

Datavisualisatie report

Data

Open data verkeersongelukken met lichamelijk letsel:

https://statbel.fgov.be/sites/default/files/files/opendata/Accident_XY/OPENDATA_MAP_2017-2022.xlsx

Variabelen metadata:

https://statbel.fgov.be/sites/default/files/files/opendata/Accident_XY/VAR_LIST_ACC_OPEN_DATA_MAP.xlsx

Refnis code¹:

https://statbel.fgov.be/sites/default/files/files/opendata/REFNIS%20code/TU_COM_REFNIS.xlsx

Info	Labels in excel	Mogelijke waarden	
Datum: jaar, maand, uur	DT_YEAR_COLLISION, DT_MONTH_COLLISION, DT_TIME		
NIS	CD_NIS		
Gewest	TX_RGN_COLLISION_NL, TX_RGN_COLLISION_FR		
Provincie	TX_PROV_COLLISION_NL, TX_PROV_COLLISION_FR		
Gemeente	TX_MUNTY_COLLISION_NL, TX_MUNTY_COLLISION_FR		
Geolocatie (x, y)	MS_X_COORD, MS_Y_COORD		
Beschrijving van kruispunt	CD_CROSSWAY, TX_CROSSWAY_NL, TX_CROSSWAY_FR	1	Op Kruispunt
		2	Buiten Kruispunt
		9	onbekend

¹ NIS-code van gewesten, provincies, bestuurlijke arrondissementen en gemeenten

Info	Labels in excel	Mogelijke waarden	
Beschrijving van weersomstandigheden	CD_WEATHER, TX_WEATHER_NL, TX_WEATHER_FR	1	Normaal
		2	Regenval
		3	Mist(zichtbaarheid minder dan 100m)
		4	Sterke wind, rukwind
		5	Sneeuwval
		6	Hagelbui
		7	Andere (dikke rook, ...)
		9	Onbekend
Beschrijving van wegconditie	CD_ROAD_CONDITION, TX_ROAD_CONDITION_NL, TX_ROAD_CONDITION_FR	1	Droog
		2	Nat, plassen
		3	Ijzel, sneeuw
		4	Proper
		5	Vuil (zand, grint, bladeren, ...)
		9	Onbekend
Bebouwde kom?	CD_BUILD_UP_AREA, TX_BUILD_UP_AREA_NL, TX_BUILD_UP_AREA_FR	1	binnen bebouwde kom
		2	buiten bebouwde kom
		9	onbekend
Lichtgesteldheid	CD_LIGHT_CONDITION, TX_LIGHT_CONDITION_NL, TX_LIGHT_CONDITION_FR	1	Dag
		2	Dageraad - schemering
		3	Nacht, openbare verlichting aanwezig en ontstoken
		4	Nacht, geen openbare verlichting aanwezig
		5	Nacht, openb. verlicht. aanwe., maar niet ontstoken
		9	Onbekend
Type weg	CD_ROAD_TYPE, CD_ROAD_TYPE_NL, CD_ROAD_TYPE_FR	1	Autosnelweg
		2	Gemeenteweg
		3	Gewestweg
		9	onbekend

Info	Labels in excel	Mogelijke waarden	
Type accident	CD_CLASS_ACCIDENTS, TX_CLASS_ACCIDENTS_NL, TX_CLASS_ACCIDENTS_FR	1	Met doden
		2	dodelijke gewonden
		3	met zwaargewonden
		4	met lichtgewonden
Type weggebruiker (zowel type1 als type2)	CD_ROAD_USR_TYPE1, TX_ROAD_USR_TYPE1_NL, TX_ROAD_USR_TYPE1_FR, CD_ROAD_USR_TYPE2, TX_ROAD_USR_TYPE2_NL, TX_ROAD_USR_TYPE2_FR	01	Autobus
		02	Andere weggebruiker
		03	Fiets
		04	Vrachtwagen
		05	Lichte vrachtauto
		06	Bromfiets
		07	Motorfiets
		08	Voetganger
		09	personenwagen
		10	één bestuurder, geen hindernis
		11	een weggebruiker tegen een hindernis
		99	onbekend
Type aanrijding	CD_COLLISION_TYPE, TX_COLLISION_TYPE_NL, TX_COLLISION_TYPE_FR	0	Geen vermelding van type aanrijding
		1	Kettingbotsin (3 bestuurders of meer)
		2	Tussen 2 bestuurders: Frontale botsing
		3	Tussen 2 bestuurders: Langs achteren
		4	Tss. 2 best.: langs opzij (voor-/achterkant-flank)
		5	Met een voetganger
		6	Weggebruiker tegen hindernis op de rijbaan
		7	Weggebruiker tegen hindernis buiten rijbaan
		8	Eén bestuurder, geen hindernis (incl.val)

