit, waardoor criminelen straffeloos hun misdaden kunnen plegen en niet vervolgd zullen worden. De crimineel zal blijven doorgaan met het plegen van delicten zonder enige consequenties waardoor het probleem zal blijven bestaan. Zodra een katvanger bekend is bij de politie en RDW zal er weer een nieuwe katvanger worden gerekruteerd en begint het hele proces opnieuw.

Er staat op dit moment een groot aantal individuen geregistreerd als katvanger bij de RDW en dit aantal blijft groeien. Het is een complex probleem omdat er sprake is van een samenwerking tussen crimineel en katvanger. Vanwege de anonimiteit van de crimineel is het onhaalbaar om hen te kunnen bereiken en het probleem op te lossen. Hierdoor wordt er vooralsnog weinig aandacht besteed aan de katvangers. Andere kwesties lijken vaak urgenter en gemakkelijker op te lossen, waardoor het probleem van katvangers laag op de prioriteitenlijst staat. Er is tot op heden ook weinig onderzoek naar voertuigkatvangers gedaan waardoor er nog weinig passende preventieve maatregelen zijn opgesteld.

Het onderzoek naar voertuigkatvangers is sociologisch relevant omdat er dieper wordt ingegaan op de ongelijkheden in de samenleving. Katvangers zijn mogelijk vaak kwetsbare individuen. Ze wonen naar verwachting vaak in achtergestelde buurten en hebben geldproblemen. Hierdoor zijn zij vaak niet in staat of bereid om de boetes te betalen. Uiteindelijk worden verschillende belastingen niet afgedragen en de boetes die open blijven staan kunnen hoog oplopen (van Noordenburg, 2022). Verder hebben katvangers mogelijk een negatieve invloed op de (verkeers)veiligheid. Voertuigen op naam van katvangers zijn vaker betrokken bij verkeersovertredingen. (RDW, z.d.-a). Er wordt vaker roekeloos mee gereden, bijvoorbeeld door het overschrijden van de maximumsnelheid. De katvangers ontvangen de boetes voor het wangedrag van de eigenlijke eigenaren, wat problemen oplevert voor de overheid. Katvangers kunnen een faciliterende rol hebben in de georganiseerde criminaliteit, zij zorgen dat criminelen buiten het zicht blijven waardoor zij ongestoord verder kunnen gaan met het plegen van misdrijven. Dit heeft effect op iedereen in de samenleving en door meer informatie hierover te verzamelen kan er een passend beleid worden gecreëerd met deze inzichten. Door inzicht te krijgen in de factoren die bijdragen aan waarom

zij betrokken raken bij deze criminele activiteiten kan er een helder beeld worden geschetst hoe dit voorkomen kan worden. Daarom wordt de volgende hoofdvraag opgesteld:

Wat kenmerkt katvangers en wat zijn de mogelijke effecten van hun faciliterende rol op de veiligheid in de samenleving?

Er worden ook deelvragen opgesteld die helpen om de probleemstelling beter te begrijpen en te kunnen beantwoorden:

1. Wat zijn de verschillen tussen katvangers en niet-katvangers en welke indicatoren dragen bij aan de mogelijke kwetsbaarheid van katvangers?
2. Wat voor impact hebben katvangers op de verkeersveiligheid?
3. Welke preventieve maatregelen kunnen met kennis over achtergrondkenmerken opgesteld om het inzetten van katvangers te voorkomen?

Het huidige onderzoek heeft als eerste het doel inzicht te verschaffen in de achtergrondkenmerken en verschillen tussen katvangers en niet-katvangers. Dit wordt gedaan door middel van literatuuronderzoek, maar ook door data te analyseren afkomstig uit het kentekenregister van de RDW en van het Centraal Bureau Statistiek (CBS). Op basis van deze informatie kunnen passende maatregelen worden getroffen die gebaseerd zijn op de meest voorkomende kenmerken. Bijvoorbeeld, als blijkt dat katvangers vaak financieel kwetsbare adolescenten zijn, kunnen maatregelen specifiek op deze groep worden afgestemd. Het tweede doel is het onderzoeken van het effect van katvangers op de samenleving en de veiligheid. Hiervoor wordt data uit het kentekenregister geanalyseerd, waar de focus ligt op de overtredingen en de voertuigkenmerken. Zo wordt er bijvoorbeeld gekeken of de voertuigen van katvangers minder vaak APK gekeurd zijn en daardoor niet veilig genoeg kunnen zijn om op de openbare weg te rijden. Het uiteindelijke doel is om het probleem vroegtijdig aan te pakken door preventieve maatregelen op te stellen, zodat katvangers niet langer fungeren als facilitators van georganiseerde criminaliteit en de samenleving veiliger wordt.

De hoofdvraag en deelvragen zullen worden beantwoord door middel van literatuuronderzoek en door data te analyseren afkomstig uit het kentekenregister van de RDW en data van het Centraal Bureau Statistiek (CBS). In het methode hoofdstuk wordt besproken hoe de data zijn verzameld, hoe de variabelen zijn geoperationaliseerd en wordt er een analyseplan opgesteld. Vervolgens worden de resultaten van de analyses besproken. Hieruit zullen conclusies worden getrokken en zullen er aanbevelingen worden gedaan voor een mogelijk beleid en preventieve maatregelen die kunnen worden getroffen.

Theoretisch kader

Wat zijn katvangers en hoe gaan zij te werk?

Een katvanger is een persoon die wordt gebruikt als dekmantel door een crimineel. Dit kan zich op verschillende manieren uiten binnen de georganiseerde criminaliteit. Een katvanger kan iemand zijn op gebied van drugscriminaliteit, witwassen of identiteitsfraude (Verspille, 2022). Synoniemen die worden gebruikt voor katvangers zijn stromannen (straw men) en geldezels (money mules). Geldezels die worden ingezet voor het witwassen van geld, stellen bankrekening ter beschikking. Geld wordt overgemaakt naar de rekening van de geldezels, waarna de criminelen dit geld opnemen of storten op de rekening van een andere geldezels. Dit zorgt ervoor dat het traceren van dit illegale geld voor instanties bijna onmogelijk wordt. Als gevolg hiervan worden criminelen minder snel betrapt en blijven ze vaak ongestraft (Florencio & Herley, 2010). Verder worden katvangers vaak ingezet in de illegale hennepcultuur. Ze laten huurwoningen op hun naam zetten om als kwekerij te dienen. Wanneer een dergelijk huis wordt ontdekt, worden de katvangers verantwoordelijk gehouden. Ook in dit geval blijven de criminelen ongestraft, omdat ze moeilijk te traceren zijn (Verspille, 2022).

De focus van dit onderzoek ligt op voertuigkatvangers. Zij zetten voertuigen op naam terwijl zij er zelf niet in rijden. Er wordt met data van de RDW gewerkt en daarom wordt dezelfde definitie gehanteerd die de RDW ook gebruikt: *Een katvanger is een natuurlijke (rechts)persoon die - in het algemeen tegen vergoeding - één of meer kentekens op zijn naam heeft staan, terwijl zij geen eigenaar of houder van het desbetreffende voertuig zijn.* Op deze manier wordt de identiteit van de daadwerkelijke houder verhuuld (*Staatsblad 2001, 82 | Overheid.nl > Officiële bekendmakingen*, 2001). Er kan dus niet worden achterhaald wie in het voertuig rijdt en deze persoon kan ook niet verantwoordelijk worden gesteld voor de verplichtingen die verbonden zitten aan de registratie van het voertuig.

Er zijn drie criteria die worden gehanteerd om iemand als katvanger te kenmerken: 1) de natuurlijke persoon of rechtspersoon in kwestie heeft niet de feitelijke beschikking over de op zijn naam geregistreerde voertuigen en zij maken niet aannemelijk dat zij de eigenaar of houder van de op zijn naam geregistreerde voertuigen zijn, 2) zij zijn niet te traceren of bieden onvoldoende verhaal, ten aanzien van de op zijn naam geregistreerde voertuigen en 3) er wordt stelselmatig niet aan één of meer aan de registratie verbonden verplichtingen voldaan (*Staatsblad 2001, 82 | Overheid.nl > Officiële bekendmakingen*, 2001). Er is voornamelijk weinig onderzoek gedaan naar voertuigkatvangers. De gevonden literatuur gaat voornamelijk over de katvangers die worden ingezet voor witwassen of drugscriminaliteit. Onderzoek naar voertuigkatvangers is alleen kleinschalig binnen de RDW gedaan. Er moet rekening worden gehouden met het feit dat de gevonden literatuur niet specifiek over voertuigkatvangers gaat, maar over verschillende soorten

katvangers. Het is relevant om deze onderzoeken te bestuderen, omdat ze inzicht kunnen geven in de eigenschappen van katvangers. Op basis van deze informatie en het huidige onderzoek kan er een duidelijker beeld worden geschetst van een katvanger. Een katvanger wordt gebruikt door criminelen als dekmantel om zelf niet vervolgd te worden. Een katvanger dient als facilitator van georganiseerde criminaliteit. Het huidige onderzoek kan bijdragen aan de schaarse literatuur over voertuigkatvangers en het construeren van maatregelen.

Financiële kwetsbaarheid van katvangers

Financiële moeilijkheden, gebrek aan bewustzijn van de gevolgen en sociaaleconomische achterstanden maken individuen kwetsbaarder voor het betrokken raken bij katvanger praktijken (Wissink, 2021). De hoofdhypothese die wordt opgesteld is dat katvangers (financieel) kwetsbaarder zijn dan niet-katvangers. Hieronder wordt er gekeken welke variabelen kunnen verklaren waarom katvangers over bepaalde kenmerken beschikken.

Er wordt ten eerste gekeken of de katvangers voldoen aan bepaalde wettelijke verplichtingen. Allereerst moet een individu motorrijtuigenbelasting (MRB), ook wel bekend als wegenbelasting, betalen voor het voertuig wat op naam staat. Als katvangers weinig te besteden hebben (Wissink, 2021), wordt verondersteld dat zij daarom niet in staat zullen zijn om de motorrijtuigenbelasting te betalen. Wanneer een persoon geen MRB betaalt, ontvangt deze persoon een boete. In dit geval is dat de katvanger. Wanneer een persoon meer dan vijf keer nalatig is in het betalen van de MRB, wordt deze individu gekenmerkt als een fiscale wanbetaler (Belastingdienst, 2023) Als gevolg hiervan kan deze persoon geen nieuwe voertuigen meer op naam registreren. Er wordt verwacht dat katvangers vaker nalatig zijn in het betalen van de MRB en daardoor vaker fiscale wanbetalers zijn dan niet-katvangers.

Ten tweede wordt gekeken naar het zogenoemde CJIB-signaal. Wanneer een individu vijf of meer boetes op grond van de Wet administratiefrechtelijke handhaving verkeersvoorschriften niet heeft betaald dan kan het Centraal Justitieel Incassobureau (CJIB) een status op een persoon plaatsen in het kentekenregister. Deze persoon kan dan geen nieuwe voertuigen meer op naam zetten (Raad van State, z.d.). De boetes kunnen zijn voor overtredingen op de weg, maar bijvoorbeeld ook voor het niet hebben van een verzekering of het niet betalen van de MRB. Aangezien katvangers vaak financiële moeilijkheden hebben, is het aannemelijk om te veronderstellen dat katvangers over het algemeen vaker een CJIB-signaal hebben, aangezien zij vaak niet in staat zijn om de boetes te betalen. Waar hier rekening mee moet worden gehouden is dat wanneer iemand als katvanger wordt geregistreerd, hij of zij geen nieuwe boetes meer opgelegd zal krijgen. Er wordt verwacht dat katvangers vaker nalatig zijn in het betalen van boetes en van de MRB door hun beperkte financiële middelen. Zij zullen daardoor vaker zowel een CJIB-signaal hebben en vaker fiscale wanbetalers zijn.

Er wordt dus verondersteld dat katvangers niet aan alle verplichtingen kunnen voldoen vanwege financiële moeilijkheden. Het is interessant om te kijken of er bepaalde kenmerken zijn van katvangers die verklaard kunnen worden door deze financiële moeilijkheden. Over het algemeen lijken katvangers te wonen in de grootste gemeenten van Nederland (Arevalo, 2015; Havenaar, 2019). Een van de mogelijke verklaringen hiervoor kan zijn dat er simpelweg meer mensen in de steden wonen en er meer criminaliteit wordt gemeten dan kleinere gemeenten waar ook minder mensen wonen. Rotterdam, Amsterdam en Den Haag hebben het grootste percentage arme inwoners (van Hulst & Hoff, 2019). Katvangers zullen in ruil voor geld eerder akkoord gaan met het op naam zetten van een voertuig, zonder altijd even bewust te zijn van de consequenties (Wissink, 2021). Voor hen kan door deze kwetsbaarheid een klein bedrag veel betekenen en denken hierdoor niet twee keer na voordat ze doorhebben waar ze aan begonnen zijn. Verder lijken katvangers vaker in achterstandswijken of armere wijken te wonen (Arevalo, 2015; Oerlemans et al., 2016). Dit komt overeen met het gegeven dat katvangers vaak minder te besteden hebben. Er zal worden onderzocht of katvangers wonen op postcodes waar een hoog percentage huishoudens valt onder de 40% laagste huishoudensinkomens in Nederland, een gehanteerde maat door het CBS. De hypothese die wordt opgesteld is dat katvangers vaker wonen op postcodes met een gemiddeld laag huishoudensinkomen dan niet-katvangers.

Overigens lijken katvangers vaak gebruik te maken van hetzelfde adres. Een van de voorwaarden voor het op naam zetten van een voertuig is dat de persoon woonachtig moet zijn in Nederland. Door het kopen van een vals huurcontract kunnen de criminelen voertuigen op naam zetten van anderen. Ook worden soms adressen gebruikt van bijvoorbeeld hotels (Van Leiden et al., 2019). Vaak worden voor verschillende katvangers hetzelfde adres gebruikt waardoor er soms een groot aantal mensen op één adres wonen. Dit wordt voornamelijk gedaan door katvangers die bewust zichzelf hebben aangesteld als een katvanger (J. de Graaf, persoonlijke communicatie, 1 februari 2023). Er wordt verwacht dat katvangers op adressen wonen waar veel anderen staan ingeschreven, bijvoorbeeld een opvangcentrum of een bedrijf. Dit hoeven niet per definitie allemaal katvangers te zijn. Toch worden door criminelen ook vaak hetzelfde adres hergebruikt voor meerdere katvangers uit gemakzucht. De hypothese die hieruit wordt opgesteld is dat katvangers op adressen wonen waar gemiddeld meer mensen op staan ingeschreven dan adressen van niet-katvangers.

Verder is het relevant om te kijken naar andere kenmerken van katvangers. Zo wordt er ook gekeken of katvangers een rijbewijs hebben (gehad) of niet. Het is niet verplicht om te beschikken over een rijbewijs om een voertuig op naam te zetten. Ook hier spelen de beperkte financiële middelen die katvangers mogelijk hebben een rol (Wissink, 2021). Rijlessen en een rijbewijs behalen kost veel geld. Er wordt daarom ook hier verwacht dat katvangers minder vaak een rijbewijs hebben omdat zij simpelweg de middelen niet hebben om hiervoor te betalen.

Achtergrondkenmerken

In dit onderzoek wordt er rekening gehouden met externe variabelen, ook wel controlevariabelen genoemd. Ze worden ook onderzocht omdat ze de uitkomsten van het onderzoek kunnen beïnvloeden. Door ze in dit onderzoek op te nemen, kan er worden uitgesloten of zij wel of geen invloed hebben op de uitkomsten. Het gaat om geslacht, leeftijden nationaliteit.

Allereerst lijken mannen vaker katvanger te zijn dan vrouwen (Raza et al., 2020; Wissink, 2021; Kuppens et al., 2006; Arevalo, 2015). Mannen zijn over het algemeen eerder geneigd zijn om een beslissing te nemen gebaseerd op de beloning die ze ervoor krijgen. Ze denken minder goed na over de gevolgen die deze beslissing zal hebben (Arevalo, 2015). Mannen tonen over het algemeen vaker crimineel gedrag dan vrouwen (Leukfeldt & Jansen, 2015).

Over de gemiddelde leeftijd is niet elk onderzoek het met elkaar eens. Het ene onderzoek geeft aan dat de gemiddelde leeftijd van katvangers tussen de 18 en 34 zit (Hülse, 2017), het andere onderzoek laat zien dat het gaat om 30 - 50-jarigen (Versprille, 2016). Wel zijn de onderzoeken het erover eens dat personen die 55+ zijn over het algemeen minder vaak slachtoffer zijn (Wissink, 2021). Over het algemeen lijkt het meer om jonge adolescenten te gaan dan om volwassenen. Met de digitalisering worden er vaak op sociale media-advertenties geplaatst waarin staat dat er snel geld kan worden verdiend. Op deze manier wordt de jongere doelgroep het meeste bereikt en kan dit verklaren waarom zij sneller in aanraking komen met katvanger praktijken (Esoimeme, 2021). Een andere verklaring waarom jongeren sneller katvanger zijn is omdat zij over het vermogen niet hebben om te zien wat hun daden voor consequenties in de toekomst hebben. Ze zullen daarom sneller akkoord gaan om katvanger te worden omdat zij zich niet bewust ervan zijn wat de gevolgen zijn. Jongeren zijn vaak ook goedgeloviger en trappen eerder in een omkooppraatje van criminelen dan volwassen (Arevalo, 2015). Jongeren zijn dus kwetsbaarder dan volwassenen.

Tot slot worden katvangers vaak ingezet door buitenlandse criminelen om naar Nederland te komen om een BSN-nummer aan te vragen en om daarbij ook direct een voertuig op naam te zetten (van Noordenburg, 2020). Ze komen vaak in groepen naar Nederland en ze verlaten het land weer zodra het voertuig op naam staat (J. de Graaf, persoonlijke communicatie, 1 februari 2023). In dit onderzoek wordt er dan ook gekeken of katvangers een Nederlandse nationaliteit hebben of juist een andere nationaliteit.

(Verkeers)veiligheid

In deze scriptie wordt ook onderzocht of katvangers de verkeersveiligheid en rechtszekerheid ondermijnen. Ze kosten de samenleving mogelijk veel geld en dienen als facilitator van georganiseerde criminaliteit. Criminelen schakelen vaak katvangers in wanneer zij een voertuig nodig hebben voor het plegen van een delict. Uit onderzoek bleek dat meer dan de helft van de groep

katvangers binnen één maand meer dan vijf voertuigen op naam had gezet (Nota katvangersproblematiek RDW, 2020). Deze voertuigen zijn veelvuldig betrokken bij verschillende vormen van criminaliteit (RDW, z.d.-a). Verder zullen zij geen boetes ontvangen en is het verleidelijker om roekeloos te gaan rijden, bijvoorbeeld het overschrijden van de maximumsnelheid. De voertuigen die op naam van katvangers staan, begaan 5,5 keer zoveel verkeersovertredingen dan gemiddeld (Nota katvangersproblematiek RDW, 2020). De tweede hoofdhypothese die wordt opgesteld is dat voertuigen op naam van katvangers een negatieve impact hebben op de verkeersveiligheid.

Verkeersveiligheid is een breed begrip. In dit onderzoek is het relevant om te kijken naar de begrippen die de RDW hanteert en hoe deze in verband staan met de verkeersveiligheid. Er wordt gekeken naar de algemene periodieke keuring (APK), WA-verzekeringen, WOK-status en of een voertuig in beslag is genomen. Ook kan er worden gekeken hoeveel boetes een persoon op naam heeft staan.

Allereerst wordt er gekeken of een voertuig een verlopen APK heeft. Bijna alle voertuigen zijn verplicht om te voldoen aan een periodieke autokeuring (APK). Als een voertuig wordt afgekeurd, betekent dit dat er iets mis is en het kan gevaarlijk zijn in het verkeer. Denk bijvoorbeeld aan versleten remmen, waardoor een voertuig mogelijk niet op tijd kan stoppen en in botsing kan komen met een ander voertuig. Voertuigen op naam van katvangers worden vaak niet APK gekeurd (Nota katvangersproblematiek RDW, 2020).

Verder moet iemand met een voertuig op naam verplicht een WA-verzekering afsluiten (RDW, z.d.-b). Dit is verplicht volgens de WAM, de ‘Wet aansprakelijkheidsverzekering Motorrijtuigen’. Met een WA-verzekering is iemand verzekerd wanneer er schade wordt veroorzaakt aan een ander voertuig of persoon. Schade aan eigen voertuig wordt niet gedekt door de WA-verzekering (Verzekering.nl B.V., z.d.). Het verzekeren van een voertuig is verplicht, maar de verzekering moet zelf worden afgesloten. Dit kan betekenen dat niet iedereen ervoor zal kiezen dit ook te doen. Wanneer een persoon geen verzekering heeft afgesloten en er schade wordt gereden zullen er boetes worden uitgegeven, deze zullen dan naar de katvanger gaan. Vaak voldoen voertuigen van katvangers niet aan de voorwaarde van een verplichte WA-verzekering (Nota katvangersproblematiek RDW, 2020). Hier wordt de verwachting gedaan dat de voertuigen van katvangers minder vaak verzekerd zijn.

Wanneer er blijkt dat een voertuig niet aan deze verplichte eisen voldoet, wordt er een boete vanuit het Centraal Justitieel Incassobureau (CJIB) gestuurd naar het adres waarop het voertuig staat geregistreerd (RDW, z.d.-b). In dit geval zal dat dus de katvanger zijn. Uit eerder onderzoek van de RDW blijkt ook dat voertuigen op naam van katvangers minder vaak APK gekeurd zijn (Nota

fenomeen katvangers RDW, 2019). Een verklaring waarom voertuigen op naam van katvangers niet voldoen aan de keurings- en verzekeringsplicht is dat de pakkans klein is voor de criminelen. Criminelen lopen alleen risico om betrapt te worden wanneer zij worden aangehouden. Er kan dan worden gezien dat het voertuig niet voldoet aan de vereisten, wat kan leiden tot verdenking van katvanger praktijken. Dergelijke controles vinden echter alleen plaats wanneer iemand om een andere reden wordt aangehouden; er worden geen willekeurige wegcontroles gedaan om te controleren of voertuigen aan deze voorwaarden voldoen.