Info	Labels in excel	Mogelijke waarden	
		9	Andere of onbekend
		A	Tussen 2 bestuurders Flank tegen flank
Obstakels?	CD_OBSTACLES, TX_OBSTACLES_NL, TX_OBSTACLES_FR	-1	Geen vermelding van type aanrijding
		-3	Geen vermelding van de aanwezigheid van een hindernis
		0	Geen hindernis
		50	Op de rijbaan: Dier (huisdier)
		51	Trein
		53	Op de rijbaan: Lading
		54	Op de rijbaan: Container
		55	Op de rijbaan: Werken (signal., aarde, toestel, ...)
		56	Buiten de rijbaan: Verkeerseiland, middenberm
		57	Verkeersdrempel, ezelsrug, put, uitholling, straatgoot
		58	Tramspoor
		59	Buiten de rijbaan: Boom
		60	Buiten de rijbaan: Verlichtingspaal
		61	Buiten de rijbaan: Andere paal
		62	Buiten de rijbaan: Vangrails overschreden
		63	Buiten de rijbaan: Vangrails niet overgestoken
		64	Buiten de rijbaan: Muur gebouw
		65	Buiten de rijbaan: Omheining
		66	Buiten de rijbaan: Gracht
		67	Andere hindenis
		68	Op de rijbaan: Dier (wild dier)
		69	Buiten de rijbaan: Dier
		70	Op de rijbaan: Verkeersdrempel, ezelsrug

Info	Labels in excel	Mogelijke waarden	
		71	Op de rijbaan: Put, straatgoot, tramspoor
		99	Onbekend

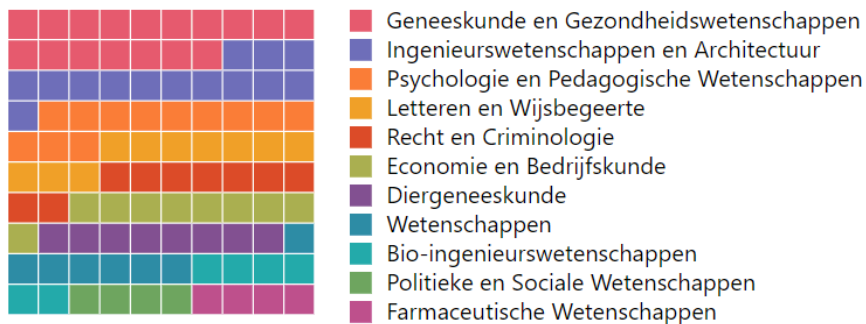
Taakverdeling

Naam	Taak
Emma Neirinck	Kaart, line chart, dataset exploratie, report
Jef Roosens	Dataloaders, correlatie matrix, per capita, readme, report
Robbe Van Rijsselberghe	Teksten site, waffle chart, github actions, totale bar chart

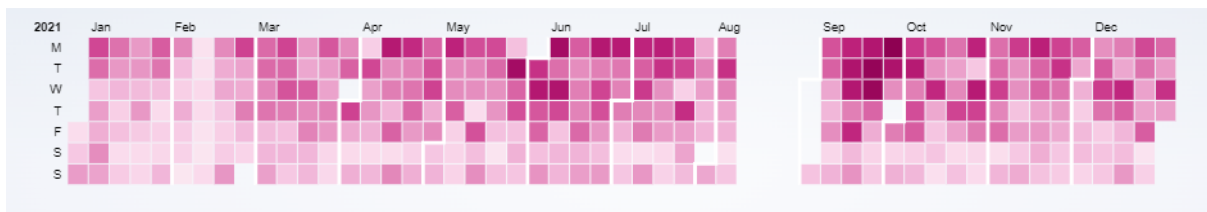
Brainstorm

Waffle chart: Aantal ongevallen per provincie / gewest

Deelnemers over de faculteiten heen



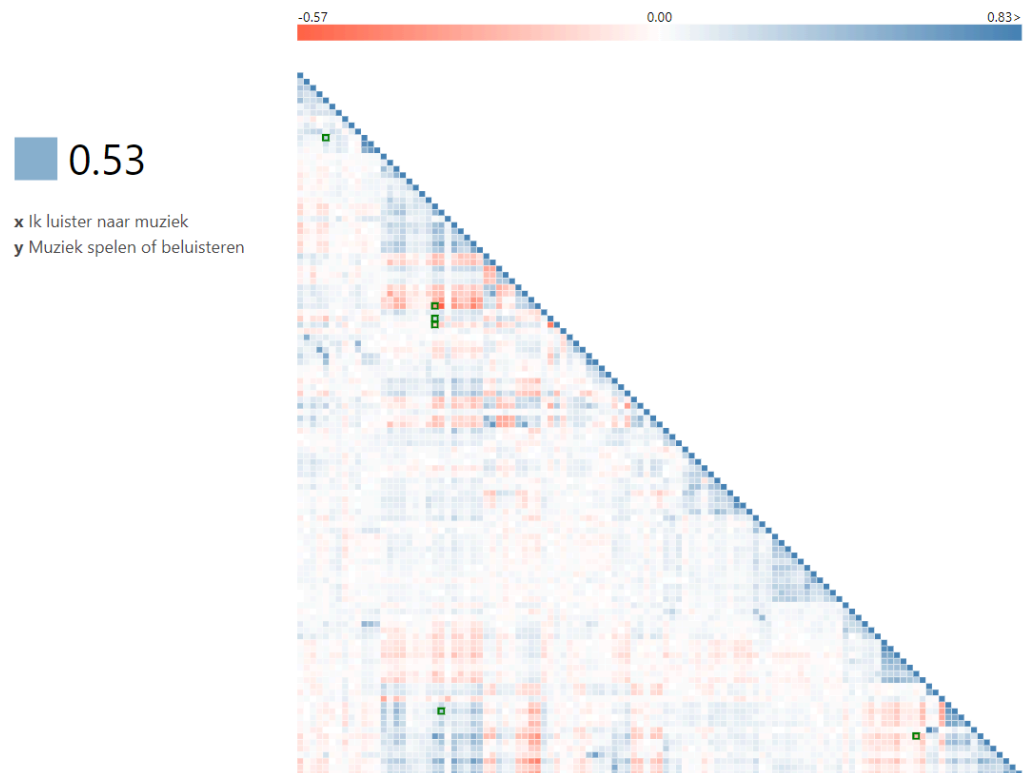
Heatmap: ongevallen per maand per jaar



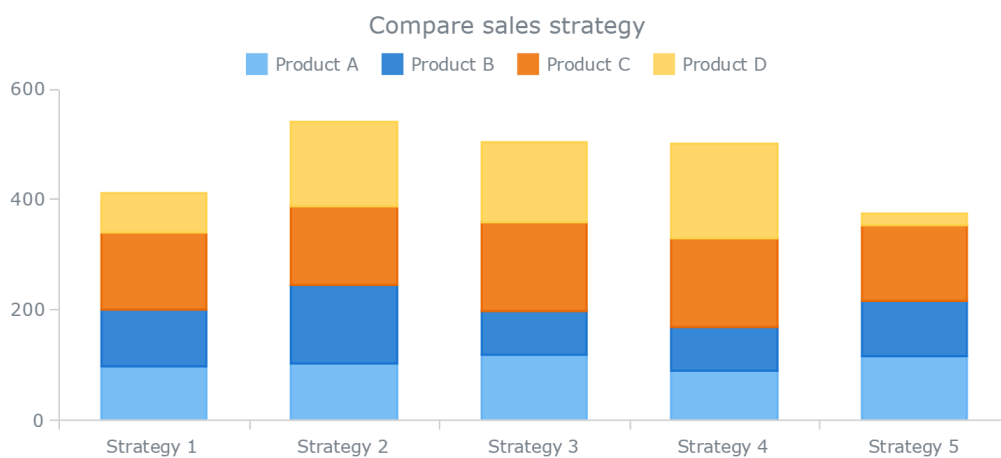
Dot map met sliders: locaties met jaar-maand sliders

<https://users.ugent.be/~bmesuere/datavisualisatie/2018-2019/groep6/>

Correlatie matrix: beschrijvingen en types



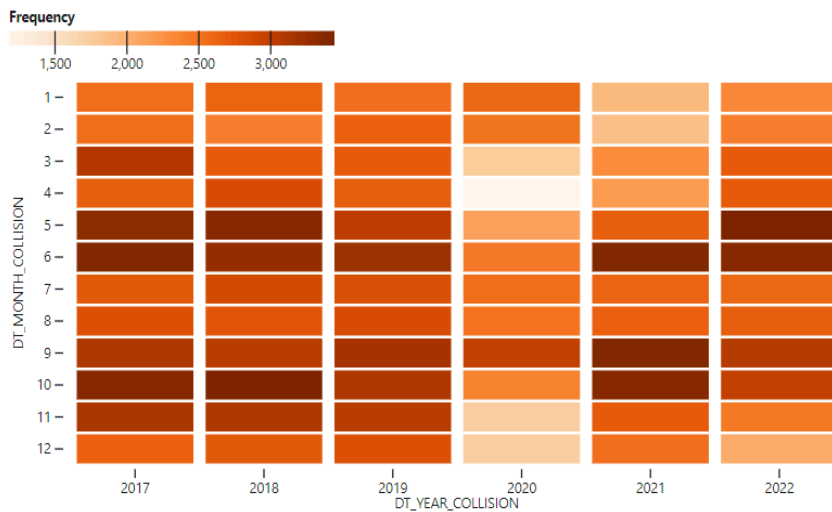
Stacked barchart: type accident (lichtgewonden tot met doden) per maand per jaar