Verder wordt er gekeken naar het aantal CJIB-boetes. Het aantal CJIB-boetes laat zien hoeveel boetes er op dit moment openstaan voor een persoon. Het is hier belangrijk om rekening te houden dat wanneer iemand wordt geregistreerd als katvanger, zij geen nieuwe boetes meer opgelegd krijgen (J. de Graaf, persoonlijke communicatie, 1 februari 2023). Er wordt verwacht dat criminelen vaak roekelozer rijden in voertuigen. Dit roekeloze rijgedrag kan leiden tot een toename van het aantal verkeersovertredingen en daarmee een hoger aantal CJIB. Het aantal CJIB-boetes kan als een indicatie van de mate van betrokkenheid van een persoon als katvanger worden gebruikt, waarbij een hoog aantal boetes mogelijk wijst op meer risicovol gedrag in voertuigen die op naam van katvangers zijn geregistreerd.

Vervolgens wordt er ook gekeken naar indicatoren die te maken hebben met de conditie van het voertuig. Allereerst wordt gekeken naar de WOK-status. Dit staat nu ook bekend als ‘verbod voor rijden op de weg’. Auto’s die een WOK-status krijgen toegeschreven mogen niet meer de weg op. Voertuigen met een WOK-status voldoen niet aan de wettelijke eisen. Vaak hebben ze schade of schade gehad, wat ervoor zorgt dat zij niet meer veilig zijn om op de weg te laten (Sluis, 2021). Ook kan een WOK-status betekenen dat het om een opgevoerd voertuig gaat. Dit geldt alleen voor scooters, brommers en motoren. Zowel de RDW als de politie kan deze status opleggen. Zodra een voertuig een WOK-status krijgt mag dit voertuig de openbare weg niet meer op (RDW, z.d.-d). Een WOK-status kan dus laten zien of katvangers vaker voertuigen op naam hebben staan die schade hebben. Dit kan een indicatie zijn dat er meer ongelukken gebeuren met voertuigen op naam van katvangers en of er dus roekelozer wordt gereden. De verwachting is dan ook dat de voertuigen van katvangers vaker dan gemiddeld deze status zullen hebben. Deze verwachting is gebaseerd op het idee dat criminelen roekelozer rijden in deze voertuigen en ze vaker betrekken bij criminele activiteiten, waardoor de voertuigen sneller beschadigd raken.

Ook kan er worden gekeken of een voertuig in beslag is genomen. De politie mag dit doen wanneer een voertuig kan worden gebruikt als bewijsmateriaal of als er duidelijk wordt dat het voertuig wordt gebruikt voor strafbare feiten. Ook kan er beslag worden gelegd op een voertuig wanneer er een snelheidsovertreding is begaan waar de bestuurder een gevaar vormde voor andere

weggebruikers. Het voertuig wordt dan in beslag genomen om te garanderen dat eventuele boetes of ontnemingsmaatregelen in de toekomst kunnen worden betaald (Ingevorderd Rijbewijs, 2023). De persoon waarop het voertuig staat ingeschreven wordt in het kentekenregister van de RDW geregistreerd en kan geen nieuwe voertuigen op naam zetten. Dit wordt een tenaamstellingsblokkade genoemd (Ministerie van Justitie en Veiligheid, 2020). Er wordt verwacht dat voertuigen van katvangers vaker in beslag zijn genomen aangezien er ook wordt verwacht dat zij vaker verkeersovertredingen maken en vaker betrokken zijn bij criminele activiteiten (RDW, z.d.-b; Nota katvangersproblematiek RDW, 2020).

Dit onderzoek

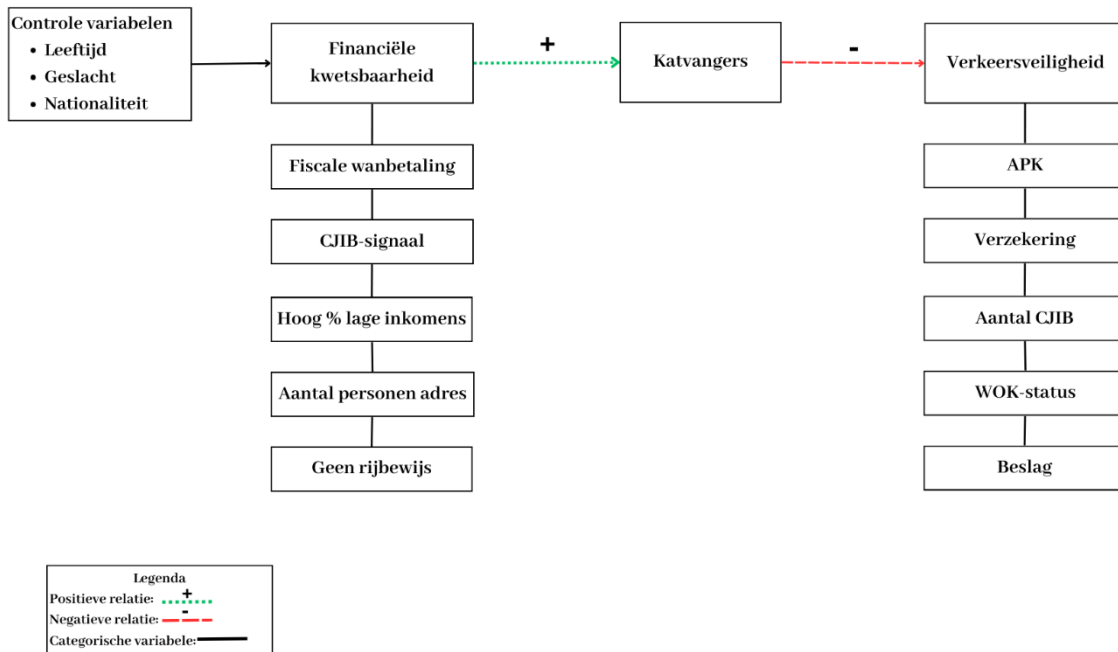
In dit onderzoek zal er worden gekeken naar de voertuigen die katvangers op naam hebben staan en naar voertuigen die op naam staan personen uit een willekeurige steekproef. Vervolgens wordt een vergelijking gemaakt tussen de kenmerken van katvangers en niet-katvangers, evenals tussen de voertuigen op naam van katvangers en niet-katvangers.

Er zijn verschillende kenmerken die zijn te associëren met voertuigkatvangers. Jongeren zijn vaak minder bewust van de gevolgen van hun daden acties en hebben vaak minder te besteden, waardoor ze eerder geneigd zijn om de beloning van snel geld te accepteren zonder volledig rekening te houden met de gevolgen. Daarnaast lijken katvangers vaak te wonen in grote steden, vaak met hogere armoedeniveaus, wat ook lijkt te wijzen op een verband tussen financiële kwetsbaarheid en betrokkenheid bij deze praktijken. De overkoepelende factor van kwetsbaarheid lijkt diverse kenmerken van voertuigkatvangers met elkaar te verbinden. *Figuur 1* toont de eerste hoofdhypothese; er wordt verwacht dat er een positieve relatie is tussen financiële kwetsbaarheid en katvangers. Dit betekent dat er wordt verwacht dat katvangers vaker financieel kwetsbaar zijn, in tegenstelling tot niet-katvangers.

Hiervoor wordt er gekeken naar de achtergrondkenmerken, dus de kenmerken die gebonden zijn aan de persoon. Er wordt gekeken of de persoon een fiscale wanbetaler is, of zij een CJIB-sigitaal hebben, of zij een rijbewijs hebben (gehad), hoeveel personen op hun woonadres staan ingeschreven en of zij behoren tot huishoudens met een laag gemiddeld inkomen. Er wordt hierbij gecontroleerd voor leeftijd, geslacht en nationaliteit.

Vervolgens wordt er gekeken naar de variabelen die gekoppeld zijn aan de voertuigen van individuen. Er wordt een vergelijking gemaakt tussen de voertuigen van katvangers en niet-katvanger. Daarbij wordt er gekeken of de voertuigen van katvangers daadwerkelijk minder vaak verzekerd zijn, meer boetes open hebben staan, een verlopen APK hebben, een WOK-status hebben en of ze in beslag zijn genomen. Dit is relevant om te analyseren, omdat het kan aantonen of katvangers meer schade rijden met hun voertuigen, waardoor de assumptie zou kunnen worden

gedaan dat zij roekelozer rijden en dus de verkeersveiligheid beïnvloeden. Het conceptuele model laat ook de tweede hoofdhypothese die opgesteld is zien. Er wordt verwacht dat er een negatieve relatie is tussen katvangers en verkeersveiligheid. Dit houdt in dat er wordt verwacht dat katvangers een negatieve invloed hebben op de verkeersveiligheid, in vergelijking met niet-katvangers.



Figuur 1: Conceptueel model

Methode

Steekproef en design

De data die wordt gebruikt in het onderzoek is verzameld uit het kentekenregister van de RDW. De datasets zijn gecreeërd in februari 2023 en zijn een momentopname van de situatie in die maand. Dit betekent dat eventuele veranderingen niet zijn opgenomen in de datasets.

Ten eerste zijn er twee datasets gecreeërd die informatie bevatten over de achtergrondkenmerken. De eerste dataset bestaat uit geïdentificeerde katvangers. Vervolgens is er een tweede dataset gemaakt met dezelfde variabelen, maar de individuen zijn nu willekeurig geselecteerd uit de Nederlandse populatie. Hierdoor kunnen de katvangers worden vergeleken met de gemiddelde samenleving. Uiteindelijk worden deze twee datasets samengevoegd om de statistische analyse uit te voeren.

Verder zijn er twee extra datasets gemaakt die specifieke informatie bevatten over de voertuigen van deze personen. Aangezien individuen vaak meerdere voertuigen op naam hebben staan, wordt door het apart behandelen van de dataset het makkelijker om de specifieke informatie over de voertuigen te onderzoeken.

Tot slot wordt er gebruikgemaakt van data van het CBS. De data laten zien welke hoeveel procent van de huishoudens in welke postcodes in de categorie 40% laagste gestandaardiseerde huishoudensinkomen vallen. Door de waarden van de postcodes toe te voegen aan de dataset met achtergrondkenmerken kan er worden onderzocht of katvangers vaker in armere buurten wonen. De meest complete cijfers van het CBS komen uit 2020 en worden in dit geval ook gebruikt.

Personen en voertuigen

De eerste dataset bestaat uit 9981 katvangers. De postcode van de katvangers gaat over hun geregistreerde woonplaats, en bestaat uit vier cijfers en beperkt is tot Nederlandse postcodes. Deze wordt uiteindelijk gebruikt voor het combineren met de data uit het CBS. De postcodes waar geen data voor beschikbaar was kregen system-missing. Vervolgens is er voor de postcodes bekend hoeveel procent van de huishoudens binnen die postcode onder het gemiddelde inkomen vallen. Voor 26 katvangers misten postcodes en zij zijn uit de dataset verwijderd. Verder zijn de missende waarden voor de variabelen katvanger, CJIB-signaal, fiscale wanbetaling, rijbewijs, aantal op adres, leeftijd, geslacht en nationaliteit ook verwijderd. Er bleven uiteindelijk 9662 katvangers over. Deze dataset wordt aangevuld met de random steekproef bestaande uit 10000 individuen. Ook hier is de data van het CBS toegevoegd. Na het verwijderen van de ontbrekende waarden in de steekproef, blijven er 5408 personen over. Dit aantal is gehalveerd omdat er in de corresponderende voertuigdataset cases zijn verwijderd omdat dit outliers waren. Deze cases zijn daarom ook verwijderd uit de dataset met niet-katvangers. Uiteindelijk bevat de gecombineerde dataset 15070 individuen. Deze dataset bevat een aantal controlevariabelen zoals leeftijd, geslacht en nationaliteit.

Naast het verwijderen van de missende waarden zijn outliers voor 'leeftijd' en 'aantal op adres' verwijderd, als ze drie standaarddeviaties van het gemiddelde afweken. In totaal zijn er 63 cases verwijderd voor de variabele 'leeftijd' en 146 cases voor 'aantal op adres'. In bijlage 3 wordt weergegeven hoe dit is gedaan.

Daarnaast zijn er twee datasets met betrekking tot voertuigkenmerken. Er zijn in totaal 7302 voertuigen op naam van katvangers. Na het verwijderen van de cases waarvan gegevens voor sommige variabelen ontbreken, blijven er 7152 cases over. Dit betekent dat er voor 2537 katvangers geen voertuig staat geregistreerd. De willekeurige steekproef bestaat uit origineel uit 124.925 cases en na het verwijderen van de cases met ontbrekende scores op bepaalde variabelen blijven er 87.453 cases over. Dit betekent dat er voor de random steekproef geldt dat één individu ongeveer 8,8 voertuigen op naam heeft staan. Dit aantal is erg hoog, waardoor er een vermoeden ontstaat dat er ook bedrijven in de steekproef zitten. Er wordt daarom ervoor gekozen om alle cases die meer dan vijf voertuigen op naam hebben uit de dataset te filteren. Het ging om in totaal om 4592 personen en deze hadden gemiddeld 111,23 voertuigen op naam staan. Een reden hiervoor kan zijn is dat er hoogstwaarschijnlijk ook bedrijven of garages zijn opgenomen in de dataset, die vaak meerdere voertuigen op naam hebben staan.

Na het verwijderen van deze cases blijven er 13.819 cases over. Dit zorgt er nu voor dat een individu uit de random steekproef gemiddeld 2,93 voertuigen op naam heeft. De verwijderde cases uit de willekeurige steekproef dataset voor voertuigkenmerken worden ook verwijderd uit de dataset met achtergrondkenmerken van niet-katvangers, in bijlage 3 wordt weergegeven hoe dit is gedaan. Uiteindelijk worden deze twee datasets ook samengevoegd. De uiteindelijke dataset heeft 20.971 cases.

Variabelen

Datasets – achtergrondkenmerken

In totaal bestaat de dataset met achtergrondkenmerken uit 9 variabelen waarvan de operationalisaties eerst kort zullen worden besproken.

Katvanger

De afhankelijke variabele is katvanger (1) versus niet-katvanger (0).

Fiscale wanbetaling

Ook wordt de variabele 'fiscale wanbetaling' gebruikt als voorspeller. Fiscale wanbetaling is een nominale variabele en wordt gecodeerd door middel van 0 = geen fiscale wanbetaler en 1 = een fiscale wanbetaler.

CJIB-signaal

CJIB-signaal laat zien wanneer een individu vijf of meer boetes op grond van de Wet administratiefrechtelijke handhaving verkeersvoorschriften niet heeft betaald. CJIB-signaal is een nominale variabele en bestaat uit twee categorieën; 0 = geen CJIB-signaal, 1 = wel CJIB-signaal.

Rijbewijs

Rijbewijs meet of een persoon momenteel een rijbewijs heeft of heeft gehad. Het is een nominale variabele en wordt geoperationaliseerd door middel van 0 = nooit rijbewijs gehad, 1 = heeft rijbewijs (gehad).

Percentage laag-inkomen huishoudens

De variabele ‘percentage laag-inkomen huishoudens’ geeft het percentage huishoudens aan die binnen deze postcode tot de 40% laagste inkomenshuishoudens behoren. Het is een continue variabele.

Aantal op adres

De variabele ‘aantal op adres’ toont aan hoeveel personen op één adres staan ingeschreven. ‘Aantal op adres’ is een continue variabele waardoor ervoor is gekozen om deze te centreren door het gemiddelde van 8,07 af te trekken.

Leeftijd

De eerste controlevariabele is de continue variabele leeftijd in jaren, door het geboortjaar van het jaar 2022 af te trekken. Zo blijft alleen de leeftijd over. Leeftijd is een continue variabele en gecentreerd door van de variabele ‘leeftijd’ de gemiddelde leeftijd van 48,59 af te trekken.

Geslacht: 0 = man en 1 = vrouw.

Nationaliteit

Er wordt in het onderzoek alleen gekeken naar binnenlandse katvangers wat inhoudt dat zij actief moeten zijn in Nederland. Er wordt hierbij ook gekeken of katvangers over het algemeen een Nederlandse nationaliteit hebben of een buitenlandse nationaliteit. De variabele ‘nationaliteit’ is een nominale variabele en bestaat uit de categorieën 0 = niet Nederlands en 1 = Nederlands.

Datasets – voertuigen

Vervolgens wordt er gekeken naar de datasets met voertuigkenmerken. Deze bestaat uit 8 variabelen.

WOK-status

Ten eerste wordt gekeken naar de WOK-status/verbod voor rijden op de weg. WOK staat voor wachten op keuring, ook wel binnen de RDW ‘verbod voor rijden op de weg’ genoemd. De RDW werkt met twee soorten WOK-statussen. In de dataset wordt geen onderscheid gemaakt tussen de verschillende soorten WOK-status. De codering van de nominale variabele WOK-status is 0= geen status en 1 = wel status.

APK-verlopen (1) en niet verlopen (0)

Aantal CJIB-boetes

Verder wordt er gekeken naar hoeveel CJIB-boetes er aan een voertuig zijn gekoppeld. Deze variabele verschilt van de variabele ‘CJIB-sigitaal’ die in de dataset met achtergrondkenmerken wordt gebruikt. De variabele ‘aantal CJIB’ geeft het huidige aantal CJIB-boetes die openstaan weer. Dit is een discrete variabele waarvan de minimale waarde 0 is en waar verder elke numerieke waarde kan worden ingevuld.

Verzekering: 0 = wel verzekerd, 1 = niet verzekerd.

Beslag: 0 = niet in beslag genomen en 1 = wel in beslag genomen.

Analyseplan

Door middel van een logistische regressie wordt er in de dataset met achtergrondkenmerken gekeken naar het verschil tussen de katvangers en niet-katvangers. De afhankelijke variabele is katvanger. In het eerste model worden de controlevariabelen leeftijd, geslacht en nationaliteit toegevoegd. In het tweede model worden de andere voorspellers toegevoegd: percentage laag inkomen huishoudens, aantal personen op adres, rijbewijs, CJIB-sigitaal en fiscale wanbetaling.

In een tweede analyse wordt er ook gekeken naar de voertuigkenmerken. Er wordt een vergelijking gemaakt tussen de voertuigen van katvangers en niet-katvangers door te kijken naar de verschillen in de descriptieve statistieken. Ook worden de gegevens geaggregeerd om te kijken of er sprake is van een verschil in bijvoorbeeld de hoeveelheid voertuigen dat een persoon op naam heeft staan, of katvangers nieuwere voertuigen op naam hebben staan en of zij gemiddeld meer boetes open hebben staan. De verschillen zullen worden geanalyseerd door middel van een logistische regressie.

Resultaten

Descriptieve statistieken persoonsniveau

Tabel 1 laat de descriptieve statistieken van de achtergrondkenmerken zien. Van de katvangers is 0,02% fiscalewanbetaler. Hieruit blijkt dat individuen uit de willekeurige dataset vaker fiscale wanbetaler lijken (0,4%). De Chi-kwadraattoets heeft een p -waarde van $<,001$ dus ook hier wordt de nulhypothese verworpen ($\text{Chi}^2(1) = 27,24$). Er is een significant verband tussen fiscale wanbetaler en katvangers.

Ten tweede wordt er gekeken of een katvanger een CJIB-signaal heeft of niet: 2,9% heeft een CJIB-signaal en van de katvangers is dit 0,5% ($\text{Chi}^2(1) = 101,54$; $p <,001$). Er is sprake van een significant verband.

De meerderheid van de katvangers 63,5% heeft geen rijbewijs. Bijna iedereen in de willekeurige steekproef heeft een rijbewijs (gehad, 99,9%). De Chi-kwadraattoets heeft een p -waarde van $<,001$ dus ook hier wordt de nulhypothese verworpen ($\text{Chi}^2(1) = 5761,24$). Er is een significant verband tussen rijbewijs en katvanger.

De variabele 'percentage laag-inkomen huishoudens' heeft een gemiddelde van 44,48 (SD = 11,63) in de dataset van de katvangers. Het minimum is 2 en het maximum is 76. In de dataset van de willekeurige individuen is iets lager, namelijk 36,97 (SD = 10,81). De t -toets heeft een p -waarde van $<,001$ ($t(15068) = 39,02$). Dit betekent dat de nulhypothese wordt verworpen en dat er een significant verschil is tussen het percentage lage inkomens voor niet-katvangers en katvangers.

Vervolgens wordt er gekeken hoeveel personen op een adres wonen. Het gemiddelde van de katvangers is hoog, namelijk 11,11. Er is sprake van veel spreiding (SD=38,36) op de schaal van 1 – 427. Er is voor deze waarde gecorrigeerd door alle waarden boven de 1000 eruit te halen, maar de hoge waarden zijn niet allemaal eruit gehaald omdat katvangers vaak valse adressen gebruiken zoals een hotel. Het meest voorkomende antwoord is 2 personen op één adres. Het gemiddelde aantal mensen wat op één adres woont van de willekeurige steekproef is een stuk lager, namelijk 2,63. De standaarddeviatie is redelijk groot met 5,88 wat betekent dat er redelijk wat spreiding is. De p -waarde is $<,001$ ($t(15068) = -16,16$). Dit betekent dat de nulhypothese wordt verworpen.

De eerste controlevariabele is leeftijd. De gemiddelde leeftijd van een katvanger is 47,32 (range: 18 - 99). In de willekeurige steekproef ligt de gemiddelde leeftijd iets hoger, namelijk 50,85 (range: 16 - 89) en dit verschil was significant ($t(15068) = 13,74$; $p <,001$).

Er zijn meer mannen dan vrouwen katvanger: 74,4% is man en 25,6% is vrouw. In de willekeurige steekproef is bijna driekwart man (70,7%). Ook hier is de $p <,001$ van de Chi-kwadraattoets, wat betekent dat de nulhypothese wordt verworpen en dat er dus een significant verband is tussen geslacht en katvanger ($\text{Chi}^2(1) = 23,96$). De gemiddeldes van de katvangers en niet-katvangers lijken ook erg overeen te komen.

De meerderheid van de katvangers is Nederlands, namelijk 87,7%. In de willekeurige dataset heeft bijna iedereen de Nederlandse nationaliteit, namelijk 99,7%. De Chi-kwadraattoets heeft een p -

waarde van $<,001$ dus ook hier wordt de nulhypothese verworpen ($\text{Chi}^2(1) = 679,83$). Er is een significant verband tussen nationaliteit en katvangers.

Dit betekent dat katvangers verschillen van niet-katvangers doordat ze specifieke kenmerken vertonen die statistisch significant afwijken. Fiscale wanbeteling, CJIB-sigitaal, het hebben van een rijbewijs, het percentage laag-inkomen van huishoudens, het aantal personen op één adres, de leeftijd, geslacht en nationaliteit van katvangers laten allemaal significante verschillen zien in vergelijking met niet-katvangers.

Tabel 1: descriptieve statistieken van de variabelen uit datasets met achtergrondkenmerken

Variabele	Katvangers						Niet-katvangers						T-test	Df	Chi-square	Df
	Mean	SD	Min	Max	Med.	Mode	Mean	SD	Min	Max	Med.	Mode				
<i>Fiscale Wanbetaling</i>	0=99,98%					0	0=99,6%					0	-	-	27,24**	1
<i>0 = niet</i>	1=,02%						1=,4%									
<i>1 = wel</i>																
<i>CJIB-signaal</i>	0=97,1%	-	-	-	-	1	0=99,5%	-	-	-	-	0	-	-	101,54**	1
<i>0 = niet</i>	1=2,9%						1=,5%									
<i>1 = wel</i>																
<i>Rijbewijs</i>	0= 63,5%	-	-	-	-	1	0=,1%	-	-	-	-	0	-	-	5761,24**	1
<i>0=nooit</i>	1=36,5%						1=99,9%									
<i>1=heeft rijbewijs (gehad)</i>																
<i>Percentage huishoudens</i>	44,48	11,63	2	76	45	46	36,97	10,81	4	74	35	29	39,02**	15068	-	-
<i>Aantal op adres</i>	11,11	38,37	1	427	2	1	2,63	5,88	1	168	2	2	-16,16**	15068	-	-
<i>Leeftijd</i>	47,32	14,54	18	99	47	53	50,85	16,12	16	89	51	52	13,74**	15068	-	-
<i>Geslacht</i>	0= 74,4%	-	-	-	-	0	0=70,7%	-	-	-	-	0	-	-	23,96**	1
<i>0=man</i>	1=25,6%						1=29,3%:									
<i>1=vrouw</i>																
<i>Nationaliteit</i>	0=12,3%	-	-	-	-	0	0=0,3%	-	-	-	-	1	-	-	679,83**	1
<i>0 = niet NL</i>	1= 87,7%						1=99,7%									
<i>1 = NL</i>																

*significant op 0,05; **significant op 0,01

Modelfit

Er wordt gekeken naar de logistische regressie en welk model de beste voorspeller is. In model 2 worden de demografische variabelen toegevoegd. *Tabel 3* laat zien dat model 2 beter voorspelt dan het lege model ($\text{Chi}^2 = 1272,43; p <,001$). Het toevoegen van de variabelen zorgt er dus voor een betere voorspelling. In model 3 worden de andere predictoren toegevoegd. De Chi-waarde is hoger dan model 2 ($\text{Chi}^2 = 10205,11; p <,001$). Een lagere Chi-waarde betekent dat er sprake is van goede pasvorm tussen het voorspelde model en de waargenomen data. Model 2 heeft in dit geval een betere fit.

Verder wordt er gekeken naar de Cox & Snell R^2 . Deze toets meet hoe goed het model de variantie in de afhankelijke variabele kan verklaren. De waarde ,08 betekent dat het model ongeveer 8% van de variantie kan voorspellen. De Cox & Snell R^2 is een stuk hoger geworden na het toevoegen van de voorspellers, namelijk ,53. Dit betekent dat model 3 dus ongeveer 53% van de variantie in katvangerschap voorspellen.

Ten tweede wordt er gekeken naar de Nagelkerke R^2 . Dit is een verbeterde versie van de Cox & Snell R^2 . Des te dichter deze waarde bij de 1 ligt, des te beter de fit van het model is. De fit van model 2 is niet heel hoog. De onafhankelijke variabelen kunnen de variabiliteit van de afhankelijke variabele met 11,1% verklaren. De Nagelkerke R^2 in model 3 stijgt naar ,73, wat betekent dat de onafhankelijke variabelen de variabiliteit van de afhankelijke variabele met 73% kunnen verklaren. De rest wordt verklaard door externe variabelen, dus variabelen die niet in het model zijn opgenomen. Ook hier is model 3 het beste.

De McFadden R^2 is ook een pseudo R^2 -maat. De McFadden R^2 in model 2 is ,08. Het model verklaart ongeveer 8% van de verschillen in de uitkomstvariabele. Het is belangrijk om te onthouden dat de pseudo R^2 -maten vaak lagere waarden aangeven, maar gebaseerd op deze resultaten kan er worden geconcludeerd dat het model met alle variabelen een betere fit heeft dan het model zonder variabelen. In model 3 stijgt de McFadden R^2 naar ,63. Model 3 kan dus ongeveer 63% van de verschillen in uitkomstvariabele verklaren.

Tot slot kan er worden gekeken naar de Hosmer & Lemeshow test. De cases worden verdeeld in groepen van 10 en er wordt gekeken naar hoeveel individuen 0 of 1 hebben gescoord en hoeveel er op basis van het model worden verwacht. De verschillen zijn in dit geval niet significant ($p <,001$). Dit zou betekenen dat het model niet goed bij de data past. De test geeft bij grote steekproeven echter snel aan dat er sprake is van significant verschillen. Aangezien hier sprake is van een grote steekproef kan er worden geconcludeerd dat de Hosmer & Lemeshow test geen goede voorspeller is van de fit. Ook in model 3 is de Hosmer & Lemeshow significant met $p <,001$. De N blijft hoog en daarom is deze test nog steeds geen goede voorspeller van de fit.

Tabel 3: parameters van de logistische regressie met achtergrondkenmerken en met katvanger als onafhankelijke variabele

Variabele	Model 1			Model 2			Model 3		
	<i>b</i> (SD)	Sig.	Exp(B)	<i>b</i> (SD)	Sig.	Exp(B)	<i>b</i> (SD)	Sig.	Exp(B)
<i>Constante</i>	,58(,02)	<,001	1,78	4,46(,25)	<,001	86,37	9,71(,47)	<,001	16414,5
<i>Leeftijd (gecentreerd)</i>	-	-	-	-,02(,00)	<,001	,98	-,03(,00)	<,001	,97
<i>Geslacht</i>	-	-	-	-,17(,04)	<,001	,85	,22(,06)	<,001	1,24
<i>Nationaliteit</i>	-	-	-	-3,98(,25)	<,001	,02	-5,24(,26)	<,001	,01
<i>Fiscale Wanbetaling</i>	-	-	-	-	-	-	-4,13(1,42)	,004	,02
<i>CJIB-sigtaal</i>	-	-	-	-	-	-	1,72(,26)	<,001	3,67
<i>Rijbewijs</i>	-	-	-	-	-	-	-7,78(,38)	<,001	,00
<i>Percentage huishoudens</i>	-	-	-	-	-	-	,06(,00)	<,001	1,07
<i>Aantal op adres (gecentreerd)</i>	-	-	-	-	-	-	,06(,01)	<,001	1,06
<i>Cox and Snell</i>	-	-	-	,08	-	-	,53	-	-
<i>Nagelkerke</i>	-	-	-	,11	-	-	,73	-	-
<i>McFadden</i>	-	-	-	,08	-	-	,63	-	-
<i>Chi-square</i>	-	-	-	1272,43	<,001	-	10205,11	<,001	-
<i>Hosmer & Lemeshow</i>	-	-	-	64,46	<,001	-	150,59	<,001	-
<i>N</i>	15070								

Assumpties en outliers

Een logistische regressie moet aan een aantal assumpties voldoen. Het is allereerst van belang dat de afhankelijke variabele binair is. De variabele 'katvanger' is in deze analyse de afhankelijke variabele en is binair. Deze eerste assumptie wordt dan ook niet geschonden.

Verder is het van belang dat de waarnemingen onafhankelijk moeten zijn van elkaar. Er is geen indicatie dat de waarnemingen gerelateerd zijn aan elkaar, dus deze assumptie wordt niet geschonden.

Ook moeten de onafhankelijke variabelen niet te sterk met elkaar gecorreleerd zijn, er mag geen sprake zijn van multicollineariteit. De VIF-waardes moeten onder de 10 liggen. Uit de analyse lijken voor alle variabelen de VIF-waarde rond de 1 te liggen. Deze assumptie wordt niet geschonden.

Tot slot wordt er gekeken naar de lineariteit tussen de continue en de log-odds. In een logistische regressie wordt verondersteld dat de uitkomst een lineaire combinatie is van de onafhankelijke variabelen. Dit zou dus betekenen dat er geen lineaire relaties moeten zijn tussen de voorspellers zelf. Zowel voor de variabele 'leeftijd' als 'percentage huishoudens' is er sprake van lineariteit tussen de log-odds (zie bijlage 3). Dit betekent dat deze assumptie wordt geschonden. Voor de variabele 'aantal op adres' is er geen sprake van lineariteit en wordt voor deze variabele de assumptie niet geschonden.

Er wordt gecontroleerd voor outliers door te kijken naar de Cook's Distance, gestandaardiseerde residuen en Mahalanobis Distance. Allereerst is er al gefilterd op een aantal variabelen. Voor zowel niet-katvangers als katvangers was de maximumleeftijd hoog.

Gebaseerd op de Cook's Distance, gestandaardiseerde residuen en Mahalanobis Distance valt op dat er ook outliers zijn (Bijlage 3). Wanneer er wordt gekeken naar de cases valt op dat er echter geen extreme waarden zijn. Gebaseerd op deze bevindingen is ervoor gekozen om de cases die hoog scoren op deze toetsen niet te verwijderen omdat zij niet opvallend scoren op de individuele variabelen.

Analyse – Logistische regressie achtergrondkenmerken

Allereerst wordt gekeken naar de variabele 'fiscale wanbetaler'. Er is sprake van een hoge negatieve helling ($b = -4,13$; $SE = 1,42$; $p = ,004$). De kans op het zijn van een katvanger neemt af wanneer iemand een fiscale wanbetaler is. De odds-ratio is laag, namelijk ,02. De kans op het zijn van een katvanger wanneer iemand een fiscale wanbetaler is, is 2% van de kans dat iemand een katvanger is wanneer deze niet wordt gezien als fiscale wanbetaler.

De helling van 'CJIB-signaal' is positief en redelijk groot ($b = 1,72$; $SE = ,26$; $p < ,001$). De kans dat iemand een katvanger is neemt met 1,72 toe wanneer iemand een CJIB-signaal heeft. De odds-ratio is hoog, namelijk 3,67. De kans dat een individu katvanger is, is 3,67 groter wanneer iemand een CJIB-signaal heeft dan een individu zonder CJIB-signaal.

Opvallend zijn de helling en odds-ratio van de variabele 'rijbewijs'. De helling is negatief en hoog ($b = -7,78$; $SE = ,38$; $p < ,001$). Wanneer een individu een rijbewijs heeft (gehad) worden de log-

odds op het zijn van een katvanger 7,78 lager, gecontroleerd voor alle andere variabelen. De odds-ratio is laag, namelijk ,00. Dit betekent dat het hebben van een rijbewijs zorgt voor een lagere kans op het zijn van een katvanger.

Vervolgens wordt er gekeken naar het percentage huishoudens die behoren tot de armere huishoudens. De helling is positief en redelijk laag ($b = ,06$; $SE = ,00$; $p < ,001$). Des te hoger het percentage is, des te hoger de kans is dat iemand een katvanger is. De odds-ratio is 1,07. Dit betekent dat de kans op katvanger 1,07 groter wordt wanneer het gebied met armere huishoudens met 1 procentpunt toeneemt.

Verder kan er worden gekeken naar aantal op adres. De helling is positief en laag ($b = ,06$; $SE = ,01$; $p < ,001$). Zodra er meer mensen op één adres wonen zullen de log-odds op katvanger toenemen met ,06. Des te meer mensen er dus op één adres wonen, des te groter de kans is dat het gaat om een katvanger.

Daarnaast wordt er gekeken naar de controle variabelen. Leeftijd heeft een negatieve en kleine helling ($b = -,03$ $SE = ,002$; $p < ,001$). Des te ouder iemand is, des te kleiner de kans wordt dat deze persoon een katvanger is. De odds-ratio ligt onder de 1 ($\text{Exp}(B) = ,97$), wat betekent dat wanneer iemand 1 jaar ouder wordt, de kans op katvanger met ,97 afneemt.

De helling voor geslacht is positief en redelijk hoog ($b = ,22$; $SE = ,06$; $p < ,001$). Wanneer een individu vrouw is, zullen de log-odds op katvanger zijn met ,22 toenemen, gegeven dat er op de rest 0 wordt gescoord. De $\text{Exp}(B)$, ofwel odds-ratio, is 1,24.

De helling in model 2 voor geslacht negatief en significant is, terwijl deze in model 3 positief is geworden. Een reden hiervoor kan zijn dat er in het derde model een variabele wordt toegevoegd die gerelateerd is aan de afhankelijke en de onafhankelijke variabele, in dit geval leeftijd. Dit wordt een confounding variabele genoemd en kan dus de richting van de helling beïnvloeden.

Bijlage 2 laat daarom zien dat er voor de variabele geslacht stapsgewijs een logistische regressie is gedaan omdat er een vermoeden is dat er sprake is van een schijnverband. De analyses laten zien dat zodra de variabele rijbewijs wordt toegevoegd, de helling van negatief naar positief verandert.

Wanneer er wordt gekeken naar nationaliteit valt op dat de helling negatief en groot is ($b = -5,24$ $SE = ,26$; $p < ,001$). De kans op het zijn van een katvanger voor iemand met een Nederlands nationaliteit is ongeveer 1% van de kans van individuen met een niet-Nederlandse nationaliteit ($\text{Exp}(B) = ,01$). Individuen met een niet-Nederlandse afkomst hebben dus een hogere kans om katvanger te zijn.

Descriptieve statistieken voertuigkenmerken

Tabel 4 laat de descriptieve statistieken van de voertuigkenmerken zien. Allereerst wordt er gekeken naar hoe vaak een ID gemiddeld voorkomt. Het gemiddelde van katvangers is 23,24 (SD = 35,59). De schaal loopt van 1 – 154. In de dataset met niet-katvangers valt op dat dit gemiddelde een stuk lager ligt, namelijk 2,93 (SD = ,94). Het maximum voertuigen op naam van de random steekproef is 4.

Uit de resultaten blijkt dat minder individuen uit de willekeurige steekproef een WOK-status hebben (2,3%) dan de katvangers (3,6%). De p -waarde van de Chi-kwadraattoets is $<,001$, wat betekent dat de nulhypothese wordt verworpen ($\text{Chi}^2(1) = 28,96$). Het aantal katvangers is hetzelfde voor een voertuig met WOK-status, als een voertuig zonder WOK-status.

Voertuigen van katvangers hebben vaak een verlopen APK (59,1%), terwijl dit voor geen van de niet-katvangers geldt. Ook hier is de p -waarde van de Chi-kwadraattoets $<,001$, wat betekent dat er een significant verband is tussen katvanger en APK ($\text{Chi}^2(1) = 10220,1$).

Het hoogste aantal CJIB-boetes op naam van een katvanger is 67. Het laagste aantal, 0, komt het meest voor. Het gemiddelde van de katvangers is ,21 (SD = 1,50) Het gemiddelde in de willekeurige steekproef ligt iets lager, namelijk ,16 (SD = ,582). Ook hier is de p -waarde van de t -toets $<,001$ wat betekent dat er een significant verschil is tussen katvanger en het aantal CJIB-boetes ($t(20969) = -3,24$).

In totaal is 73,7% van de voertuigen van katvangers niet verzekerd. De meerderheid van de voertuigen op naam van katvangers blijken dus niet verzekerd te zijn. In de dataset bestaande uit willekeurige individuen heeft iets meer dan de helft hun voertuig wel verzekerd, namelijk 67,9%. Dit verschil is significant ($\text{Chi}^2(1) = 3276,71$; $p <,001$)

Tot slot is 7,4% van de voertuigen van katvangers in beslag genomen. Slechts ,6% van de willekeurige steekproef heeft beslag gelegd gekregen op hun voertuig. De p -waarde van de Chi-kwadraattoets is $<,001$ ($\text{Chi}^2(1) = 887,74$). Er is een significant verband tussen katvanger en verzekering.

Tabel 4: descriptieve statistieken van de variabelen uit datasets met voertuigkenmerken

Variabele	Katvangers						Niet-katvangers						T-test	Df	Chi-square	Df
	Mean	SD	Min	Max	Med	Mode	Mean	SD	Min	Max	Med	Mode				
<i>ID geaggregeerd</i>	23,24	35,59	1	154	5	1	2,93	,94	1	4	3	3	-	-	-	-
<i>WOK-status</i> 0 = geen status 1 = wel status	0=96,4% 1=3,6%	-	-	-	-	0	0=97,7% 1=2,3%	-	-	-	-	0	-	-	28,96**	1
<i>Apk verlopen</i> 0 = niet verlopen 1 = verlopen	0= 40,9% 1=59,1%	-	-	-	-	0	0=100% 1=0%:	-	-	-	-	0	-	-	10220,1**	1
<i>Aantal CJIB</i>	,21	1,50	0	67	0	0	,16	,58	0	20	0	0	-3,24**	20969	-	-
<i>Verzekerd</i> 0 = verzekerd 1 = niet verzekerd	0=26,3% 1=73,7%	-	-	-	-	1	0=67,9% 1=32,1%	-	-	-	-	0	-	-	3276,71**	1
<i>Beslag</i> 0 = niet in beslag 1 = in beslag	0=92,6% 1=7,4%	-	-	-	-	0	0=99,7% 1=,3%	-	-	-	-	0	-	-	887,74**	1

*significant op 0,05; **significant op 0,01

Modelfit

Ook voor de samengevoegde dataset met voertuigkenmerken wordt er een logistische regressie gedaan.

In model 2 worden de predictoren toegevoegd. *Tabel 5* laat zien dat model 2 beter voorspelt dan het lege model ($\text{Chi}^2 = 12327,87; p <,001$). Het toevoegen van de variabelen zorgt er dus voor een betere voorspelling. Dit betekent dat dit model beter is dan het model zonder de variabelen. Het toevoegen van de variabelen zorgt er dus voor dat er een betere voorspelling kan worden gedaan.

Voor de modelfit wordt er ook worden gekeken naar de Cox & Snell R^2 en de Nagelkerke R^2 . De Cox & Snell R^2 in model 2 is ,44, het model kan dus ongeveer 44% van de variantie in de afhankelijke variabele voorspellen. De Nagelkerke R^2 is ,62 en ligt redelijk dicht bij de 1. De fit van dit model is dus redelijk goed. De onafhankelijke variabelen kunnen de variabiliteit van de afhankelijke variabele met 62% verklaren. De rest wordt verklaard door externe variabelen, dus variabelen die niet in het model zijn opgenomen.

Verder is er ook nog de McFadden R^2 . Dit is ook een pseudo R^2 -maat. De McFadden R^2 in model 2 is ,46. Het model verklaart ongeveer 46% van de verschillen in de uitkomstvariabele. Gebaseerd op deze resultaten kan ook hier worden geconcludeerd dat het model met alle variabelen een betere fit heeft dan het model zonder variabelen.

Tot slot kan wordt er gekeken naar de Hosmer & Lemeshow test. De verschillen tussen de groepen zijn in dit geval significant ($p = ,056$). Dit zou betekenen dat het model goed bij de data past.

Assumpties en outliers

Wederom moet er ook gecontroleerd worden voor de assumpties. Ook hier is de variabele 'katvanger' de afhankelijke variabele en deze variabele is binair. De eerste assumptie wordt niet geschonden.

De tweede assumptie is dat de cases onafhankelijk van elkaar moeten zijn. Deze dataset bestaat uit voertuigen die gebonden zijn aan de individuen uit de eerste datasets. Meerdere voertuigen kunnen dus wel gebonden zijn aan één individu. Elk voertuig heeft zijn eigen kenmerken. De cases zijn dus onafhankelijk van elkaar wat betekent dat deze assumptie niet wordt geschonden.

De derde assumptie is dat er geen sprake mag zijn van multicollineariteit. Hiervoor wordt er weer gekeken naar de VIF-waarden. Een VIF-waarde is goed wanneer deze onder de 10 is. De VIF-waarden van de logistische regressie liggen allemaal rond de 1. Dit betekent dat er geen sterke samenhang is tussen de variabelen, er is geen multicollineariteit. De derde assumptie wordt niet geschonden.

Tot slot moet er geen sprake mag zijn van lineariteit tussen de continue variabelen en de log-odds. In deze logistische regressie is 'aantal CJIB' de enige continue variabele. Voor deze variabele is er sprake van lineariteit, deze assumptie wordt dus geschonden.

Er wordt ook hier gecontroleerd voor outliers door te kijken naar de Cook's Distance, gestandaardiseerde residuen en Mahalanobis Distance. In bijlage 3 wordt meer informatie gegeven

over hoge waardes die deze toetsen laten zien voor sommige cases. Deze hoge waardes betekenen dat er wel een aantal outliers zijn. Er wordt echter hier er ook voor gekozen om ze niet uit de dataset te verwijderen. Wanneer er wordt gekeken naar de analyse lijken ze niet veel invloed uit te oefenen. De outliers scoren op het aantal CJIB-boetes vaak iets hoger, maar deze informatie kan juist veel vertellen.

Tabel 5: parameters van de logistische regressies met achtergrondkenmerken en met katvanger als onafhankelijke variabele

Variabele	Model 1			Model 2		
	<i>b</i> (SD)	Sig.	Exp(B)	<i>b</i> (SD)	Sig.	Exp(B)
<i>Constante</i>	-,66(,02)	<,001	,52	-1,87(,30)	<,001	,15
<i>WOK-status</i>	-	-	-	1,09(,10)	<,001	2,98
<i>APK</i>	-	-	-	22,59(605,20)	,970	6480199525
<i>Aantal CJIB</i>	-	-	-	,25(,03)	<,001	1,28
<i>Verzekerd</i>	-	-	-	,37(,04)	<,001	1,33
<i>Beslag</i>	-	-	-	3,24(,17)	<,001	25,63
<i>Cox and Snell</i>	-	-	-	,44	-	-
<i>Nagelkerke</i>	-	-	-	,62	-	-
<i>McFadden</i>	-	-	-	,46	-	-
<i>Chi-square</i>	-	-	-	12327,87	,00	-
<i>Hosmer & Lemeshow</i>	-	-	-	7,55	,06	-
<i>N</i>	20971					

Analyse – Logistische regressie voertuigkenmerken

Allereerst werd er gekeken of er sprake was van interessante verschillen tussen de variabelen. Uit de descriptieve variabelen blijken er bij sommige variabelen interessante verschillen te zijn tussen niet-katvangers en katvangers die verder kunnen worden onderzocht. Voor deze variabelen is er een logistische regressie uitgevoerd.