Logboek

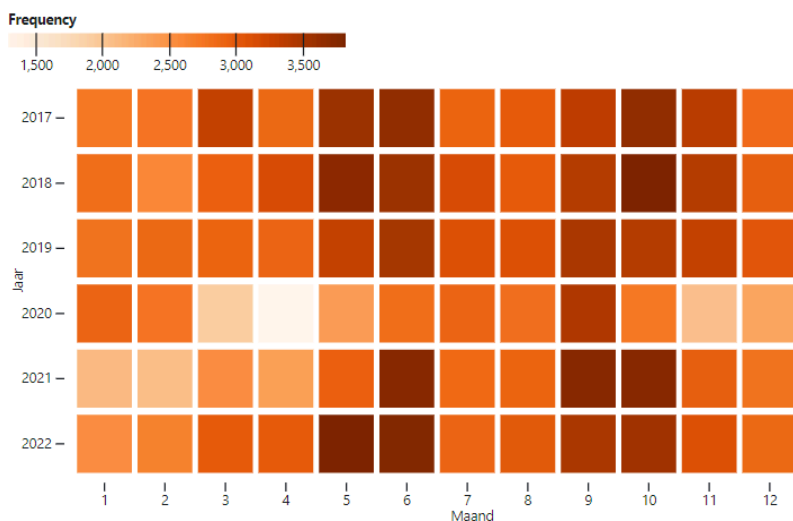
Heatmap ongevallen per maand per jaar

Ongevallen per maand per jaar



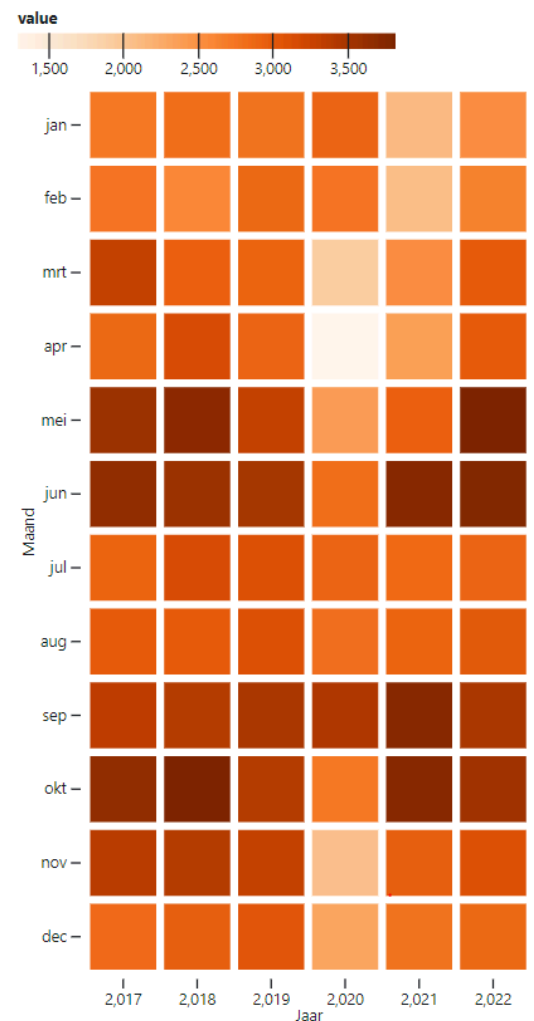
Versie 1

Versie 2 werd in de feedback getoond, we wisselden maand en jaar as om. Uit feedback bleek toch beter terug te wisselen met x-as jaar en y-as maand. De maanden werden ook beter in verkorte naam vorm gezet in plaats van een maand nummering. Door de maanden als string voor te stellen werd dit ook voor de jaartallen gedaan en verschenen plots decimale separators terwijl dit niet gewenst was in versie 3.



Versie 2

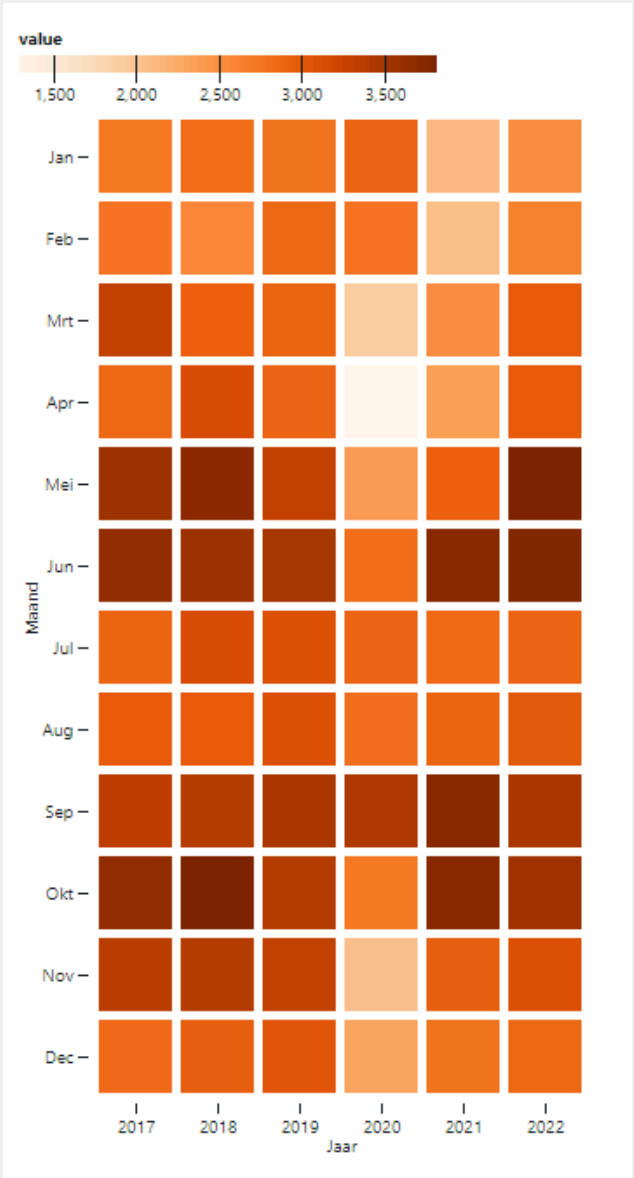
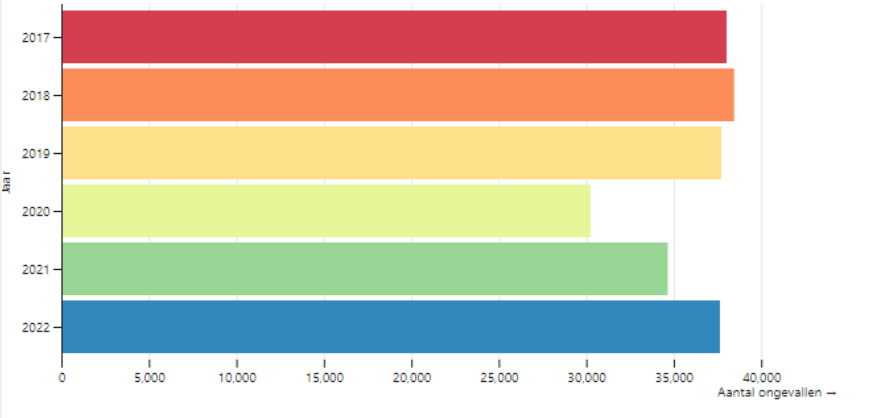
Deze eerste versie had nog de letterlijke labels uit de dataset. Zowel jaar als maand waren beide integers. Er werd gekeken om met logaritmische schaal te werken, maar de waarden waren in dezelfde grootte marge.