Ten eerste wordt gekeken naar de WOK-status. De helling is positief en groot ($b= 1,09$; $SE= ,10$; $p <,001$) en significant. De odds-ratio is 2,98, wat betekent dat de kans op het zijn van een katvanger 2,98 keer hoger is voor een voertuig met een WOK-status dan voor een voertuig zonder WOK-status, na controle voor de andere variabelen.

Ten tweede wordt er gekeken naar APK. De helling is erg groot en positief ($b= 22,60$; $SE=605,20$; $p =,970$), hier is de helling niet significant. Wanneer een voertuig een verlopen APK heeft zal de log-odds op het behoren tot een katvanger met 22,60 toenemen. De odds-ratio van de variabele APK is echter heel hoog. Dit kan komen omdat niet-katvangers allemaal een 1 scoren op de variabele APK. Dit betekent niet dat deze waarden incorrect zijn, aangezien dit nu eenmaal de manier is waarop de niet-katvangers scoorden.

Ook wordt gekeken naar het aantal CJIB-boetes dat aan een voertuig is gekoppeld. De helling is positief en redelijk klein ($b=,25$; $SE=,03$; $p <,001$). De odds-ratio is 1,28, wat betekent dat de kans op het zijn van een katvanger met 1,28 toeneemt voor elk extra CJIB-boete dat aan een voertuig is gekoppeld.

De helling van 'beslag' is positief en groot ($b=3,24$; $SE=,17$; $p <,001$). De odds-ratio is ook hoog, namelijk 25,63. Dit betekent dat de kans 25,63 groter voor een voertuig om te behoren tot een katvanger te zijn wanneer er beslag is gelegd op het voertuig.

Tot slot wordt gekeken of het voertuig voldoet aan de verplichte WA-verzekering. De helling is positief en redelijk klein ($b= ,37$; $SE= ,04$; $p <,001$). De odds-ratio is 1,33. Het hebben van een WA-verzekering geeft een grotere kans op het zijn van een katvanger.

Conclusie en discussie

Het doel van het onderzoek was om meer informatie te verschaffen over katvangers. Er werd onderzoek gedaan naar hun achtergrondkenmerken en wat voor rol zij hebben in de samenleving. De hoofdvraag die werd opgesteld was: *“Wat kenmerkt katvangers en wat zijn de mogelijke effecten van hun faciliterende rol op de veiligheid in de samenleving?”*

Er zijn ook verschillende deelvragen opgesteld. De eerste deelvraag waarop antwoord wordt gegeven is: *“Wat zijn de verschillen tussen katvangers en niet-katvangers en welke indicatoren dragen bij aan de mogelijke kwetsbaarheid van katvangers?”*.

Uit literatuuronderzoek blijkt dat katvangers vaak kwetsbaarder zijn dan niet-katvangers. Dit is zowel op financieel gebied, maar uit zich ook in een gebrek aan bewustzijn (Wissink, 2021). Deze verwachting werd daarom ook gedaan voor het huidige onderzoek.

Katvangers als fiscale wanbetalers

Om de eerste deelvraag te beantwoorden is er een logistische regressie gedaan tussen katvangers en niet-katvangers. Allereerst blijkt dat katvangers over het algemeen minder vaak (,02%) worden gekenmerkt als een fiscale wanbetaler dan niet-katvangers (,4%). De kans dat een katvanger een fiscale wanbetaler is dus lager dan wanneer iemand geen katvanger is. Dit komt niet overeen met de verwachting over financieel kwetsbare katvangers. Katvangers hebben vaak beperkte financiële middelen waardoor er werd verwacht dat zij nalatig zouden zijn in het betalen van de motorrijtuigenbelasting (MRB).

Een verklaring waarom katvangers minder vaak een fiscale wanbetaler zijn is omdat zij vaak maar van korte duur worden ingezet als katvanger. De MRB moet per kwartaal worden betaald. Wanneer de MRB meer dan vijf keer niet wordt betaald, dan pas krijgt een individu de status fiscale wanbetaler (Belastingdienst, 2023). Een katvanger moet dus 15 maanden lang de MRB niet hebben betaald voordat deze status wordt toegekend. Een voertuig op naam wordt vaak maar voor één doeleinde gebruikt, waardoor het gebruik van 15 maanden onwaarschijnlijk lijkt. Een niet-katvanger heeft een voertuig vaak voor meerdere jaren waardoor het veel waarschijnlijker lijkt dat de MRB vergeten wordt om te betalen.

Een andere verklaring hiervoor zou kunnen zijn dat criminelen op de hoogte zijn van het feit dat iemand al een fiscale wanbetaler is. Wanneer dit voor hen bekend is, is het niet interessant om deze persoon in te zetten als katvanger omdat het risico groter wordt dat iemand wordt geregistreerd als katvanger. Des te eerder iemand deze status krijgt, des te sneller er onderzoek naar deze persoon kan worden gedaan. Fiscale wanbetalers worden hoogstwaarschijnlijk vaker onderworpen aan controle en toezicht van de belastingdienst of andere autoriteiten. Dit verhoogde toezicht kan het risico op betrokkenheid bij illegale activiteiten ontmoedigen, inclusief het fungeren als katvanger. Ze zouden

het zelf dus ook kunnen afwijzen omdat zij al in de problemen zitten en niet nog meer aandacht naar zichzelf willen trekken.

Andere vormen van financiële kwetsbaarheid

Verder laat het onderzoek wel zien dat katvangers vaker een CJIB-signaal hebben, namelijk 2,9%, terwijl ,5% van de niet-katvangers een CJIB-signaal hebben. Dit signaal krijgen zij wanneer er iemand vijf of meer boetes niet hebben betaald. Een reden waarom zij vaker dit signaal hebben dan niet-katvangers kan worden verklaard door de financiële beperkingen die katvangers vaak hebben. Zij zijn vaak katvangers geworden om wat te verdienen en kunnen dus niet de boetes die binnenkomen voor dat voertuig betalen. Hierdoor zullen zij vaker dit signaal hebben.

Uit dit onderzoek blijkt ook dat katvangers over het algemeen minder vaak beschikken over een rijbewijs, of nog nooit een rijbewijs hebben gehad (63,5%). Van de niet-katvangers heeft slecht ,1% niet of nooit een rijbewijs gehaald. Dit verschil kan ook worden verklaard door de financiële kwetsbaarheid van katvangers. Rijlessen en een rijbewijs zelf kost veel geld, waardoor het minder haalbaar is voor mensen met minder geld om hun rijbewijs te behalen. Deze bevindingen ondersteunen dus de verwachting die werd gedaan op basis van de literatuur.

Tot slot werd er onderzoek gedaan naar de huishoudens waar de katvangers toe behoorden. De verwachting die werd gedaan is dat zij vaak in armere buurten wonen vanwege de financiële kwetsbaarheid. De resultaten van dit onderzoek laten zien dat de kans op katvangerschap hoger wordt wanneer iemand op een postcode woont die behoort tot de 40% laagste gemiddelde inkomen in Nederland. Katvangers wonen dus vaker in armere buurten dan niet-katvangers wat ook hier kan worden verklaard door de financiële kwetsbaarheid.

Achtergrondkenmerken

Naast de financiële kwetsbaarheid werd er ook gekeken naar verschillende kenmerken waarover katvangers beschikken. Er werd gecontroleerd voor drie verschillende variabelen. Literatuuronderzoek toont aan dat mannen vaker slachtoffer van katvangerschap lijken te raken dan vrouwen (Raza et al., 2020; Wissink, 2021; Kuppens et al., 2006; Arevalo, 2015). De resultaten van dit onderzoek wijken af van de gevonden literatuur. In de dataset is de meerderheid man, zowel voor katvangers als niet-katvangers. De analyse toont echter aan dat de kans groter is voor vrouwen om een katvanger te zijn dan mannen. Een reden hiervoor kan zijn dat er sprake is van een confounding variabele die niet is opgenomen in de dataset, maar die wel het verband tussen katvanger en geslacht kan beïnvloeden.

Uit de extra analyse die hiervoor is uitgevoerd bleek dat de variabele rijbewijs de hellingrichting veranderde van negatief naar positief. De helling blijft negatief wanneer alleen de variabelen geslacht en rijbewijs in de analyse worden meegenomen, maar wanneer leeftijd, nationaliteit, fiscale wanbetaling en CJIB-signaal ook worden meegenomen, laat de analyse zien dat de

helling van geslacht verandert. Dit gedeelte vereist vervolgonderzoek. Er kan nu namelijk moeilijk worden herleid waarom de helling van richting verandert. Er is geen sprake van een te hoge multicollineariteit, de variabelen zijn dus niet te sterk met elkaar gecorreleerd.

De tweede controlevariabele is leeftijd. Er werd de verwachting gemaakt dat katvangers over het algemeen jonge adolescenten zijn (Esoimeme, 2021). De literatuur is verdeeld over de gemiddelde leeftijd van katvangers, maar ze zijn het er wel over eens dat des te ouder een individu, des te onwaarschijnlijker het is dat zij katvanger zullen zijn (Wissink, 2021). De resultaten van dit onderzoek komen overeen met de bevindingen uit de gevonden literatuur. Jongeren zijn over het algemeen beïnvloedbaarder dan volwassenen en zijn makkelijker te strikken voor dit soort criminele klussen. Ook zijn zij vaak minder bewust van de consequenties van hun eigen gedrag en keuzes (Arevalo, 2015).

De laatste controlevariabele is nationaliteit. Onderzoek toont aan dat katvangers ook vaak buitenlanders zijn die worden ingezet door criminelen (van Noordenburg, 2020). Er werd daarom ook de verwachting gedaan dat katvangers vaak een niet-Nederlandse nationaliteit hebben. De resultaten van dit onderzoek zijn niet in lijn met de gevonden literatuur. Katvangers lijken voor het merendeel een Nederlandse nationaliteit te hebben, namelijk 87,7%. Wel ligt dit percentage iets lager dan die van niet-katvangers, waar 99,7% een Nederlandse nationaliteit heeft.

Toch blijkt uit de analyse dat de kans op het zijn van een katvanger groter is voor niet-Nederlandse individuen. Een reden hiervoor kan zijn dat de economische omstandigheden van niet-Nederlandse individuen hen vatbaarder maken voor het fungeren als katvanger. Het is ook belangrijk om te overwegen of er sprake is van sociaal onrecht dat de kansen op werk voor niet-Nederlandse individuen beïnvloedt, waardoor ze mogelijk eerder betrokken raken bij criminele activiteiten zoals het fungeren als katvanger.

Impact op de verkeersveiligheid

De tweede deelvraag die is opgesteld luidt als volgt: *“Wat voor impact hebben katvangers op de verkeersveiligheid?”*. Voor deze deelvraag is er gekeken naar verschillende variabelen gekoppeld aan de voertuigen die verkeersveiligheid kunnen beïnvloeden. Ook hier is er een logistische analyse gedaan om het verschil tussen voertuigen van katvangers en niet-katvangers te onderzoeken.

De voertuigen op naam van katvangers worden vaak gebruikt bij criminele activiteiten. Dit zorgt er op zichzelf al voor dat de veiligheid in gevaar komt omdat katvangers een soort facilitator zijn voor criminaliteit. De criminelen komen vaak weg met hun daden omdat zij door de katvangers anoniem kunnen blijven.

Naast de algemene veiligheid beïnvloeden deze voertuigen ook de verkeersveiligheid. Doordat er geen consequenties zijn voor de persoon die in het voertuig rijdt, is het makkelijker om roekelozer te rijden en om niet te voldoen aan de verplichtingen. De bestuurder zal nooit voor de

boetes opdraaien, maar de katvanger wel (Wissink, 2021). Er wordt daarom allereerst de verwachting gedaan dat de voertuigen van katvangers over het algemeen minder vaak APK-gekeurd zullen zijn. Uit de resultaten van dit onderzoek wordt deze verwachting ondersteunt. Meer dan de helft van de voertuigen van katvangers is niet APK-gekeurd (59,1%), terwijl 100% van de voertuigen van niet-katvangers wel is APK-gekeurd. Uit de analyse blijkt dan ook dat de kans groter is dat een voertuig behoort tot een katvanger wanneer deze niet APK-gekeurd is. Deze conclusie moet voorzichtig worden getrokken omdat het resultaat niet significant is.

Het niet APK-keuren van een voertuig kan een negatieve invloed hebben op de verkeersveiligheid. Des te langer een voertuig rondrijdt zonder gekeurd te worden, des te groter de kans is dat dit voertuig gebreken heeft. Denk hier bijvoorbeeld aan versleten remmen of slechte verlichting. Deze gebreken kunnen gemakkelijk ongelukken veroorzaken. De voertuigen van katvangers ondergaan minder vaak keuringen, wat resulteert in een verhoogd risico op gebreken in vergelijking met voertuigen van niet-katvangers. De voertuigen van katvangers hebben dus sneller een negatieve impact op de verkeersveiligheid.

Toch is er ook een aantal voertuigen van katvangers die wel APK-gekeurd zijn. Sommige criminelen zouden ervoor kunnen kiezen om hun voertuigen wel te laten keuren, als ze van plan zijn om ze voor een langere periode te gebruiken. Ze willen mogelijk voorkomen dat ze bij een aanhouding door de politie worden verdacht van het inzetten van een katvanger, vooral als het voertuig bijvoorbeeld niet verzekerd is en niet de vereiste APK-keuring heeft ondergaan. De pakkans wordt kleiner wanneer zij wel aan alle verplichtingen voldoen.

Naast de APK-keuring moet een voertuig ook verplicht verzekerd zijn. Ook hier werd de verwachting gedaan dat de voertuigen van katvangers minder vaak verzekerd worden omdat de pakkans voor de criminelen klein is. De resultaten tonen aan dat de grote meerderheid van de voertuigen van katvangers niet verzekerd zijn, namelijk 73,7%, waardoor de verwachting die is gedaan wordt ondersteund.

Toch blijkt ook een groot deel, namelijk 32,1%, van de niet-katvangers een voertuig hebben die niet verzekerd is. Dit aantal is laag, gezien het verzekeren van een voertuig verplicht is en hier anders een boete voor wordt uitgedeeld. Tijdens de stage aan de RDW werd er ook een klantreis gedaan. Dit is de weg die een klant aflegt bij het aanschaffen van een dienst of een product. De klantreis die werd uitgevoerd betrof het kopen van een tweedehands voertuig. Hieruit bleek dat veel mensen vergeten dat na het overschrijven van een voertuig ook verplicht zijn om het voertuig te verzekeren. Dit zou het geval kunnen zijn bij de niet-katvangers. Veel mensen kopen een tweedehands voertuig en zouden dus sneller kunnen vergeten hun voertuig te verzekeren.

Ten derde werd er gekeken naar de WOK-status, nu ook wel bekend als 'verbod voor rijden op de weg'. Er werd de verwachting gemaakt dat de voertuigen van katvangers vaker een WOK-status hebben omdat de criminelen er roekelozer in rijden. De resultaten tonen aan dat in vergelijking met

niet-katvangers (2,3%) de voertuigen van katvangers vaker een WOK-status hebben (3,6%). Toch is dit percentage erg klein. Het is belangrijk om op te merken dat een WOK-status uitzonderlijk is en relatief weinig voorkomt. Het is dus nog steeds mogelijk dat criminelen meer ongelukken veroorzaken, maar dat de schade niet zodanig ernstig is dat een WOK-status wordt toegekend aan het voertuig.

Vanwege de verwachting dat criminelen waarschijnlijk roekelozer rijgedrag vertonen in de voertuigen, werd ook verwacht dat ze gemiddeld vaker verkeersovertredingen zouden begaan en daardoor frequenter boetes zouden ontvangen. De resultaten tonen aan dat katvangers gemiddeld een hoger aantal CJIB-boetes heeft (,21) dan niet-katvangers (,16). De verwachting wordt met deze bevindingen ondersteund.

Tot slot wordt er verwacht dat er vaker beslag wordt gelegd op deze voertuigen met de veronderstelling dat criminelen vaker verkeersovertredingen maken (Nota katvangersproblematiek RDW, 2020). De resultaten tonen aan dat een slechts 7,4% van de voertuigen van katvangers in beslag is genomen. Dit geldt ook voor de voertuigen van niet-katvangers, maar dit aantal ligt lager dan dat bij de voertuigen van katvangers, namelijk slechts ,3%.

Preventie van katvangers

De laatste deelvraag waarop een antwoord wordt gegeven is: *“Welke preventieve maatregelen kunnen met kennis over achtergrondkenmerken en trends over tijd worden opgesteld om het inzetten van katvangers te voorkomen?”*.

Er wordt allereerst gekeken naar bestaande maatregelen en hoe effectief deze op dit moment zijn. Het proces van een voertuig op naam zetten blijkt relatief makkelijk te zijn. Een van de voorwaarden is dat de persoon woonachtig moet zijn in Nederland en dus ingeschreven moet staan bij de Basisregistratie Personen (BRP). Wanneer er sprake is van twijfel met betrekking tot het adres van iemand kan de gemeente hier verder onderzoek naar doen. Bij ongeveer 1% van de inwoners van een gemeente is het adres in onderzoek. Uit dat onderzoek blijkt dat 85% van die inwoners niet meer op het opgegeven adres woont. Adresfraude is eenvoudig, bijvoorbeeld via illegale huurcontracten op platforms zoals Marktplaats (J. de Graaf, persoonlijke communicatie, 1 februari 2023). Zolang een adres in onderzoek is, wordt het aanvragen van een tenaamstelling geweigerd.

Een andere mogelijk is het aanhouden van een persoon zolang het niet duidelijk dat deze persoon daadwerkelijk woonachtig is in Nederland, of een nep-adres heeft doorgegeven. Het nadeel van deze aanpak is echter dat het proces van het achterhalen van het adres lang kan duren, en de verkoper van het voertuig heeft niet altijd de bereidheid om hierop te wachten. Ook kost het veel moeite en mankracht om mensen aan te houden. Daarom heeft het weigeren van de tenaamstelling de voorkeur (nota katvangersproblematiek RDW, 2020). Deze aanpak is efficiënter, maar ook hier blijft

de crimineel achter de katvanger onbekend, de katvanger zelf zal dan worden aangehouden, niet de crimineel.

Verder wordt er bij een rijbewijsomwisseling binnen 6 weken na inschrijving bij de BRP wordt ook het adres onderzocht. Als de persoon niet meer in Nederland woont, wordt de aanvraag geweigerd. De aanvraag zal hoe dan ook worden geweigerd als de aanvrager de status emigratie heeft (Nota fenomeen katvangers RDW, 2019).

Een persoon met de status katvanger in het kentekenregister kan geen voertuigen meer op naam zetten. Als deze persoon bij de RDW een ander adres opgeeft dan het geregistreerde adres, wordt dit alternatieve adres gecontroleerd. Indien het overeenkomt met een adres wat al verbonden staat aan een ander individu met de status katvanger wordt de aanvraag afgewezen (Nota fenomeen katvangers RDW, 2019).

Ondanks deze maatregelen zijn er nog duizenden katvangers. Het langdurige adresonderzoek levert vaak geen resultaten op en het stopzetten van de tenaamstelling lijkt niet volledig te werken. Er moet een maatregel komen voor het feit dat het zo gemakkelijk is om een voertuig op naam te kunnen zetten.

Voor het aanvragen van een voertuig is een Burgerservicenummer (BSN) nodig. Deze wordt verkregen wanneer iemand staat ingeschreven bij Basisregistratie Personen (BRP). Het aanvragen van een BSN is eenvoudig en snel in Nederland, met 19 gemeentes die een RNI-loket hebben. Hier kunnen buitenlanders hun BSN-aanvragen en ontvangen deze binnen een korte termijn. Dit maakt het mogelijk voor buitenlanders om snel een BSN aan te vragen, een voertuig op naam te zetten, en het land te verlaten (Fieldlab katvangers RDW, 2019).

Vanuit het Landelijk Intelligence- en expertisecentrum Voertuigcriminaliteit (LIV) is een fieldlab in Den Haag gestart vanwege het opvallend hoge aantal BSN-aanvragen. Een BSN werd hier erg snel geschreven. Iemand weigeren die een BSN aanvraagt is niet toegestaan, dus zelfs wanneer het duidelijk is dat het gaat om een katvanger. Na constatering van het snelle en eenvoudige proces, is het beleid aangepast. Nu duurt het langer om een BSN te verkrijgen, met gericht onderzoek en meer vragen, vooral bij vermoeden van katvangers (Fieldlab katvangers RDW, 2019). Deze vertraging en intensievere controle maken het minder aantrekkelijk om een BSN aan te vragen, met mogelijk een afname van het aantal katvangers als gevolg.

Dit beleid is op dit moment alleen werkzaam in de gemeente Den Haag en nog niet in andere gemeenten. Wat opviel is dat de hoeveelheid aanvragen in Den Haag daalde, maar weer opliep in andere gemeenten. Er is hier dus mogelijk sprake van een waterbedeffect (Fieldlab katvangers RDW, 2019). Op één plek wordt criminaliteit bestreden, maar op een andere plek zal hierdoor de criminaliteit toenemen. Om dit beleid effectief te maken, moet het worden toegepast in alle gemeenten waar BSN kan worden aangevraagd.

Dit beleid is alleen effectief voor buitenlandse katvangers; mensen met een bestaand BSN kunnen het aanvragen overslaan. Het onderzoek laat zien dat de meerderheid van de katvangers een Nederlandse nationaliteit heeft. Het is daarom van belang om het proces op een ander punt bemoeilijken. Het blijft echter moeilijk om het registratieproces van voertuigen lastiger te maken, gezien de impact op niet-katvangers.

Het is interessanter om te zorgen voor een snellere detectie van katvangers in het kentekenregister. Onderzoek toont aan dat voertuigen van katvangers vaak niet aan verplichtingen voldoen, zoals APK-keuringen of dat de boetes niet worden betaald. Het is daarom belangrijker dat deze voertuigen sneller worden opgespoord door de politie en de RDW. Een snellere detectie kan leiden tot een snellere opsporing van de bestuurder en dus de crimineel. Wanneer duidelijk is welk voertuig op naam staat van een katvanger, kan er worden gescreend voor dit voertuig op de weg. Identificatie van voertuigen op naam van een katvanger maakt gerichte screening mogelijk, waardoor de crimineel achter de katvanger sneller kan worden aangehouden.

De politie maakt momenteel gebruik van automatische kentekenherkenning (ANPR). Met behulp van camera's worden kentekens in het verkeer vastgelegd. Deze techniek vergelijkt kentekens met bekende personen in politiedatabases. Wanneer iemand sneller wordt geregistreerd als katvanger kan er accurater worden gecontroleerd op dit kenteken. Verbeterde registratie van katvangers maakt nauwkeurigere controle mogelijk. Een regeling waarbij de politie een persoon mag staande houden bij een vermoeden van betrokkenheid bij katvangers is noodzakelijk voor effectieve opsporing. Ook hier is samenwerking tussen de politie en de RDW vereist. De RDW registreert de voertuigkatvangers in het kentekenregister, terwijl de politie dit nergens specifiek registreert. De politie moet daaro gebruik maken van die registratiegegevens om de katvangers op te sporen.

Tot slot is het van belang om de verkeersveiligheid effectiever te beschermen. Het is noodzakelijk om aanvullende maatregelen te treffen, aangezien de bestaande verplichtingen op zichzelf niet lijken te volstaan. Een eerste stap is het oprichten van een speciaal politieteam gericht op het identificeren van voertuigen gekoppeld aan katvangers. Het is hierbij noodzakelijk dat er een sterkere samenwerking komt tussen de RDW, politie, verzekeringsmaatschappijen en het CJIB. Het is van belang dat alle instanties evenveel informatie hebben. Dit zou kunnen resulteren in een snellere reactie bij het aanpakken van katvangers. Aangezien er op dit moment geen reguliere wegcontroles plaatsvinden voor bijvoorbeeld de geldigheid van de APK of verzekering, kan een speciaal team zich hierop focussen.

Met behulp van de bevindingen op de deelvragen kan de hoofdvraag: *“Wat kenmerkt katvangers en wat zijn de mogelijke effecten van hun faciliterende rol op de veiligheid in de samenleving?”* worden beantwoord.

Katvangers zijn doorgaans jonge Nederlandse mannen met financiële moeilijkheden. Ze hebben vaak geen rijbewijs, wonen in armere wijken en betalen vaak hun boetes niet. Katvangers zijn

facilitators voor criminaliteit omdat door hen de criminelen anoniem blijven. Dit heeft een negatief effect op de maatschappij, aangezien ze crimineel gedrag hierdoor ongestraft laten. Bovendien hebben voertuigkatvangers een nadelige invloed op de verkeersveiligheid. Criminelen zetten deze voertuigen in bij illegale activiteiten en rijden vaker roekeloos omdat de pakkans klein is. Ze verschuilen zich achter de katvangers en ontlopen zo de gevolgen van hun daden.

Beperkingen

Er is een aantal beperkingen aan het onderzoek die kort zullen worden besproken. Deze beperkingen kunnen namelijk invloed hebben op de interpretatie van de resultaten en kunnen duidelijk maken op welke gebieden vervolgonderzoek zich moet richten.

Een beperking is de ongelijke verdeling van mannen en vrouwen in de random sample. Er zijn meer mannen dan vrouwen opgenomen in het onderzoek, namelijk 73,1%. Dit grote verschil kan leiden tot vertekende resultaten en maakt generaliseerbaarheid moeilijker. De data zijn verkregen door een random steekproef te doen uit het kentekenregister, maar in vervolgonderzoek is het noodzakelijk om erop te letten dat deze steekproef goed wordt uitgevoerd en dat er met zekerheid kan worden geconstateerd dat het gaat om een random sample.

Een andere beperking is dat wanneer een individu een CJIB-signaal krijgt in het kentekenregister, deze persoon geen nieuwe boetes meer opgelegd zal krijgen. Uit het onderzoek blijkt dat katvangers vaker een CJIB-signaal hebben. Dit kan ervoor zorgen dat het aantal boetes gemiddeld lager liggen dan verwacht en dat daarom de trend van het aantal boetes op naam van katvangers gedaald is.

Het is belangrijk op te merken dat zodra een individu de status van katvanger krijgt in het kentekenregister, er geen nieuwe boetes aan hen worden opgelegd. Ook moet er rekening worden gehouden dat er niet met zekerheid kan worden gezegd of het specifiek om verkeersovertredingen gaat, omdat katvangers ook boetes kunnen ontvangen voor andere zaken, zoals het niet hebben van een verzekering. Het zou interessant zijn voor vervolgonderzoek om dieper in te gaan op het soort boetes dat katvangers ontvangen, om definitief vast te stellen of zij daadwerkelijk vaker betrokken zijn bij verkeersovertredingen en dus potentieel gevaarlijker zijn in het verkeer.

Daarnaast blijkt ook een groot aantal van de niet-katvangers niet verzekerd te zijn. Dit is opmerkelijk, aangezien dit een verplichting is waar menig mens wel aan wil en kan voldoen. Het is dus interessant om dit verder te onderzoeken wat de redenen hiervoor kunnen zijn, of dat het ligt aan de willekeurige steekproef die gebruikt is in dit onderzoek.

De grootste beperking van dit onderzoek is de verschillende soorten datasets en hoe deze data zijn verzameld. Er is niet duidelijk hoe de random steekproef is uitgevoerd. Ook bevatte de dataset hoogstwaarschijnlijk bedrijven waarvoor is gekozen om deze eruit te filteren. Dit kon niet met zekerheid worden vastgesteld waardoor de kans aanwezig is dat er bruikbare informatie verloren is

gegaan door ze eruit te filteren. Voor vervolgonderzoek wordt er aangeraden om tijdens het samenstellen van de dataset onderscheid te maken tussen individuen en bedrijven zodat er achteraf geen data verloren gaan. Dit onderzoek vormt een eerste stap in het verkennen van voertuigkatvangers. Het gebrek aan uitgebreid onderzoek naar voertuigkatvangers is toe te schrijven aan de relatieve onbekendheid van dit fenomeen en de complexiteit ervan, waardoor het bedenken van effectieve oplossingen veel tijd en moeite kost. Hierdoor heeft dit fenomeen vaak geen prioriteit.

Dit onderzoek had als voornaamste doel meer kennis te verkrijgen over het fenomeen voertuigkatvangers en daarmee bij te dragen aan de beperkte wetenschappelijke kennis over dit fenomeen. Gebaseerd op de verworven kennis had het onderzoek het doel om passende maatregelen op te stellen waarbij specifiek wordt gekeken naar de kenmerken van katvangers uit de resultaten van het onderzoek. Op deze manier kan het probleem effectief, maar ook efficiënt worden aangepakt.

Bovendien draagt dit onderzoek bij aan het vergroten van het bewustzijn over de bredere implicaties van voertuigkatvangpraktijken in de samenleving. Ondanks de relatieve onbekend over voertuigkatvangers hebben zij een grote impact op de samenleving. Zij spelen een substantiële rol in de georganiseerde criminaliteit en faciliteren het voortbestaan hiervan. Het vergroten van bewustwording door middel van uitgebreider onderzoek kan wellicht bijdragen aan het voorkomen dat mensen sneller slachtoffer worden, aangezien zij nu beter bekend zijn met het fenomeen.

Er wordt gehoopt dat dit onderzoek dient als een waardevolle bron en ook als een fundament voor vervolgonderzoek naar katvangers. Met meer kennis kan het probleem effectief worden aangepakt wat bijdraagt aan de veiligheid in de samenleving.

Literatuurlijst

- Arevalo, B.C. (2015). Money mules: Facilitators of financial crime. An explorative research on money mules (MA thesis). Utrecht: Universiteit Utrecht
- Belastingdienst. (2023, 29 maart). Belastingen op auto en motor.
https://www.belastingdienst.nl/wps/wcm/connect/bldcontentnl/belastingdienst/privé/auto_en_vervoer/belastingen_op_auto_en_motor/
- Esoimeme, E. E. (2021). Identifying and reducing the money laundering risks posed by individuals who have been unknowingly recruited as money mules. *Journal of Money Laundering Control*, 24(1), 201–212. <https://doi.org/10.1108/jmlc-05-2020-0053>
- Florencio, D., & Herley, C. (2010). Phishing and money mules.
<https://doi.org/10.1109/wifs.2010.5711465>
- Hülse, R. (2017). The money mule: Its discursive construction and the implications. *Vanderbilt Journal of Transnational Law*, 50, 1007-1032.
- Ingevorderd Rijbewijs. (2023). Wanneer neemt de politie mijn auto in? Ingevorderd Rijbewijs.
<https://www.ingevorderd-rijbewijs.nl/inbeslagname-auto-terugkrijgen/>
- Leukfeldt, R., & Jansen, J. (2015). Cyber Criminal Networks and Money Mules: An Analysis of Low-Tech and High-Tech Fraud Attacks in the Netherlands. *Zenodo (CERN European Organization for Nuclear Research)*. <https://doi.org/10.5281/zenodo.56210>
- Ministerie van Justitie en Veiligheid. (2020, 24 november). *Vraag en antwoord Beslag*. Beslag | Openbaar Ministerie. <https://www.om.nl/onderwerpen/beslag/faq>
- Oerlemans, J.J., Custers, B.H.M., Pool, R.L.D., & Cornelisse, R. (2016). Cybercrime en witwassen: Bitcoins, online dienstverleners en andere witwasmethoden bij banking malware en ransomware. Den Haag: WODC.
- Raad van State. (z.d.). Uitspraak 202106271/1/A2.
<https://www.raadvanstate.nl/uitspraken/@132437/202106271-1-a2/>

- Raza, M. S., Zhan, Q., & Rubab, S. (2020). Role of money mules in money laundering and financial crimes a discussion through case studies. *Journal of Financial Crime*, 27(3), 911–931.
<https://doi.org/10.1108/jfc-02-2020-0028>
- RDW. (z.d.-a). *Dossier voertuigcriminaliteit*. . <https://www.rdw.nl/over-rdw/actueel/dossiers/voertuigcriminaliteit>
- RDW. (z.d.-b). *Verplichtingen bij gebruik*.
<https://www.rdw.nl/particulier/voertuigen/auto/verplichtingen-bij-gebruik>
- RDW. (z.d.-c). *APK*. <https://www.rdw.nl/particulier/voertuigen/auto/apk>
- RDW. (z.d.-d). Keuren na verbod voor rijden over de weg.
<https://www.rdw.nl/particulier/voertuigen/brommer/keuren/keuren-na-verbod-voor-rijden-over-de-weg>
- Sluis, J. (2021). Voorkom een miskoop en check de WOK-status. *Vaartland.nl*.
<https://www.vaartland.nl/blog/auto-advies/auto-wetgeving/wok-status-controleren>
- Van Hulst, B., & Hoff, S. (2019). Armoede in Kaart. Sociaal en Cultureel Planbureau.
<https://digitaal.scp.nl/armoedeinkaat2019/waar-wonen-de-armen-in-nederland/>
- Van Leiden, I., Wolsink, J., & Ferwerda, H. (2019). *CRIMINELEN ACHTER HET STUUR: Aard en omvang van het gebruik van huurmotorvoertuigen voor criminele activiteiten*. Ministerie van Justitie en Veiligheid.
- Van Noordenburg, C. (2020). De Roemeense katvanger als faciliteerder van mobiel banditisme. *Centrum voor Criminaliteitspreventie en Veiligheid (CCV)*.
<https://hetccv.nl/onderwerpen/mobiele-bendes/documenten/de-roemeense-katvanger-als-faciliteerder-van-mobiel-banditisme/>
- Van Noordenburg, C. (2022, 19 augustus). Waarom het belangrijk blijft katvangers te dwarsbomen en hoe je dat dan doet. *Centrum voor Criminaliteitspreventie en Veiligheid (CCV)*.
<https://hetccv.nl/nieuws/waarom-het-belangrijk-blijft-katvangers-te-dwarsbomen-en-hoe-je-dat-dan-doet/>
- Versprille, I. (2022, 20 oktober). *De katvanger in beeld*. ToeZine. <https://www.toezine.nl/de-katvanger-in-beeld/>

Verzekering.nl B.V. (z.d.). *Wat is een WA-verzekering?* Verzekering.nl.

<https://www.verzekering.nl/autoverzekering/wat-is-een-wa-verzekering/>

Wissink, I. B. (2021). 'If it's too good to be true, it's too good to be true': In *Wegwijzer Jeugd en Veiligheid*. Bureau Halt & Betaalvereniging Nederland. <https://wegwijzerjeugdenveiligheid.nl/if-its-too-good-to-be-true-its-too-good-to-be-true>

Bijlage 1 – operationalisaties variabelen

Dataset achtergrondkenmerken

Missing waarden verwijderen

Beide datasets bestaan uit dezelfde variabelen wat betekent dat er voor beide datasets dezelfde stappen zijn genomen om de missing data uit de dataset te halen. Dit is gedaan om voor elke variabele een dummy te maken en vervolgens een nieuwe variabele genaamd 'Miss' aan te maken die bestaat uit het product van deze dummy variabelen. Vervolgens wordt er gefilterd op miss < 1. Op deze manier worden de missing waarden verwijderd.

```
RECODE BURGELYKE_STAAT GESLACHT POSTCODE_NUMERIEK AANTAL_OP_ADRES RYB_IND CJIB_SIGNAAL  
  ANDERS_DAN_NL_NAT Leeftijd REDEN_EIND_NP(SYSMIS=1) (ELSE=0) INTO DumBurgSt DumGes DumPost DumAantAd DumRij  
  DumCJIB DumEtn DumLeef DumReNP.  
EXECUTE.
```

```
COMPUTE Miss=DumBurgSt + DumLeef + DumGes + DumPost + DumAantAd + DumRij + DumCJIB + DumEtn.  
EXECUTE.
```

```
USE ALL.  
COMPUTE filter_$=(Miss < 1).  
VARIABLE LABELS filter_$ 'Miss < 1 (FILTER)'.  
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.  
FORMATS filter_$ (f1.0).  
FILTER BY filter_$.
```

Katvanger

In de dataset met katvangers is er een nieuwe variabele gecreëerd met de waarde 0 genaamd katvanger. De syntax hieronder toont dat de cases onder de 10000 hiervoor zijn geselecteerd.

```
COMPUTE Katvanger=ID < 10000 = 1|  
EXECUTE.
```

Dit is ook gedaan voor de dataset die bestaat uit de willekeurige steekproef, maar daar heeft de waarde katvanger een 0 gekregen.

```
COMPUTE Katvanger=ID < 10000 = 0.  
EXECUTE.
```

Vervolgens zijn de twee datasets samengevoegd. 0 = niet-katvanger en 1 = katvanger. Er zijn in totaal 9976 (50,7%) niet katvangers en 9715 (49,3) katvangers.

Katvanger					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	9976	50,7	50,7	50,7
	1,00	9715	49,3	49,3	100,0
	Total	19691	100,0	100,0	

Na het verwijderen van de outliers wordt dit de nieuwe verdeling:

Katvanger

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	niet-katvanger	5408	35,9	35,9	35,9
	katvanger	9662	64,1	64,1	100,0
	Total	15070	100,0	100,0	

Fiscale Wanbetaling

Inhoudelijk is er aan de variabele fiscale wanbetaling niets aangepast. 0 = geen fiscale wanbetaler en 1 = wel fiscale wanbetaler.

FISC_WANBETALER_SIGNAAL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	9979	100,0	100,0	100,0
	1	2	,0	,0	100,0
	Total	9981	100,0	100,0	

FISC_WANBETALER_SIGNAAL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	9660	100,0	100,0	100,0
	1	2	,0	,0	100,0
	Total	9662	100,0	100,0	

99,98% van de katvangers is geen fiscale wanbetaler. ,02% wordt wel bestempeld als fiscale wanbetaler. Deze verdeling is niet veranderd nadat de missing waarden eruit zijn gehaald. Bijna iedereen heeft geen fiscale wanbetaler signaal.

FISC_WANBETALER_SIGNAAL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	9965	99,7	99,7	99,7
	1	35	,4	,4	100,0
	Total	10000	100,0	100,0	

FISC_WANBETALER_SIGNAAL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	5389	99,6	99,6	99,6
	1	19	,4	,4	100,0
	Total	5408	100,0	100,0	

Hieruit blijkt dat de dataset met willekeurige mensen meer fiscale wanbetalers heeft dan de dataset met katvangers. ,4% wordt bestempeld als katvanger. Toch blijft de meerderheid (99,6%) geen fiscale wanbetaler. Deze verdeling verschilt wederom nauwelijks met de gegevens voordat de missing waarden zijn verwijderd.

CJIB-sigitaal

Voor de variabele CJIB-sigitaal is inhoudelijk niks veranderd. 0 = geen CJIB-sigitaal en 1 = wel CJIB-sigitaal.

CJIB_SIGNAAL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	9678	97,0	97,0	97,0
	1	303	3,0	3,0	100,0
	Total	9981	100,0	100,0	

CJIB_SIGNAAL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	9386	97,1	97,1	97,1
	1	276	2,9	2,9	100,0
	Total	9662	100,0	100,0	

Hier valt ook op dat na het verwijderen van de missing waarden de verdeling bijna hetzelfde blijft. 97,1%, de grote meerderheid van de katvangers, heeft geen CJIB-sigitaal. 2,9% heeft wel een CJIB-sigitaal gekregen in het kentekenregister.

CJIB_SIGNAAL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	9953	99,5	99,5	99,5
	1	47	,5	,5	100,0
	Total	10000	100,0	100,0	

CJIB_SIGNAAL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	5383	99,5	99,5	99,5
	1	25	,5	,5	100,0
	Total	5408	100,0	100,0	

Ook voor de data met random personen geldt dat de grote meerderheid, bijna iedereen, geen CJIB-sigitaal heeft (99,5%). Slechts ,5% heeft wel een CJIB-sigitaal. De verdeling is hier niet veranderd na het verwijderen van de missende waarden.

Rijbewijs

RECODE RYB_IND (0=1) (1=0).
EXECUTE.

		RYB_IND			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	6229	62,4	62,4	62,4
	1	3752	37,6	37,6	100,0
Total		9981	100,0	100,0	

		RYB_IND			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	6132	63,5	63,5	63,5
	1	3530	36,5	36,5	100,0
Total		9662	100,0	100,0	

Van de variabele rijbewijs is de codering omgedraaid. 0 = nooit rijbewijs gehad en 1 = heeft rijbewijs (gehad). Na het verwijderen van de missing waarden blijft de verdeling vrijwel hetzelfde. In de dataset met katvangers heeft 63,5% een nooit een rijbewijs gehad en 36,5% heeft een rijbewijs (gehad).

		RYB_IND			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	14	,1	,1	,1
	1	9986	99,9	99,9	100,0
Total		10000	100,0	100,0	

		RYB_IND			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	7	,1	,1	,1
	1	5401	99,9	99,9	100,0
Total		5408	100,0	100,0	

Ook hier verandert de verdeling eigenlijk niet na het verwijderen van de missing waarden. In de dataset met random individuen valt op dat vrijwel iedereen een rijbewijs heeft (gehad), namelijk 99,9%. Maar 14 mensen (,1%) heeft nooit een rijbewijs gehad, maar toch ooit een voertuig op naam gehad.

Postcode & percentage huishouden

De variabele postcode is toegevoegd om te kunnen combineren met de data van het CBS. De data van het CBS laat zien hoeveel procent van de huishoudens in een postcode onder 40% van het gemiddelde huishoudensinkomen vallen. De postcodes uit eigen dataset zijn vervolgens oplopend gesorteerd en gekopieerd in een Excel bestand. Vervolgens zijn de postcodes in de dataset van het CBS ook oplopend gesorteerd en geplakt in het Excel bestand samen met hun corresponderende waarden. Er is daarna gebruik gemaakt van een formule om de data met elkaar te matchen voor elke postcode. De postcodes waar geen data voor beschikbaar was kreeg de waarde 99. Soms misten er ook waarden, deze kregen dan de waarde 0. Deze waarden zijn geplakt in de dataset met achtergrondkenmerken. Vervolgens zijn de waarde 0 en 99 gehercodeerd naar system-missing. Uiteindelijk is er dus voor de postcodes bekend hoeveel procent van de huishoudens onder het gemiddelde inkomen vallen.

Aantal op adres

Voor deze variabele wordt er een aparte filter aangemaakt. In de dataset komen namelijk een paar cases voor waar 234157 individuen op één adres wonen. De postcode hiervan is ook 0 en zorgen voor verwarring in de data. Dit zijn enorme outliers en worden daarom gefilterd uit de data.

USE ALL.

COMPUTE filter_\$=(Miss < 1 & AANTAL_OP_ADRES < 234157).

VARIABLE LABELS filter_\$ 'Miss < 1 & AANTAL_OP_ADRES < 234157 (FILTER)'.
VALUE LABELS filter_\$ 0 'Not Selected' 1 'Selected'.

FORMATS filter_\$ (f1.0).

FILTER BY filter_\$.

EXECUTE.

Leeftijd

Leeftijd wordt gemeten in geboortjaar. Dit maakt het interpreteren van de leeftijd moeilijker, waardoor er een nieuwe variabele is gemaakt. Er is een nieuwe variabele genaamd leeftijd gemaakt door het geboortjaar van 2022 af te trekken. Er is gekozen voor 2022 omdat de data is verzameld aan het begin van 2023. Dit maakt een nieuwe variabele met de leeftijd in jaren. Dit wordt voor zowel de dataset met katvangers als de dataset met de willekeurige steekproef gedaan.

COMPUTE Leeftijd=2022 - GEBOORTEJAAR.

EXECUTE.

Statistics			Statistics		
Leeftijd			Leeftijd		
N	Valid	9981	N	Valid	9662
	Missing	0		Missing	0
Mean		47,3466	Mean		47,3234
Median		47,0000	Median		47,0000
Mode		52,00	Mode		53,00
Std. Deviation		14,49248	Std. Deviation		14,53956
Minimum		18,00	Minimum		18,00
Maximum		104,00	Maximum		99,00

De gemiddelde leeftijd van katvangers voordat de missing waarden zijn verwijderd is 47,35. De standaarddeviatie is 14,49. De spreiding is normaal. De jongste katvanger is 18 en de oudste katvanger is 104. De meest voorkomende leeftijd is 52. Na het verwijderen van de missing waarden uit de dataset zijn verwijderd is de gemiddelde leeftijd iets lager, namelijk 47,32. De standaarddeviatie wordt een beetje hoger, 14,54, maar de spreiding blijft normaal. Het minimum en maximum verandert niet. De meest voorkomende leeftijd is nu 53.