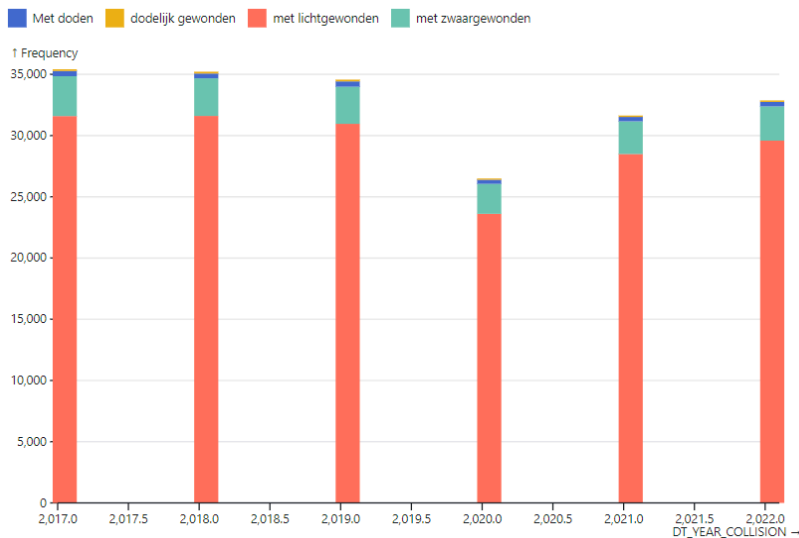
Versie 3

Finale versie

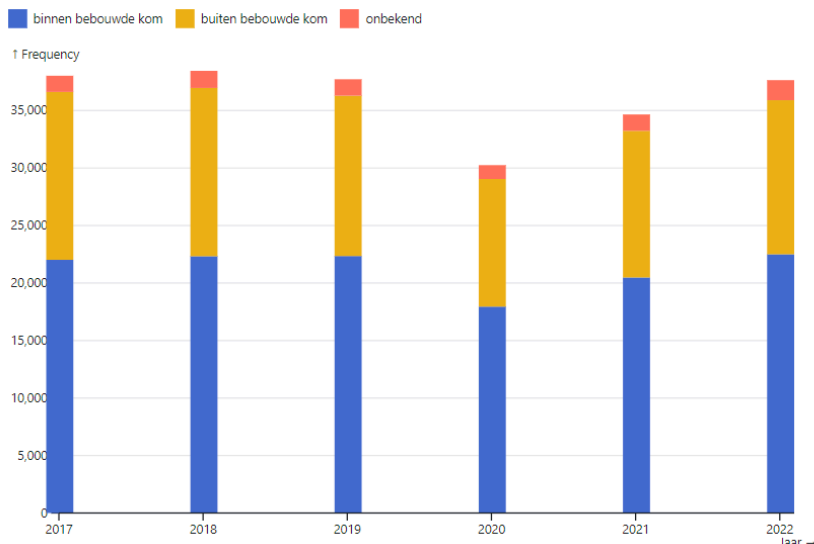
Na de maand benamingen met hoofdletters en de jaartallen zonder decimaal seperator te tonen kwamen we aan onze finale versie. Deze eerste grafiek die we hadden gaf ons ook de aanzet om ook een algemene bar chart te maken om de algemene cumulatief aantal ongevallen per jaar te tonen.