Statistics			Statistics		
Leeftijd			Leeftijd		
N	Valid	10000	N	Valid	5408
	Missing	0		Missing	0
Mean		51,3585	Mean		50,8530
Median		52,0000	Median		51,0000
Mode		59,00	Mode		52,00
Std. Deviation		16,41798	Std. Deviation		16,12344
Minimum		16,00	Minimum		16,00
Maximum		122,00	Maximum		89,00

De gemiddelde leeftijd van de random steekproef is iets hoger dan die van katvangers, namelijk 51,36. Na het verwijderen van de missing waarden uit de dataset verandert het gemiddelde vrijwel niet (50,85). De meest voorkomende leeftijd is 51. Het minimum blijft in beide 16, maar het maximum verandert wel. Voor de missing waarden zijn verwijderd was de oudste persoon 122, na het verwijderen is de oudste persoon 89.

Geslacht

```
RECODE GESLACHT (1=0) (2=1).
EXECUTE.
```

GESLACHT

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	7313	73,3	74,5	74,5
	1	2497	25,0	25,5	100,0
	Total	9810	98,3	100,0	
Missing	System	171	1,7		
Total		9981	100,0		

GESLACHT

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	7190	74,4	74,4	74,4
	1	2472	25,6	25,6	100,0
	Total	9662	100,0	100,0	

Geslacht werd gemeten door 1 = man en 2 = vrouw. Voor de analyses is er besloten om deze te hercoderen door middel van 0 = man en 1 = vrouw. Voordat de missing waarden werden verwijderd was 73,3% man en 25% vrouw. 171 (1,7%) van de cases zijn missende waarden en zijn uit de dataset gehaald. De meerderheid van de katvanger zijn dus man. Na het verwijderen van de missing waarden blijft de verhouding ongeveer hetzelfde. De meerderheid blijft man (74,4%) en een kwart blijft vrouw (25,6%).

GESLACHT

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	7151	71,5	71,6	71,6
	1	2843	28,4	28,4	100,0
	Total	9994	99,9	100,0	
Missing	System	6	,1		
Total		10000	100,0		

GESLACHT

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	3825	70,7	70,7	70,7
	1	1583	29,3	29,3	100,0
	Total	5408	100,0	100,0	

De dataset bestaande uit de willekeurige steekproef heeft een soortgelijke verdeling van geslacht. De meerderheid is hier ook man (71,5%) en een kwart is ongeveer vrouw (28,4). Er zijn in totaal ,1% missing waarden. Nadat deze zijn verwijderd blijft de verdeling vrijwel hetzelfde. 70,7% is man en 29,3 is vrouw.

Nationaliteit

**RECODE ANDERS_DAN_NL_NAT (0=1) (1=0).
EXECUTE.**

De codering van variabele is gespiegeld, dit zorgt ervoor dat 0 = niet Nederlands en 1 = wel Nederlands.

ANDERS_DAN_NL_NAT					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	1370	13,7	13,7	13,7
	1	8611	86,3	86,3	100,0
Total		9981	100,0	100,0	

ANDERS_DAN_NL_NAT					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	1189	12,3	12,3	12,3
	1	8473	87,7	87,7	100,0
Total		9662	100,0	100,0	

De meerderheid van de katvangers heeft een Nederlandse nationaliteit, namelijk 87,7%. Toch heeft ook een aantal een andere nationaliteit, in totaal 12,3%. Deze verdeling was bijna hetzelfde voordat de missing waarden werden verwijderd.

ANDERS_DAN_NL_NAT					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	32	,3	,3	,3
	1	9968	99,7	99,7	100,0
Total		10000	100,0	100,0	

ANDERS_DAN_NL_NAT					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	16	,3	,3	,3
	1	5392	99,7	99,7	100,0
Total		5408	100,0	100,0	

Na het verwijderen van de missende waarden verandert er niets aan de verdeling. In de data van de willekeurige steekproef valt op dat bijna iedereen een Nederlandse nationaliteit heeft (99,7%).

Dataset voertuigkenmerken

Ook voor deze datasets zijn de missing data op dezelfde manier uit de datasets gehaald. Eerst zijn er dummy variabelen gemaakt. Vervolgens is er een nieuwe variabele gemaakt genaamd 'miss'.

Vervolgens wordt de data gefilterd door $miss < 1$.

```
RECODE WOK_STATUS APK_VERLOPEN BPM_BEDRAG BOUWJAAR AANTAL_CJIB VoertuigClassificatie VERZEKERD_IND(SYSMIS=1)
(ELSE=0) INTO DumWok DumAPK DumBPM DumBouw DumAanCJIB DumVoerCla DumVer.
EXECUTE.

COMPUTE Miss=DumWok + DumAPK + DumBPM + DumBouw + DumAanCJIB + DumVoerCla + DumVer.
EXECUTE.

DATASET ACTIVATE DataSet1.
USE ALL.
COMPUTE filter_$=(Miss < 1 & VoertuigClassificatie > 0).
VARIABLE LABELS filter_$ 'Miss < 1 & VoertuigClassificatie > 0 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

In de dataset bestaande uit de willekeurige steekproef is er ook gefilterd op aantal ID < 5. Dit zorgt ervoor dat cases met enorm veel voertuigen op naam eruit worden gefilterd. Dit zullen voornamelijk bedrijven zijn en zij zorgen ervoor dat het gemiddelde aantal voertuigen enorm omhoogging. Dit is gedaan door een geaggregeerde variabele van ID aan te maken. Deze liet zien hoe vaak een ID-nummer voorkwam en hoeveel voertuigen dus bij dat ID-nummer hoort. Vervolgens is er gefilterd door deze variabele < 5 te doen. Zie onderstaande syntax.

```
AGGREGATE
/OUTFILE=* MODE=ADDVARIABLES
/BREAK=ID
/N_BREAKID=N.

USE ALL.
COMPUTE filter_$=(Miss < 1 AND VoertuigClassificatieRandom > 0 AND N_BREAKID < 5).
VARIABLE LABELS filter_$ 'Miss < 1 AND VoertuigClassificatieRandom > 0 AND N_BREAKID < 5 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

WOK-status

Er is inhoudelijk niets veranderd aan de codering van de variabele WOK-status. De codering van de nominale variabele WOK-status is 0= geen status en 1 = wel status.

WOK_STATUS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	geen status	7045	96,5	96,5	96,5
	wel status	256	3,5	3,5	100,0
Total		7301	100,0	100,0	

WOK_STATUS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	geen status	6896	96,4	96,4	96,4
	wel status	256	3,6	3,6	100,0
Total		7152	100,0	100,0	

Na het verwijderen van de missing waarden valt op dat er aan de verdeling vrij weinig verandert. De meerderheid van de katvangers heeft geen WOK-status, namelijk 95,6%. Slechts 4,4% heeft wel een WOK-status.

WOK_STATUS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	121001	96,9	96,9	96,9
	1	3924	3,1	3,1	100,0
Total		124925	100,0	100,0	

WOK_STATUS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	13501	97,7	97,7	97,7
	1	318	2,3	2,3	100,0
Total		13819	100,0	100,0	

De aantallen van de willekeurige steekproef komen aardig overeen met die van de katvangers. Na het verwijderen van de missende waarden heeft 97,7% geen WOK-status en 2,3% wel een wokstatus.

APK

Ook wordt er inhoudelijk niets veranderd aan de variabele APK. De variabele APK is een nominale variabele en wordt gecodeerd door middel van 0 = niet verlopen en 1 = verlopen.

APK_VERLOPEN

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	niet verlopen	3072	42,1	42,1	42,1
	verlopen	4229	57,9	57,9	100,0
Total		7301	100,0	100,0	

APK_VERLOPEN

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	niet verlopen	2928	40,9	40,9	40,9
	verlopen	4224	59,1	59,1	100,0
Total		7152	100,0	100,0	

Voordat de missing waarden eruit werden gehaald had de meerderheid van de katvangers een verlopen APK, namelijk 57,9%. 42,1% had geen verlopen APK. Ook na het verwijderen van de missende waarden heet de meerderheid een verlopen APK (59,1%). 40,9% heeft geen verlopen APK.

APK_VERLOPEN

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	124925	100,0	100,0	100,0

APK_VERLOPEN

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	13819	100,0	100,0	100,0

In de dataset zonder katvangers heeft niemand een verlopen APK. 100% heeft hun voertuig APK gekeurd gekregen.

Aantal CJIB-boetes

Statistics			Statistics		
AANTAL_CJIB			AANTAL_CJIB		
N	Valid	7301	N	Valid	7152
	Missing	0		Missing	0
Mean		,20	Mean		,21
Median		,00	Median		,00
Mode		0	Mode		0
Std. Deviation		1,482	Std. Deviation		1,497
Minimum		0	Minimum		0
Maximum		67	Maximum		67

AANTAL_CJIB				
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	6800	93,1	93,1
	1	269	3,7	96,8
	2	76	1,0	97,9
	3	56	,8	98,6
	4	29	,4	99,0
	5	19	,3	99,3
	6	9	,1	99,4
	7	1	,0	99,4
	8	7	,1	99,5
	9	7	,1	99,6
	10	3	,0	99,7
	11	5	,1	99,7
	12	4	,1	99,8
	15	3	,0	99,8
	18	4	,1	99,9
	19	2	,0	99,9
	20	1	,0	99,9
	21	1	,0	99,9
	27	1	,0	99,9
	29	2	,0	100,0
	32	1	,0	100,0
	67	1	,0	100,0
Total	7301	100,0	100,0	

AANTAL_CJIB				
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	6652	93,0	93,0
	1	268	3,7	96,8
	2	76	1,1	97,8
	3	56	,8	98,6
	4	29	,4	99,0
	5	19	,3	99,3
	6	9	,1	99,4
	7	1	,0	99,4
	8	7	,1	99,5
	9	7	,1	99,6
	10	3	,0	99,7
	11	5	,1	99,7
	12	4	,1	99,8
	15	3	,0	99,8
	18	4	,1	99,9
	19	2	,0	99,9
	20	1	,0	99,9
	21	1	,0	99,9
	27	1	,0	99,9
	29	2	,0	100,0
	32	1	,0	100,0
	67	1	,0	100,0
Total	7152	100,0	100,0	

Het aantal CJIB-boetes wat een katvanger op naam heeft wordt gemeten met een continue variabele.

Het gemiddelde is ,21. De meest voorkomende waarde is 0. Het hoogste aantal is 67. Na het

verwijderen van de missing waarden is het gemiddelde hoger geworden, van ,20 naar ,21.

Statistics

AANTAL_CJIB		
N	Valid	124925
	Missing	0
Mean		,10
Median		,00
Mode		0
Std. Deviation		,511
Minimum		0
Maximum		26

Statistics

AANTAL_CJIB		
N	Valid	13819
	Missing	0
Mean		,16
Median		,00
Mode		0
Std. Deviation		,582
Minimum		0
Maximum		20

AANTAL_CJIB

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	117319	93,9	93,9	93,9
	1	5226	4,2	4,2	98,1
	2	1417	1,1	1,1	99,2
	3	462	,4	,4	99,6
	4	232	,2	,2	99,8
	5	115	,1	,1	99,9
	6	65	,1	,1	99,9
	7	28	,0	,0	100,0
	8	17	,0	,0	100,0
	9	10	,0	,0	100,0
	10	8	,0	,0	100,0
	11	4	,0	,0	100,0
	12	4	,0	,0	100,0
	13	5	,0	,0	100,0
	14	2	,0	,0	100,0
	15	4	,0	,0	100,0
	16	2	,0	,0	100,0
17	2	,0	,0	100,0	
19	1	,0	,0	100,0	
20	1	,0	,0	100,0	
26	1	,0	,0	100,0	
Total		124925	100,0	100,0	

AANTAL_CJIB

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	12271	88,8	88,8	88,8
	1	1151	8,3	8,3	97,1
	2	259	1,9	1,9	99,0
	3	75	,5	,5	99,5
	4	26	,2	,2	99,7
	5	18	,1	,1	99,9
	6	11	,1	,1	99,9
	7	1	,0	,0	99,9
	8	3	,0	,0	100,0
	9	2	,0	,0	100,0
	11	1	,0	,0	100,0
	20	1	,0	,0	100,0
	Total		13819	100,0	100,0

Het gemiddelde verandert van ,10 naar ,16 nadat de missing waarden zijn verwijderd. De meest voorkomende waarde is 0. Het maximum is 20. De standaarddeviatie is laag (.582) wat betekent dat er weinig spreiding is.

Verzekering

De codering van de variabelen verzekering is gespiegeld, waardoor 0 = niet verzekerd en 1 = verzekerd.

```
RECODE VERZEKERD_IND (0=1) (1=0).  
EXECUTE.
```

		VERZEKERD_IND			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	verzekerd	1881	25,8	25,8	25,8
	niet verzekerd	5420	74,2	74,2	100,0
	Total	7301	100,0	100,0	

		VERZEKERD_IND			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	verzekerd	1881	26,3	26,3	26,3
	niet verzekerd	5271	73,7	73,7	100,0
	Total	7152	100,0	100,0	

Na het verwijderen van de missende waarden valt op dat de meerderheid van de katvangers geen verzekerd voertuig hebben, namelijk 73,7%. 26,3% heeft het voertuig wel verzekerd.

		VERZEKERD_IND			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	51183	41,0	41,0	41,0
	1	73742	59,0	59,0	100,0
	Total	124925	100,0	100,0	

		VERZEKERD_IND			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	verzekerd	9380	67,9	67,9	67,9
	niet verzekerd	4439	32,1	32,1	100,0
	Total	13819	100,0	100,0	

In de dataset met random personen heeft de meerderheid hun voertuig wel verzekerd, namelijk 67,9%. 32,1% heeft hun voertuig niet verzekerd.

Beslag

BESLAG_IND					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	6757	92,5	92,5	92,5
	1	544	7,5	7,5	100,0
	Total	7301	100,0	100,0	

BESLAG_IND					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	6625	92,6	92,6	92,6
	1	527	7,4	7,4	100,0
	Total	7152	100,0	100,0	

Beslag wordt gecodeerd door 0 = niet in beslag en 1 = in beslag. Er is niets veranderd aan deze codering. De verdeling verandert vrij weinig na het verwijderen van de missende waarden. De meerderheid heeft geen beslag op hun voertuig (90,8%). Slechts van 9,2% van de voertuigen van katvangers ligt er beslag op het voertuig.

BESLAG_IND					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	124258	99,5	99,5	99,5
	1	667	,5	,5	100,0
	Total	124925	100,0	100,0	

BESLAG_IND					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	13776	99,7	99,7	99,7
	1	43	,3	,3	100,0
	Total	13819	100,0	100,0	

Ook hier veranderd de verdeling vrijwel niet na het verwijderen van de missende waarden. Wederom heeft de meerderheid geen beslag op hun voertuig (99,7%). Slechts ,3% heeft wel een beslag op hun voertuig.

Bijlage 2 – analyses

Verschiltoetsen achtergrondkenmerken en katvanger

Chi-kwadraattoetsen

RYB_IND * Katvanger Crosstabulation

Count

		Katvanger		Total
		niet-katvanger	katvanger	
RYB_IND	0	7	6132	6139
	1	5401	3530	8931
Total		5408	9662	15070

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	5761,243 ^a	1	,000		
Continuity Correction ^b	5758,619	1	,000		
Likelihood Ratio	7579,172	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	5760,860	1	,000		
N of Valid Cases	15070				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 2203,03.

b. Computed only for a 2x2 table

CJIB_SIGNAAL * Katvanger Crosstabulation

Count

		Katvanger		Total
		niet-katvanger	katvanger	
CJIB_SIGNAAL	0	5383	9386	14769
	1	25	276	301
Total		5408	9662	15070

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	101,542 ^a	1	<,001		
Continuity Correction ^b	100,323	1	<,001		
Likelihood Ratio	126,357	1	<,001		
Fisher's Exact Test				<,001	<,001
Linear-by-Linear Association	101,536	1	<,001		
N of Valid Cases	15070				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 108,02.

b. Computed only for a 2x2 table

ANDERS_DAN_NL_NAT * Katvanger Crosstabulation

Count

		Katvanger		Total
		niet-katvanger	katvanger	
ANDERS_DAN_NL_NAT	0	16	1189	1205
	1	5392	8473	13865
Total		5408	9662	15070

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	679,833 ^a	1	<,001		
Continuity Correction ^b	678,201	1	<,001		
Likelihood Ratio	973,487	1	<,001		
Fisher's Exact Test				<,001	<,001
Linear-by-Linear Association	679,788	1	<,001		
N of Valid Cases	15070				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 432,42.

b. Computed only for a 2x2 table

FISC_WANBETALER_SIGNAAL * Katvanger Crosstabulation

Count

		Katvanger		Total
		niet-katvanger	katvanger	
FISC_WANBETALER_SIGNAAL	0	5389	9660	15049
	1	19	2	21
Total		5408	9662	15070

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	27,238 ^a	1	<,001		
Continuity Correction ^b	24,914	1	<,001		
Likelihood Ratio	27,551	1	<,001		
Fisher's Exact Test				<,001	<,001
Linear-by-Linear Association	27,236	1	<,001		
N of Valid Cases	15070				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 7,54.

b. Computed only for a 2x2 table

GESLACHT * Katvanger Crosstabulation

Count

		Katvanger		Total
		niet-katvanger	katvanger	
GESLACHT	0	3825	7190	11015
	1	1583	2472	4055
Total		5408	9662	15070

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	23,961 ^a	1	<,001		
Continuity Correction ^b	23,774	1	<,001		
Likelihood Ratio	23,783	1	<,001		
Fisher's Exact Test				<,001	<,001
Linear-by-Linear Association	23,960	1	<,001		
N of Valid Cases	15070				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 1455,17.

b. Computed only for a 2x2 table

T-toetsen

Group Statistics

	Katvanger	N	Mean	Std. Deviation	Std. Error Mean
% van huishoudens dat onder de 40% laagste huishoudens NL	niet-katvanger	5408	36,97	10,806	,147
	katvanger	9662	44,48	11,629	,118

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	t	df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						One-Sided p	Two-Sided p			Lower	Upper
% van huishoudens dat onder de 40% laagste huishoudens NL	Equal variances assumed	68,254	<,001	-39,020	15068	,000	,000	-7,515	,193	-7,893	-7,138
	Equal variances not assumed			-39,835	11892,178	,000	,000	-7,515	,189	-7,885	-7,145

Group Statistics

	Katvanger	N	Mean	Std. Deviation	Std. Error Mean
AANTAL_OP_ADRES	niet-katvanger	5408	2,63	5,883	,080
	katvanger	9662	11,11	38,367	,390

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	t	df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						One-Sided p	Two-Sided p			Lower	Upper
AANTAL_OP_ADRES	Equal variances assumed	933,381	<,001	-16,158	15068	<,001	<,001	-8,485	,525	-9,515	-7,456
	Equal variances not assumed			-21,297	10456,724	<,001	<,001	-8,485	,398	-9,266	-7,704

Group Statistics

	Katvanger	N	Mean	Std. Deviation	Std. Error Mean
Leeftijd	niet-katvanger	5408	50,8530	16,12344	,21925
	katvanger	9662	47,3234	14,53956	,14792

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	t	df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
						One-Sided p	Two-Sided p			Lower	Upper
Leeftijd	Equal variances assumed	94,041	<,001	13,739	15068	<,001	<,001	3,52956	,25690	3,02602	4,03311
	Equal variances not assumed			13,345	10259,614	<,001	<,001	3,52956	,26448	3,01113	4,04800

Verschiltoetsen voertuigkenmerken en katvanger

Chi-kwadraattoetsen

Katvanger * WOK_STATUS Crosstabulation

Count

		WOK_STATUS		Total
		0	1	
Katvanger	,00	13501	318	13819
	1,00	6896	256	7152
Total		20397	574	20971

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	28,925 ^a	1	<,001		
Continuity Correction ^b	28,447	1	<,001		
Likelihood Ratio	27,826	1	<,001		
Fisher's Exact Test				<,001	<,001
Linear-by-Linear Association	28,923	1	<,001		
N of Valid Cases	20971				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 195,76.

b. Computed only for a 2x2 table

Katvanger * APK_VERLOPEN Crosstabulation

Count

		APK_VERLOPEN		Total
		0	1	
Katvanger	,00	13819	0	13819
	1,00	2928	4224	7152
Total		16747	4224	20971

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	10220,100 ^a	1	,000		
Continuity Correction ^b	10216,428	1	,000		
Likelihood Ratio	11391,608	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	10219,613	1	,000		
N of Valid Cases	20971				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 1440,56.

b. Computed only for a 2x2 table

Katvanger * VERZEKERD_IND Crosstabulation

Count

		VERZEKERD_IND		Total
		verzekerd	niet verzekerd	
Katvanger	,00	9380	4439	13819
	1,00	1881	5271	7152
Total		11261	9710	20971

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3276,710 ^a	1	,000		
Continuity Correction ^b	3275,038	1	,000		
Likelihood Ratio	3364,726	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	3276,554	1	,000		
N of Valid Cases	20971				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 3311,52.

b. Computed only for a 2x2 table

Katvanger * BESLAG_IND Crosstabulation

Count

		BESLAG_IND		Total
		0	1	
Katvanger	,00	13776	43	13819
	1,00	6625	527	7152
Total		20401	570	20971

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	887,742 ^a	1	<,001		
Continuity Correction ^b	885,075	1	<,001		
Likelihood Ratio	889,098	1	<,001		
Fisher's Exact Test				<,001	<,001
Linear-by-Linear Association	887,700	1	<,001		
N of Valid Cases	20971				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 194,39.

b. Computed only for a 2x2 table

Katvanger * VoertuigClassificatieRandom Crosstabulation

Count

		VoertuigClassificatieRandom					Total
		Bedrijfsauto	Bromfiets	Driewielig motorrijtuig	Motorfiets (met zijspan)	Personenauto	
Katvanger	,00	582	1471	36	45	11685	13819
	1,00	441	1620	11	132	4948	7152
Total		1023	3091	47	177	16633	20971

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	769,669 ^a	4	<,001
Likelihood Ratio	737,325	4	<,001
Linear-by-Linear Association	564,391	1	<,001
N of Valid Cases	20971		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 16,03.