Stacked barchart: attributen bij ongeval



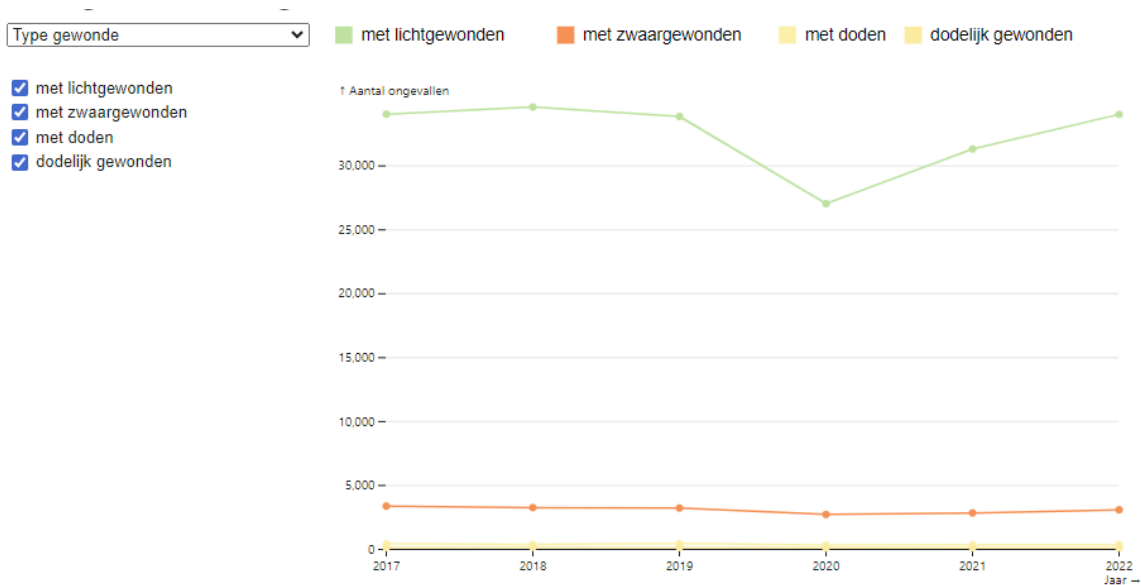
Een eerste poging om de attributen voor te stellen deden we een bar chart per attribuut. Alleen zagen deze plots er nog niet in orde uit. De labels waren hier ook nog de originele vanuit de dataset, x-as had nog de decimale separators en er waren nog ticks aangeduid die ongewenst waren.



Deze versie werd gezien op de feedback, de x-as werd juist gezet. Al snel bleek dat bar chart hier niet de juiste optie was om deze soort data te tonen. Een bar chart per attribuut was ook niet overzichtelijk. Daarnaast was het ook heel moeilijk om de kleinere categorieën van een attribuut duidelijk terug te vinden in de grafiek. (Deze grafiek voor bebouwde kom is nu wel nog duidelijk genoeg.)

Op aanraden vormden we deze bar charts om naar line charts. Hiernaast kozen we er ook voor om dit interactief te maken, we laten de gebruiker nu kiezen welk attribuut men te zien wil krijgen en welke categorieën ze in de chart willen terug zien. Enkele problemen dat we nog steeds hadden:

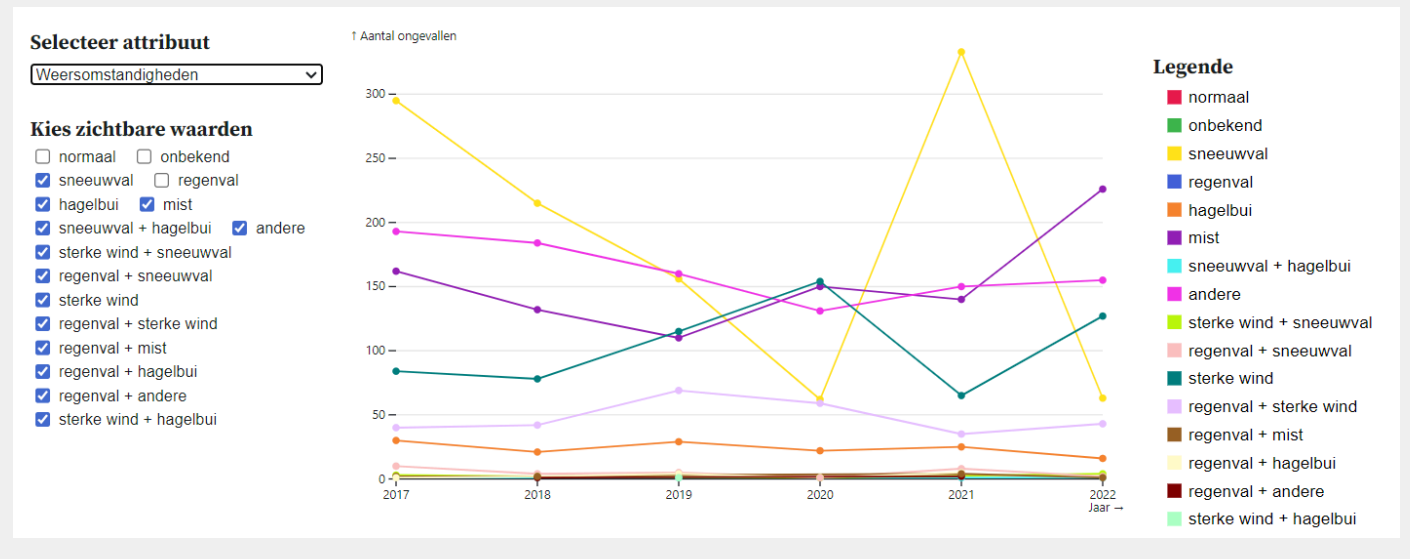
- Sommige attributen hadden veel categorieën waardoor de legende onleesbaar klein werd
- Het standaard kleurenschema voor categorische kleuren van d3 niet voldeed aan het aantal. Deze bevat maar 10 kleuren. We moesten dus nog goede kleurenschema's vinden om een duidelijk onderscheid te kunnen zien tussen de categorieën. Zoals te zien is in onderstaande afbeelding zijn de categorieën niet duidelijk te onderscheiden op basis van kleur.
- Hetzelfde probleem zoals voordien met de bar charts, wanneer alle categorieën initieel aangevinkt staan kon men vaak geen onderscheid maken in de kleinere categorieën
- Nog niet aan data cleaning gedaan, deze kon er vaak heel onleesbaar uitzien.



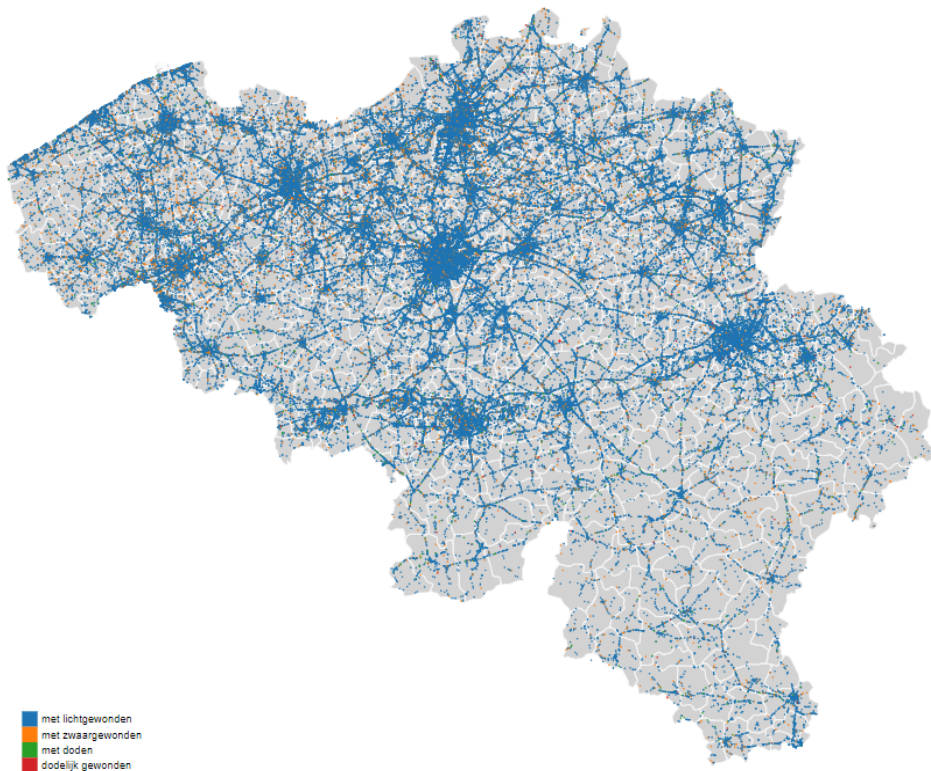
Finale versie

De legende staat nu rechts van de grafiek waardoor alle categorieën nu leesbaar te zien waren, ook hebben nu een kleurenschema groot genoeg om alle categorieën een onderscheiden kleur te geven. Bij elk attribuut staan de meest voorkomende categorieën niet aangevinkt waardoor de gebruiker deze ook in één oogopslag te zien krijgt. Indien alle categorieën aangeduid werden bestaat de kans dat de gebruiker niet de moeite deed om dieper te kijken in de kleinere categorieën. Zoals we nu voor “Weersomstandigheden” kunnen zien dat voor “sneeuwval” de grafiek abnormaal gedrag vertoont door de jaren. Met minieme data cleaning zijn de categorieën nu ook veel leesbaarder.

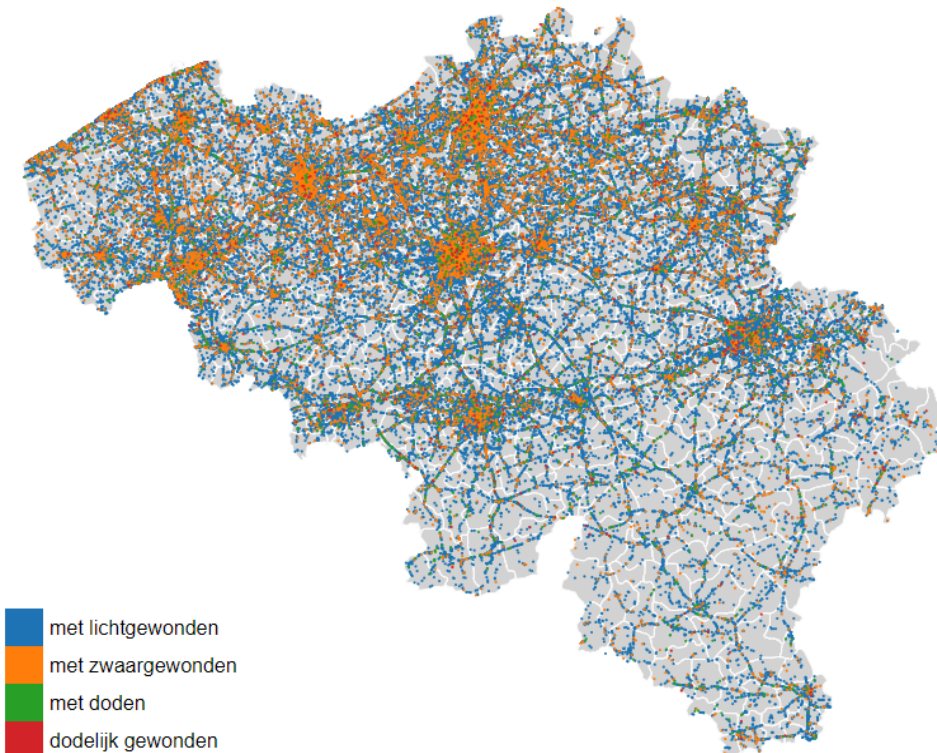
Tijdens de feedback werd voorgesteld om ook categorieën samen te nemen onder “andere”, maar hiervoor was al een categorie. En elke “+” notatie was een combinatie van 2 categorieën dus dan moesten we een keuze maken onder welke categorie we deze moesten brengen, dit hebben we niet gedaan om een manipulatie van de realiteit te vermijden.