T-toetsen

Group Statistics

		Katvanger	N	Mean	Std. Deviation	Std. Error Mean
BOUWJAAR	,00		13819	2007,79	7,978	,068
	1,00		7152	1997,93	10,029	,119

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	t	df	Significance One-Sided p	Significance Two-Sided p	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
										Lower	Upper
BOUWJAAR	Equal variances assumed	1123,553	<,001	77,530	20969	,000	,000	9,861	,127	9,612	10,110
	Equal variances not assumed			72,172	11939,224	,000	,000	9,861	,137	9,593	10,129

Group Statistics

		Katvanger	N	Mean	Std. Deviation	Std. Error Mean
AANTAL_CJIB	,00		13819	,16	,582	,005
	1,00		7152	,21	1,497	,018

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
		F	Sig.	t	df	Significance One-Sided p	Significance Two-Sided p	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
										Lower	Upper
AANTAL_CJIB	Equal variances assumed	54,437	<,001	-3,243	20969	<,001	,001	-,047	,014	-,075	-,019
	Equal variances not assumed			-2,554	8288,491	,005	,011	-,047	,018	-,083	-,011

Analyse 1a – logistische regressie achtergrondkenmerken

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-,024	,014	2,816	1	,093	,976

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	1272,428	3	<,001
	Block	1272,428	3	<,001
	Model	1272,428	3	<,001

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	18401,719 ^a	,081	,111

a. Estimation terminated at iteration number 7 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	64,462	8	<,001

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1 ^a	Leeftijd_C	-,019	,001	281,648	1	<,001	,981	,979	,983
	GESLACHT	-,165	,039	17,842	1	<,001	,848	,785	,915
	ANDERS_DAN_NL_NAT	-3,976	,253	247,820	1	<,001	,019	,011	,031
	Constant	4,459	,252	312,927	1	<,001	86,376		

a. Variable(s) entered on step 1: Leeftijd_C, GESLACHT, ANDERS_DAN_NL_NAT.

Block 2: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	10205,111	5	,000
	Block	10205,111	5	,000
	Model	11477,538	8	,000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	8196,608 ^a	,533	,731

a. Estimation terminated at iteration number 10 because parameter estimates changed by less than ,001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	150,558	8	<,001

Variables in the Equation

Step 1 ^a		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
	Leeftijd_C	-,028	,002	248,520	1	<,001	,973	,969	,976
	GESLACHT	,215	,059	13,376	1	<,001	1,240	1,105	1,391
	ANDERS_DAN_NL_NAT	-5,242	,257	414,790	1	<,001	,005	,003	,009
	% van huishoudens dat onder de 40% laagste huishoudens NL	,063	,002	677,214	1	<,001	1,065	1,060	1,070
	RYB_IND	-7,777	,380	418,904	1	<,001	,000	,000	,001
	CJIB_SIGNAAL	1,716	,263	42,580	1	<,001	5,561	3,322	9,310
	FISC_WANBETALER_SIGNAAL	-4,132	1,424	8,422	1	,004	,016	,001	,261
	AantalAdres_C	,058	,005	144,815	1	<,001	1,060	1,050	1,070
	Constant	9,727	,466	436,223	1	<,001	16768,864		

a. Variable(s) entered on step 1: % van huishoudens dat onder de 40% laagste huishoudens NL, RYB_IND, CJIB_SIGNAAL, FISC_WANBETALER_SIGNAAL, AantalAdres_C.

Analyse 1b – logistische regressie van geslacht en de overige variabelen

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	GESLACHT	-,185	,038	23,931	1	<,001	,831
	Constant	,631	,020	994,528	1	<,001	1,880

a. Variable(s) entered on step 1: GESLACHT.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	GESLACHT	-,218	,038	32,355	1	<,001	,804
	Leeftijd	-,016	,001	192,549	1	<,001	,984
	Constant	1,409	,060	549,113	1	<,001	4,093

a. Variable(s) entered on step 1: Leeftijd.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	GESLACHT	-,165	,039	17,842	1	<,001	,848
	Leeftijd	-,019	,001	281,648	1	<,001	,981
	ANDERS_DAN_NL_NAT	-3,976	,253	247,820	1	<,001	,019
	Constant	5,412	,260	432,187	1	<,001	224,175

a. Variable(s) entered on step 1: ANDERS_DAN_NL_NAT.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	GESLACHT	-,169	,039	18,620	1	<,001	,845
	Leeftijd	-,019	,001	284,226	1	<,001	,981
	ANDERS_DAN_NL_NAT	-4,014	,256	246,587	1	<,001	,018
	FISC_WANBETALER_SIG NAAL	-3,925	,969	16,401	1	<,001	,020
	Constant	5,461	,263	429,948	1	<,001	235,289

a. Variable(s) entered on step 1: FISC_WANBETALER_SIGNAAL.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	GESLACHT	-,156	,039	15,861	1	<,001	,855
	Leeftijd	-,018	,001	254,400	1	<,001	,982
	ANDERS_DAN_NL_NAT	-4,029	,256	247,662	1	<,001	,018
	FISC_WANBETALER_SIG NAAL	-4,682	1,026	20,829	1	<,001	,009
	CJIB_SIGNAAL	1,921	,225	72,892	1	<,001	6,830
	Constant	5,400	,264	419,062	1	<,001	221,395

a. Variable(s) entered on step 1: CJIB_SIGNAAL.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	GESLACHT	,132	,054	5,920	1	,015	1,141
	Leeftijd	-,029	,002	313,723	1	<,001	,972
	ANDERS_DAN_NL_NAT	-5,418	,256	447,497	1	<,001	,004
	FISC_WANBETALER_SIG NAAL	-3,933	1,218	10,420	1	,001	,020
	CJIB_SIGNAAL	1,947	,247	61,919	1	<,001	7,008
	RYB_IND	-7,695	,379	411,479	1	<,001	,000
	Constant	13,575	,469	839,235	1	<,001	786274,798

a. Variable(s) entered on step 1: RYB_IND.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	GESLACHT	,190	,056	11,642	1	<,001	1,209
	Leeftijd	-,030	,002	323,364	1	<,001	,970
	ANDERS_DAN_NL_NAT	-5,386	,257	440,268	1	<,001	,005
	FISC_WANBETALER_SIG NAAL	-3,798	1,209	9,863	1	,002	,022
	CJIB_SIGNAAL	1,874	,251	55,680	1	<,001	6,517
	RYB_IND	-7,740	,379	416,057	1	<,001	,000
	AANTAL_OP_ADRES	,058	,005	149,809	1	<,001	1,060
	Constant	13,403	,470	813,900	1	<,001	661716,679

a. Variable(s) entered on step 1: AANTAL_OP_ADRES.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	GESLACHT	,215	,059	13,376	1	<,001	1,240
	Leeftijd	-,028	,002	248,520	1	<,001	,973
	ANDERS_DAN_NL_NAT	-5,242	,257	414,790	1	<,001	,005
	FISC_WANBETALER_SIG NAAL	-4,132	1,424	8,422	1	,004	,016
	CJIB_SIGNAAL	1,716	,263	42,580	1	<,001	5,561
	RYB_IND	-7,777	,380	418,904	1	<,001	,000
	AANTAL_OP_ADRES	,058	,005	144,815	1	<,001	1,060
	% van huishoudens dat onder de 40% laagste huishoudens NL	,063	,002	677,214	1	<,001	1,065
	Constant	10,577	,478	489,594	1	<,001	39217,283

a. Variable(s) entered on step 1: % van huishoudens dat onder de 40% laagste huishoudens NL.

Analyse 1c – logistische regressie met alleen geslacht en rijbewijs

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	GESLACHT	-,042	,048	,790	1	,374	,958
	RYB_IND	-7,199	,379	361,148	1	<,001	,001
	Constant	6,786	,378	321,639	1	<,001	885,103

a. Variable(s) entered on step 1: RYB_IND.

Analyse 2 – logistische regressie voertuigkenmerken

Allereerst wordt de data geaggregeerd om te controleren of er sprake is van interessante verschillen. Hierop gebaseerd wordt er besloten of er een logistische regressie wordt uitgevoerd en voor welke variabelen. De syntax voor het aggregeren zien er als volgt uit:

AGGREGATE

```
/OUTFILE=* MODE=ADDVARIABLES
/BREAK=ID
/N_BREAKID=N.
```

Hieruit blijkt tussen sommige variabelen interessante verschillen te zitten waardoor er wordt besloten om een logistische regressie voor die variabelen te doen.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-,659	,015	2044,549	1	,000	,518

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	12327,868	5	,000
	Block	12327,868	5	,000
	Model	12327,868	5	,000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	14587,340 ^a	,444	,615

a. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	7,548	3	,056

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
							Lower	Upper
Step 1 ^a								
WOK_STATUS	1,093	,095	131,911	1	<,001	2,983	2,476	3,595
APK_VERLOPEN	22,592	605,196	,001	1	,970	6480199525,0	,000	.
AANTAL_CJIB	,246	,026	90,251	1	<,001	1,279	1,216	1,346
VERZEKERD_IND	,374	,044	71,514	1	<,001	1,454	1,333	1,586
BESLAG_IND	3,244	,168	371,443	1	<,001	25,629	18,427	35,644
Constant	-1,872	,029	4234,368	1	,000	,154		

a. Variable(s) entered on step 1: WOK_STATUS, APK_VERLOPEN, AANTAL_CJIB, VERZEKERD_IND, BESLAG_IND.

Bijlage 3 – assumpties en outliers

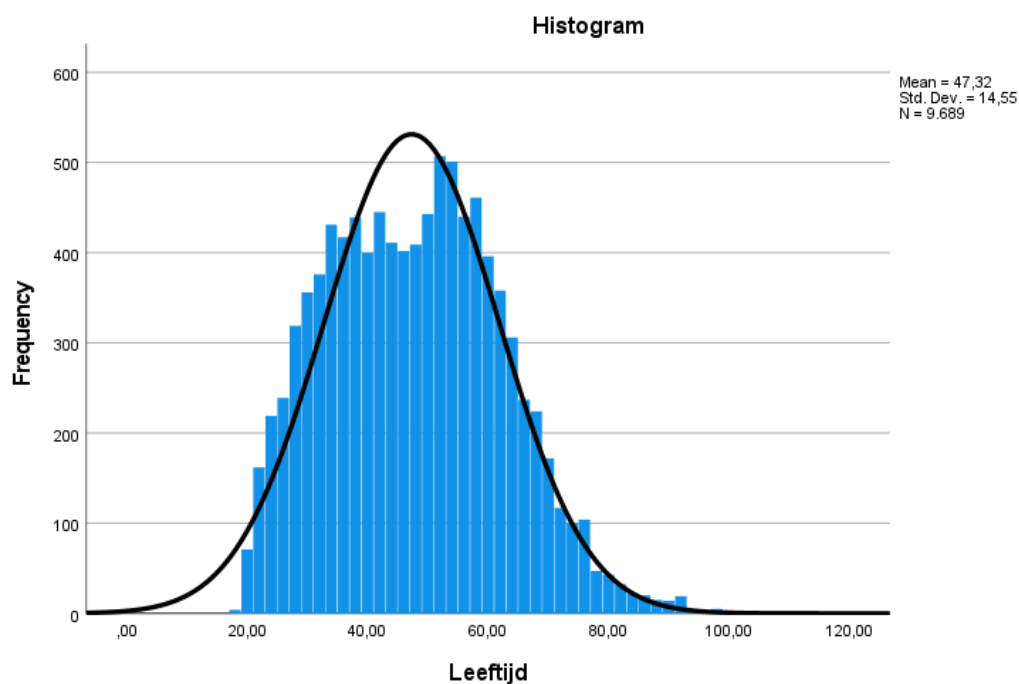
Outliers verwijderen

Uit de descriptieve statistieken bleek al gauw dat er voor de variabelen leeftijd en aantal op adres in de dataset met achtergrondkenmerken. Er is voor gekozen om drie standaarddeviaties boven en onder het gemiddelde eruit te filteren. Dit betekent dat iedereen boven de leeftijd van 90 eruit is gefilterd voor niet-katvangers en voor katvangers is iedereen ouder dan 100 eruit gefilterd.

Katvangers

Statistics

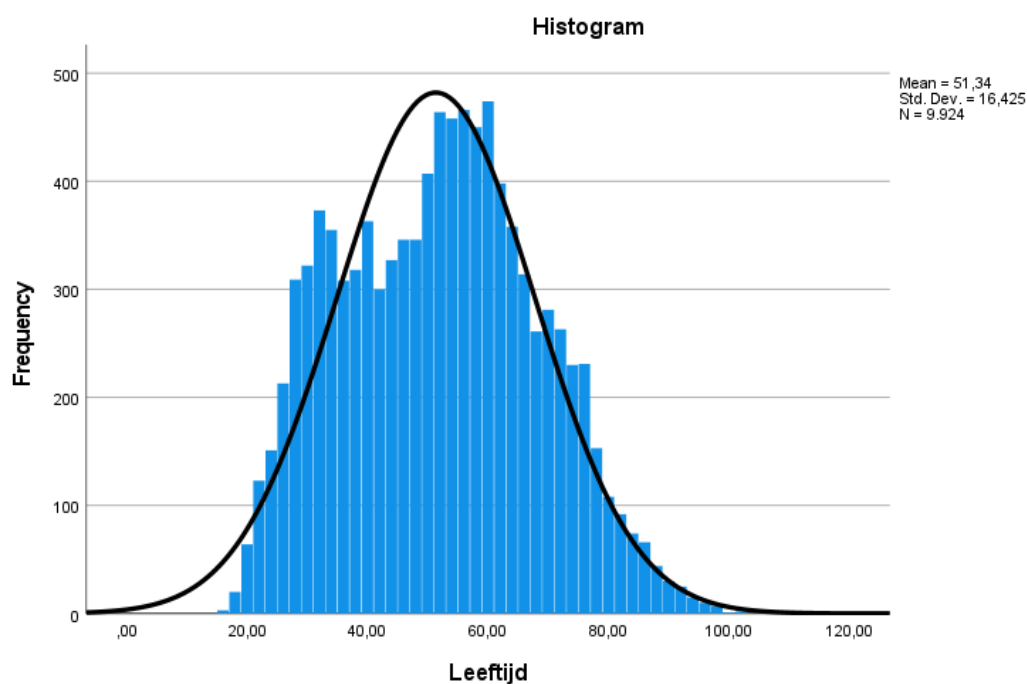
Leeftijd		
N	Valid	9689
	Missing	0
Mean		47,3157
Median		47,0000
Mode		53,00
Std. Deviation		14,54978
Minimum		18,00
Maximum		104,00



Niet-katvangers

Statistics

Leeftijd		
N	Valid	9924
	Missing	0
Mean		51,3416
Median		52,0000
Mode		59,00
Std. Deviation		16,42508
Minimum		16,00
Maximum		107,00



```
IF (Katvanger = 0) & (Leeftijd > 90) DELETE3 = 1.  
IF (Katvanger = 0) & (Leeftijd < 90) DELETE3 = 0.  
IF (Katvanger = 1) & (Leeftijd > 100) DELETE3 = 1.  
IF (Katvanger = 1) & (Leeftijd < 100) DELETE3 = 0.  
EXE.
```

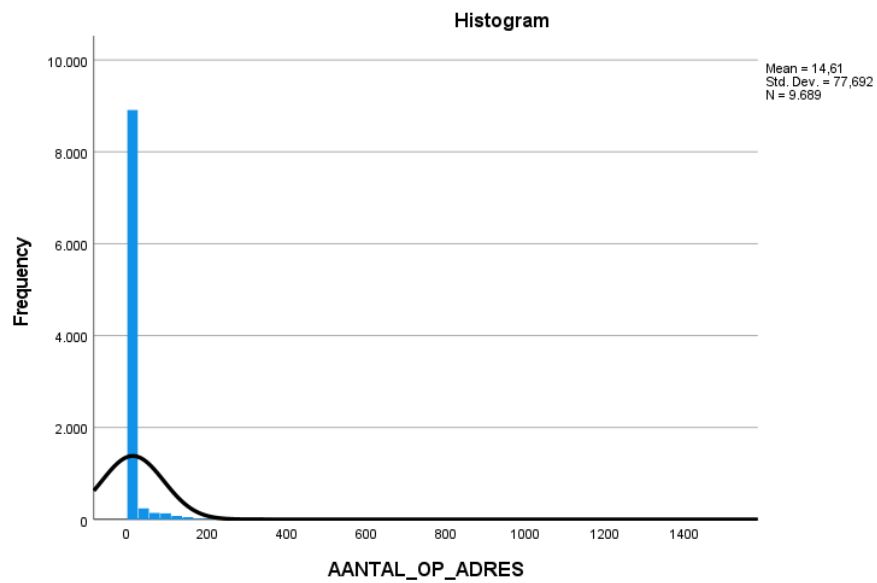
Ook geldt dit voor (niet)katvangers en de variabele 'aantal op adres'. Sommige adressen hebben 234157 personen op het adres wonen, de postcodes van deze personen onbekend zijn waardoor de waarde van 234157 automatisch aan die personen zijn toegekend tijdens het creëren van de dataset. Deze waarden zijn eruit gefilterd. Toch blijft de maximumwaarde erg hoog. Hier is er ook voor gekozen om drie standaarddeviaties boven en onder het gemiddelde eruit te filteren.

Katvangers

Statistics

AANTAL_OP_ADRES

N	Valid	9689
	Missing	0
Mean		14,61
Median		2,00
Mode		1
Std. Deviation		77,692
Minimum		1
Maximum		1456

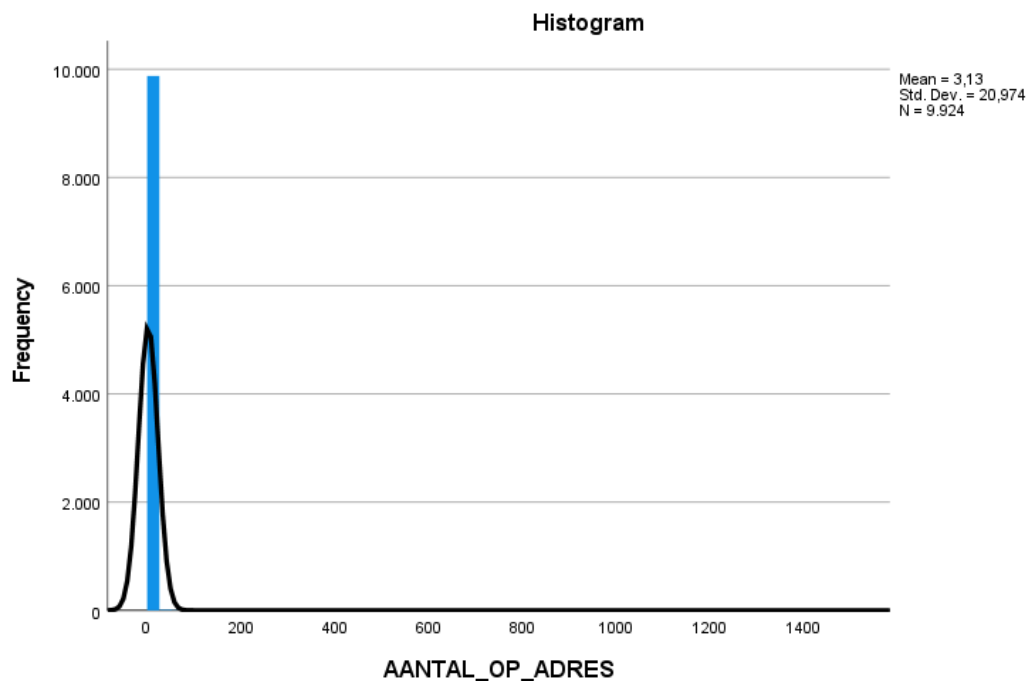


Niet-katvangers

Statistics

AANTAL_OP_ADRES

N	Valid	9924
	Missing	0
Mean		3,13
Median		2,00
Mode		2
Std. Deviation		20,974
Minimum		1
Maximum		1454



```
IF (Katvanger = 0) & (AANTAL_OP_ADRES > 247) DELETE2 = 1.
IF (Katvanger = 0) & (AANTAL_OP_ADRES < 247) DELETE2 = 0.
IF (Katvanger = 1) & (AANTAL_OP_ADRES > 1000) DELETE2 = 1.
IF (Katvanger = 1) & (AANTAL_OP_ADRES < 1000) DELETE2 = 0.
EXE.
```

Tot slot is er ook in de dataset met niet-katvangers gefilterd. Sommige cases hadden wel honderden voertuigen op naam staan. Het vermoeden is dat dit ging om bedrijven. Deze zijn eruit gefilterd omdat zij de analyse sterk beïnvloedde. Dit is gedaan door te filteren zodat iedereen minder dan 5 voertuigen op naam heeft staan. De syntax laat zien hoe dit is gedaan.

```
IF (N_BREAK < 5) & (Katvanger = 0) DELETE = 0.
IF (N_BREAK >= 5) & (Katvanger = 0) DELETE = 1.
IF (Katvanger = 1) DELETE = 0.
EXE.
```

Er is eerst een break variabele aangemaakt. Deze liet zien hoeveel een ID gemiddeld voorkwam in de dataset aangezien dit liet zien hoe veel auto's een persoon had. Wanneer dit meer dan 5 was, zijn deze eruit gefilterd. Voor katvangers is dit niet gedaan, aangezien het bij hen logischer is als zij meerdere voertuigen op naam hebben staan.

Logistische regressie achtergrondkenmerken

Assumpties

Correlations

		GESLACHT	% van huishoudens dat onder de 40% laagste huishoudens NL	AANTAL_OP_ADRES	RYB_IND	CJIB_SIGNAAL	ANDERS_DAN_NL_NAT	FISC_WANBETALER_SIGNAAL	Leeftijd	Katvanger	
Spearman's rho	GESLACHT	Correlation Coefficient	1,000	-.032**	-.061**	.055**	-.028**	.060**	-.015	-.062**	-.040**
		Sig. (2-tailed)	.	<.001	<.001	<.001	<.001	<.001	.072	<.001	<.001
		N	15070	15070	15070	15070	15070	15070	15070	15070	15070
% van huishoudens dat onder de 40% laagste huishoudens NL		Correlation Coefficient	-.032**	1,000	-.093**	-.149**	.035**	-.107**	.000	-.055**	.311**
		Sig. (2-tailed)	<.001	.	<.001	<.001	<.001	<.001	.970	<.001	.000
		N	15070	15070	15070	15070	15070	15070	15070	15070	15070
AANTAL_OP_ADRES		Correlation Coefficient	-.061**	-.093**	1,000	.095**	.052**	-.019*	.002	.008	-.069**
		Sig. (2-tailed)	<.001	<.001	.	<.001	<.001	.017	.818	.314	<.001
		N	15070	15070	15070	15070	15070	15070	15070	15070	15070
RYB_IND		Correlation Coefficient	.055**	-.149**	.095**	1,000	-.065**	-.223**	.031**	.053**	-.618**
		Sig. (2-tailed)	<.001	<.001	<.001	.	<.001	<.001	<.001	<.001	.000
		N	15070	15070	15070	15070	15070	15070	15070	15070	15070
CJIB_SIGNAAL		Correlation Coefficient	-.028**	.035**	.052**	-.065**	1,000	.028**	.058**	-.097**	.082**
		Sig. (2-tailed)	<.001	<.001	<.001	<.001	.	<.001	<.001	<.001	<.001
		N	15070	15070	15070	15070	15070	15070	15070	15070	15070
ANDERS_DAN_NL_NAT		Correlation Coefficient	.060**	-.107**	-.019*	-.223**	.028**	1,000	-.002	-.120**	-.212**
		Sig. (2-tailed)	<.001	<.001	.017	<.001	<.001	.	.796	<.001	<.001
		N	15070	15070	15070	15070	15070	15070	15070	15070	15070
FISC_WANBETALER_SIGNAAL		Correlation Coefficient	-.015	.000	.002	.031**	.058**	-.002	1,000	-.012	-.043**
		Sig. (2-tailed)	.072	.970	.818	<.001	<.001	.796	.	.141	<.001
		N	15070	15070	15070	15070	15070	15070	15070	15070	15070
Leeftijd		Correlation Coefficient	-.062**	-.055**	.008	.053**	-.097**	-.120**	-.012	1,000	-.106**
		Sig. (2-tailed)	<.001	<.001	.314	<.001	<.001	<.001	.141	.	<.001
		N	15070	15070	15070	15070	15070	15070	15070	15070	15070
Katvanger		Correlation Coefficient	-.040**	.311**	-.069**	-.618**	.082**	-.212**	-.043**	-.106**	1,000
		Sig. (2-tailed)	<.001	.000	<.001	.000	<.001	<.001	<.001	<.001	.
		N	15070	15070	15070	15070	15070	15070	15070	15070	15070

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Correlations

		GESLACHT	% van huishoudens dat onder de 40% laagste huishoudens NL	AANTAL_OP_ADRES	RYB_IND	CJIB_SIGNAAL	ANDERS_DAN_NL_NAT	FISC_WANBETALER_SIGNAAL	Leeftijd	Katvanger	
GESLACHT	Pearson Correlation	1	-.031**	-.047**	.055**	-.028**	.060**	-.015	-.055**	-.040**	
		Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	.072	<.001	<.001	
		N	15070	15070	15070	15070	15070	15070	15070	15070	
% van huishoudens dat onder de 40% laagste huishoudens NL	Pearson Correlation	-.031**	1	-.041**	-.144**	.034**	-.103**	.000	-.055**	.303**	
		Sig. (2-tailed)	<.001	.	<.001	<.001	<.001	.992	<.001	.000	
		N	15070	15070	15070	15070	15070	15070	15070	15070	
AANTAL_OP_ADRES	Pearson Correlation	-.047**	-.041**	1	-.057**	.027**	-.064**	-.007	.051**	.131**	
		Sig. (2-tailed)	<.001	<.001	.	<.001	.001	<.001	.407	<.001	<.001
		N	15070	15070	15070	15070	15070	15070	15070	15070	
RYB_IND	Pearson Correlation	.055**	-.144**	-.057**	1	-.065**	-.223**	.031**	.056**	-.618**	
		Sig. (2-tailed)	<.001	<.001	<.001	.	<.001	<.001	<.001	<.001	.000
		N	15070	15070	15070	15070	15070	15070	15070	15070	
CJIB_SIGNAAL	Pearson Correlation	-.028**	.034**	.027**	-.065**	1	.028**	.058**	-.095**	.082**	
		Sig. (2-tailed)	<.001	<.001	.001	<.001	.	<.001	<.001	<.001	
		N	15070	15070	15070	15070	15070	15070	15070	15070	
ANDERS_DAN_NL_NAT	Pearson Correlation	.060**	-.103**	-.064**	-.223**	.028**	1	-.002	-.111**	-.212**	
		Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	.	.796	<.001	<.001
		N	15070	15070	15070	15070	15070	15070	15070	15070	
FISC_WANBETALER_SIGNAAL	Pearson Correlation	-.015	.000	-.007	.031**	.058**	-.002	1	-.012	-.043**	
		Sig. (2-tailed)	.072	.992	.407	<.001	<.001	.796	.	.150	<.001
		N	15070	15070	15070	15070	15070	15070	15070	15070	
Leeftijd	Pearson Correlation	-.055**	-.055**	.051**	.056**	-.095**	-.111**	-.012	1	-.111**	
		Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	.150	.	<.001
		N	15070	15070	15070	15070	15070	15070	15070	15070	
Katvanger	Pearson Correlation	-.040**	.303**	.131**	-.618**	.082**	-.212**	-.043**	-.111**	1	
		Sig. (2-tailed)	<.001	.000	<.001	.000	<.001	<.001	<.001	<.001	.
		N	15070	15070	15070	15070	15070	15070	15070	15070	

** Correlation is significant at the 0.01 level (2-tailed).

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	1,230	,019		64,138	,000		
	GESLACHT	-,037	,009	-,034	-4,303	<,001	,994	1,006
	ANDERS_DAN_NL_NAT	-,399	,014	-,226	-28,424	<,001	,985	1,015
	Leeftijd	-,004	,000	-,138	-17,402	<,001	,985	1,015
2	(Constant)	1,449	,018		78,614	,000		
	GESLACHT	,023	,006	,021	3,908	<,001	,986	1,014
	ANDERS_DAN_NL_NAT	-,620	,010	-,350	-61,789	,000	,911	1,098
	Leeftijd	-,003	,000	-,104	-18,963	<,001	,970	1,031
	% van huishoudens dat onder de 40% laagste huishoudens NL	,007	,000	,169	30,541	<,001	,952	1,050
	AANTAL_OP_ADRES	,001	,000	,083	15,202	<,001	,983	1,017
	RYB_IND	-,644	,006	-,660	-116,179	,000	,908	1,101
	CJIB_SIGNAAL	,114	,019	,033	6,068	<,001	,982	1,019
	FISC_WANBETALER_SIG NAAL	-,323	,070	-,025	-4,624	<,001	,995	1,005

a. Dependent Variable: Katvanger

Allereerst wordt er gecontroleerd of er sprake is van een te sterke samenhang (multicollineariteit) tussen de onafhankelijke variabelen. Zowel bij de Pearsons als Spearman's correlatie is er geen sprake van een te sterke samenhang. Geen van de correlaties is groter dan -,7 of ,7.

Er wordt echter ook gekeken naar de VIF-waardes. Een VIF-waarde boven de 10 betekent dat de variabele een te sterke samenhang heeft. De VIF-waardes liggen gemiddeld rond de 1, wat betekent dat er geen sprake is van een te sterke samenhang. Deze assumptie wordt dus niet geschonden.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	GESLACHT	,204	,059	12,158	1	<,001	1,227
	RYB_IND	-7,751	,380	416,314	1	<,001	,000
	CJIB_SIGNAAL	1,741	,262	44,232	1	<,001	5,703
	ANDERS_DAN_NL_NAT	-5,233	,257	413,642	1	<,001	,005
	FISC_WANBETALER_SIG NAAL	-4,150	1,429	8,434	1	,004	,016
	LNpercentage by % van huishoudens dat onder de 40% laagste huishoudens NL	,013	,001	678,382	1	<,001	1,013
	AANTAL_OP_ADRES by LNAantalAdres	,013	,001	113,607	1	<,001	1,013
	LNLeeftijd by Leeftijd	-,006	,000	242,574	1	<,001	,994
	Constant	10,944	,471	539,139	1	<,001	56637,816

a. Variable(s) entered on step 1: GESLACHT, RYB_IND, CJIB_SIGNAAL, ANDERS_DAN_NL_NAT, FISC_WANBETALER_SIGNAAL, LNpercentage * % van huishoudens dat onder de 40% laagste huishoudens NL, AANTAL_OP_ADRES * LNAantalAdres, LNLeeftijd * Leeftijd.

COMPUTE LNLeeftijd=LN(Leeftijd).

EXECUTE.

COMPUTE LNpercentage=LN(PercentageHuishoudens).

EXECUTE.

COMPUTE LNAantalAdres=LN(AANTAL_OP_ADRES).

EXECUTE.

Om te controleren of er sprake is van lineariteit tussen de continue variabelen en de log-odds worden er drie nieuwe variabelen gemaakt (zie syntax). Vervolgens wordt er een lineaire regressie gedaan. Hierin worden drie interactie variabelen gemaakt. De eerste is tussen leeftijd en de nieuwe variabele LNleeftijd. De tweede is percentage huishoudens met LNpercentage. De derde is aantal op adres met LNAantalAdres. Wanneer de interactieterm een significante *p*-waarde heeft (<,001), dan betekent dat dat er sprake is van lineariteit tussen die continue variabele en de log-odds. Zowel de interactie van leeftijd, van percentage huishoudens en aantal op adres zijn significant, wat betekent dat er sprake is van lineariteit tussen de log-odds. Deze assumptie wordt dus geschonden.

Outliers

Allereest wordt er gekeken naar de gestandaardiseerde residuen. Alle waarden onder -3 en boven 3 zijn outliers. Er zijn heel wat outliers, de grootste waarden zijn -1,26579 en -1,13528. Het gaat om ID-nummers 5492 en 310.

Ten tweede wordt er gekeken naar de Cook's Distance (CD). De cutoff value van de Cook's Distance wordt berekend door $\frac{4}{n} = \frac{4}{15070} = 0,000265$. Gebaseerd op de CD zijn er heel wat outliers, maar

een reden hiervoor kan zijn omdat de N heel groot is en hierdoor de vuistregel heel klein wordt. Er worden ook hier besloten geen outliers te verwijderen omdat er geen opvallende waarden blijken te zijn voor deze cases.

Tot slot wordt er gekeken naar de Mahalanobis Distance. Gebaseerd op de Chi-kwadraatverdeling kijkt deze maat welke variabelen outliers zijn. Er is eerst een lineaire regressie gedaan en hierin is de Mahalanobis Distance aangevinkt. Vervolgens is er een nieuwe variabele gecreëerd (zie syntax). Wanneer een waarde onder de ,0001 ligt, wordt deze gezien als outlier. Ook hier lijken er weer een groot aantal te zijn, maar er wordt wederom voor gekozen om ze niet te verwijderen. Ze lijken niet de analyse negatief te beïnvloeden en de cases hebben ook geen opvallende waarden.

```
COMPUTE Probability=1-CDF.CHISQ(MAH_1,12).  
EXECUTE.
```

Logistische regressie voertuigkenmerken

Outliers verwijderen

Tot slot is er ook in de dataset met niet-katvangers gefilterd. Sommige cases hadden wel honderden voertuigen op naam staan. Het vermoeden is dat dit ging om bedrijven. Deze zijn eruit gefilterd omdat zij de analyse sterk beïnvloedde. Dit is gedaan door te filteren zodat iedereen minder dan 5 voertuigen op naam heeft staan. De syntax laat zien dat er eerst een break variabele is gemaakt die liet zien hoe vaak een ID-nummer voorkwam. Hoe vaak een ID-nummer voorkomt laat zien hoeveel voertuigen behoren tot dat ID-nummer. Vervolgens zijn de ID-nummers die vaker dan vijf keer voorkwamen eruit gefilterd.

Statistics

N_BREAKVoertuig

N	Valid	73634
	Missing	0
Mean		111,23
Median		27,00
Mode		5
Std. Deviation		181,448
Minimum		5
Maximum		892

Het gemiddelde aantal voertuigen op naam van de niet-katvangers met meer dan vijf voertuigen op naam was 111.23.

```
IF (Katvanger = 1) ID2=ID + 10000.  
EXECUTE.
```

```
IF (SYSMIS(ID2)) ID2 = ID.  
EXE.
```

```
IF (N_BREAK < 5) & (Katvanger = 0) DELETE =0.  
IF (N_BREAK >= 5) & (Katvanger = 0) DELETE =1.  
IF (Katvanger = 1) DELETE =0.  
EXE.
```


Assumpties

		Correlations							
		WOK_STATUS	APK_VERLOP EN	BOUWJAAR	AANTAL_CJIB	VERZEKERD_I ND	BESLAG_IND	VoertuigClassif icatieRandom	Katvanger
WOK_STATUS	Pearson Correlation	1	-.045**	.047**	.000	.053**	.008	-.128**	.037**
	Sig. (2-tailed)		<.001	<.001	.967	<.001	.252	<.001	<.001
	N	20971	20971	20971	20971	20971	20971	20971	20971
APK_VERLOPEN	Pearson Correlation	-.045**	1	-.591**	-.052**	.497**	.124**	.116**	.698**
	Sig. (2-tailed)	<.001		.000	<.001	.000	<.001	<.001	.000
	N	20971	20971	20971	20971	20971	20971	20971	20971
BOUWJAAR	Pearson Correlation	.047**	-.591**	1	.072**	-.533**	-.029**	-.085**	-.472**
	Sig. (2-tailed)	<.001	.000		<.001	.000	<.001	<.001	.000
	N	20971	20971	20971	20971	20971	20971	20971	20971
AANTAL_CJIB	Pearson Correlation	.000	-.052**	.072**	1	-.109**	.081**	.053**	.022**
	Sig. (2-tailed)	.967	<.001	<.001		<.001	<.001	<.001	.001
	N	20971	20971	20971	20971	20971	20971	20971	20971
VERZEKERD_IND	Pearson Correlation	.053**	.497**	-.533**	-.109**	1	.094**	-.033**	.395**
	Sig. (2-tailed)	<.001	.000	.000	<.001		<.001	<.001	.000
	N	20971	20971	20971	20971	20971	20971	20971	20971
BESLAG_IND	Pearson Correlation	.008	.124**	-.029**	.081**	.094**	1	-.046**	.206**
	Sig. (2-tailed)	.252	<.001	<.001	<.001	<.001		<.001	<.001
	N	20971	20971	20971	20971	20971	20971	20971	20971
VoertuigClassificatieRandom	Pearson Correlation	-.128**	.116**	-.085**	.053**	-.033**	-.046**	1	-.164**
	Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001		<.001
	N	20971	20971	20971	20971	20971	20971	20971	20971
Katvanger	Pearson Correlation	.037**	.698**	-.472**	.022**	.395**	.206**	-.164**	1
	Sig. (2-tailed)	<.001	.000	.000	.001	.000	<.001	<.001	
	N	20971	20971	20971	20971	20971	20971	20971	20971

** . Correlation is significant at the 0.01 level (2-tailed).

		Correlations								
		WOK_STATUS	APK_VERLOP EN	BOUWJAAR	AANTAL_CJIB	VERZEKERD_I ND	BESLAG_IND	VoertuigClassif icatieRandom	Katvanger	
Spearman's rho	WOK_STATUS	Correlation Coefficient	1,000	-.045**	.046**	-.009	.053**	.008	-.134**	.037**
		Sig. (2-tailed)	.	<.001	<.001	.171	<.001	.252	<.001	<.001
		N	20971	20971	20971	20971	20971	20971	20971	20971
APK_VERLOPEN	WOK_STATUS	Correlation Coefficient	-.045**	1,000	-.571**	-.129**	.497**	.124**	.131**	.698**
		Sig. (2-tailed)	<.001	.	.000	<.001	.000	<.001	<.001	.000
		N	20971	20971	20971	20971	20971	20971	20971	20971
BOUWJAAR	WOK_STATUS	Correlation Coefficient	.046**	-.571**	1,000	.165**	-.584**	-.041**	-.106**	-.463**
		Sig. (2-tailed)	<.001	.000	.	<.001	.000	<.001	<.001	.000
		N	20971	20971	20971	20971	20971	20971	20971	20971
AANTAL_CJIB	WOK_STATUS	Correlation Coefficient	-.009	-.129**	.165**	1,000	-.244**	.031**	.109**	-.064**
		Sig. (2-tailed)	.171	<.001	<.001	.	<.001	<.001	<.001	<.001
		N	20971	20971	20971	20971	20971	20971	20971	20971
VERZEKERD_IND	WOK_STATUS	Correlation Coefficient	.053**	.497**	-.584**	-.244**	1,000	.094**	-.032**	.395**
		Sig. (2-tailed)	<.001	.000	.000	<.001	.	<.001	<.001	.000
		N	20971	20971	20971	20971	20971	20971	20971	20971
BESLAG_IND	WOK_STATUS	Correlation Coefficient	.008	.124**	-.041**	.031**	.094**	1,000	-.049**	.206**
		Sig. (2-tailed)	.252	<.001	<.001	<.001	<.001	.	<.001	<.001
		N	20971	20971	20971	20971	20971	20971	20971	20971
VoertuigClassificatieRandom	WOK_STATUS	Correlation Coefficient	-.134**	.131**	-.106**	.109**	-.032**	-.049**	1,000	-.174**
		Sig. (2-tailed)	<.001	<.001	<.001	<.001	<.001	<.001	.	<.001
		N	20971	20971	20971	20971	20971	20971	20971	20971
Katvanger	WOK_STATUS	Correlation Coefficient	.037**	.698**	-.463**	-.064**	.395**	.206**	-.174**	1,000
		Sig. (2-tailed)	<.001	.000	.000	<.001	.000	<.001	<.001	.
		N	20971	20971	20971	20971	20971	20971	20971	20971

** . Correlation is significant at the 0.01 level (2-tailed).

Coefficients^a

Model		Unstandardized Coefficients		Standardized	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	9,264	,612		15,140	<,001		
	WOK_STATUS	,195	,014	,067	13,851	<,001	,987	1,013
	APK_VERLOPEN	,732	,007	,619	99,499	,000	,596	1,679
	BOUWJAAR	-,005	,000	-,095	-14,913	<,001	,570	1,754
	AANTAL_CJIB	,026	,002	,055	11,338	<,001	,980	1,021
	VERZEKERD_IND	,027	,006	,028	4,752	<,001	,651	1,537
	BESLAG_IND	,346	,014	,119	24,355	<,001	,970	1,031

a. Dependent Variable: Katvanger

REGRESSION

```

/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA COLLIN TOL
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT Katvanger
/METHOD=ENTER WOK_STATUS APK_VERLOPEN BOUWJAAR AANTAL_CJIB VERZEKERD_IND BESLAG_IND
  VoertuigClassificatie
/SAVE MAHAL COOK LEVER DFBETA.

```

Allereerst wordt er gekeken of er niet een te sterke samenhang is tussen de onafhankelijke variabelen in de dataset. Wanneer er wordt gecontroleerd voor zowel de Pearson's als de Spearman's correlatie valt op dat geen van de correlatie groter dan $-.7$ of $.7$ is. Wanneer dit wel het geval zou zijn zou er sprake zijn van een sterke samenhang. Ook wordt er hiervoor gekeken naar de VIF-waardes. Een VIF-waarde boven de 10 betekent dat er sprake is van een een te sterke samenhang. Soms wordt ook de waarde 5 hiervoor al gehanteerd. Kijkend naar de VIF-waardes van de uitgevoerde lineaire regressie valt op dat alle VIF-waardes rond de 1 zijn. Dit betekent dat er geen sprake is van multicollineariteit. Deze assumptie wordt niet geschonden.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	WOK_STATUS	,738	,340	4,697	1	,030	2,091
	APK_VERLOPEN	21,993	3779,409	,000	1	,995	3561026538,8
	VERZEKERD_IND	,683	,206	10,989	1	<,001	1,980
	BESLAG_IND	2,836	,359	62,246	1	<,001	17,049
	AANTAL_CJIB by LNaantalCJIB	,128	,019	46,678	1	<,001	1,137
	Constant	-1,717	,069	626,213	1	<,001	,180

a. Variable(s) entered on step 1: WOK_STATUS, APK_VERLOPEN, VERZEKERD_IND, BESLAG_IND, AANTAL_CJIB * LNaantalCJIB .

```

COMPUTE LNaantalCJIB=LN(AANTAL_CJIB).
EXECUTE.

```

Om te controleren of er sprake is van lineariteit tussen de continue variabelen en de log-odds worden er een nieuwe variabele gemaakt (zie syntax). Vervolgens wordt er een lineaire regressie gedaan.

Hierin wordt een interactie variabelen gemaakt tussen ‘aantal CJIB’ met de nieuwe variabele LNaantalCJIB. Deze interactie heeft een significante p -waarde ($<,001$). Dit betekent dat er wel sprake is van lineariteit tussen de onafhankelijke variabelen en de log-odds. Deze assumptie wordt geschonden.

Outliers

Allereest wordt er gekeken naar de gestandaardiseerde residuen. Alle waardes onder -3 en boven 3 zijn outliers. Er zijn in totaal 8 outliers volgens de gestandaardiseerde residuen. De grootste heeft een ZRE_1 van -6,34906. Het gaat hier om individu nummer 5383. Een reden hiervoor kan zijn dat deze persoon best wel wat boetes op naam heeft, namelijk 5. Gezien de schaal is dit best veel. Toch worden deze outliers niet verwijderd omdat zij niet opvallend afwijkend scores. Hun scores kunnen inhoudelijk veel betekenen voor dit onderzoek.

Ten tweede wordt er gekeken naar de Cook’s Distance (CD). De cutoff value van de Cook’s Distance wordt berekend door $\frac{4}{n} = \frac{4}{20971} = 0,0001908$. Gebaseerd op de CD zijn er heel wat outliers, maar een reden hiervoor kan zijn omdat de N heel groot is en hierdoor de vuistregel heel klein wordt. Er worden ook hier besloten geen outliers te verwijderen omdat er geen opvallende waarden blijken te zijn.

Tot slot wordt er gekeken naar de Mahalanobis Distance. Gebaseerd op de Chi-kwadratverdeling kijkt deze maat welke variabelen outliers zijn. Er is eerst een lineaire regressie gedaan en hierin is de Mahalanobis Distance aangevinkt. Vervolgens is er een nieuwe variabele gecreëerd (zie syntax). Wanneer een waarde onder de ,0001 ligt, wordt deze gezien als outlier. Ook hier lijken er weer een aantal te zijn, maar er wordt wederom voor gekozen om ze niet te verwijderen. Ze lijken niet de analyse negatief te beïnvloeden en lijken ook geen opvallende waarden te hebben.

```
COMPUTE Probability=1-CDF.CHISQ(MAH_1,6).  
EXECUTE.
```