Kaart: geolocaties per type slachtoffer



☒ met lichtgewonden ☒ met zwaargewonden ☒ met doden ☒ dodelijk gewonden

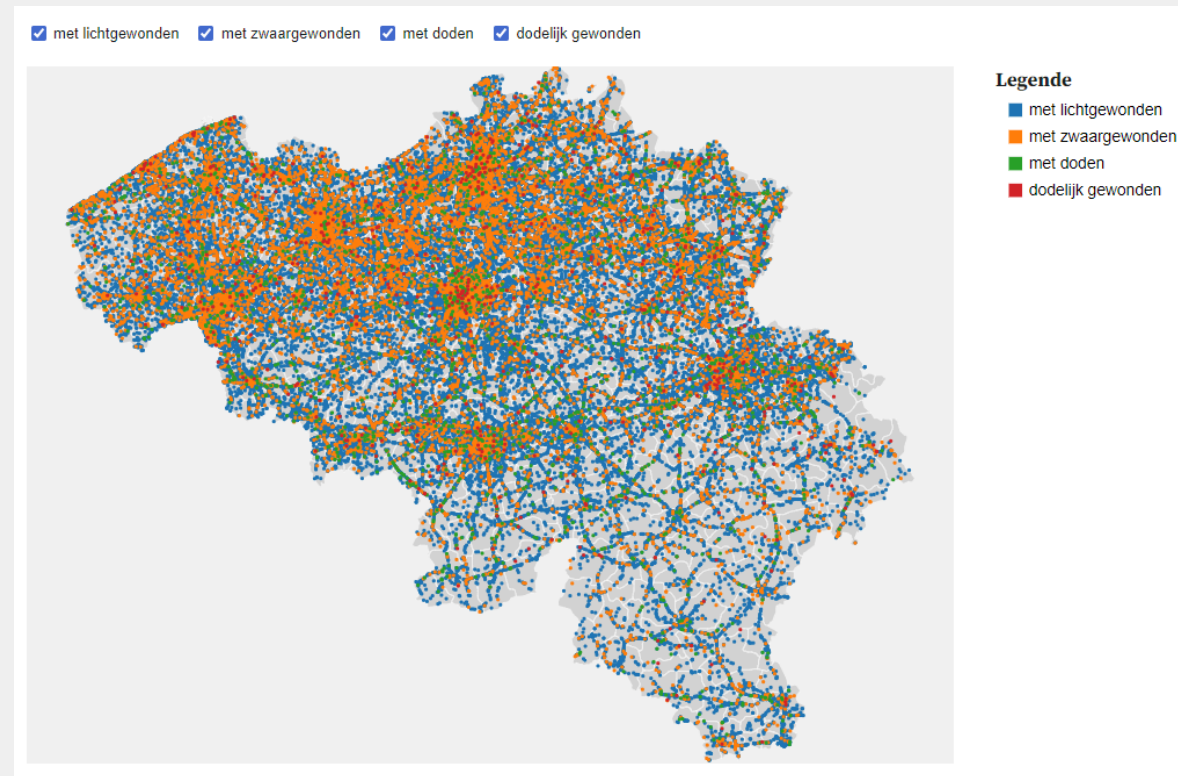


Een eerste versie van de geolocaties voor de ongevallen bracht al wat problemen met zich mee. Ten eerste stond nergens in de dataset vermeldt welke vorm de coördinaten hadden, deze stonden namelijk niet in (longitude, latitude). Na zoeken bleek het lambertcoördinatensysteem gebruikt te zijn, hiervan bestonden ook nog 2 versies. Een coördinatensysteem van 1972 en 2008, een eerste vermoeden was dat de recentste versie gehanteerd werd, maar het bleek de versie van 1972 te zijn. Een ander probleem was dat de datapunten van “met lichtgewonden” over de kleinere categorieën geplot werden.

Nu werd input geïntegreerd in de kaart en de volgorde waarin de datapunten geplot werden per categorie werd aangepast zodat “met lichtgewonden” niet overheerst. Een nieuw probleem kwam echter naar boven. Momenteel werken we met svg's en het aan- en afvinken van de checkboxen duurde heel lang omdat de aanpassing in de kaart voor te tonen categorieën tijd nodig had om opnieuw te plotten. Er werd nog niet genoeg preprocessing gedaan voor de datapunten waardoor de verandering in de kaart vertraagd werd. Een oplossing voor dit werd voorgesteld in de feedback, werk met canvas, zodat alle datapunten niet apart getekend werden maar met pixels.

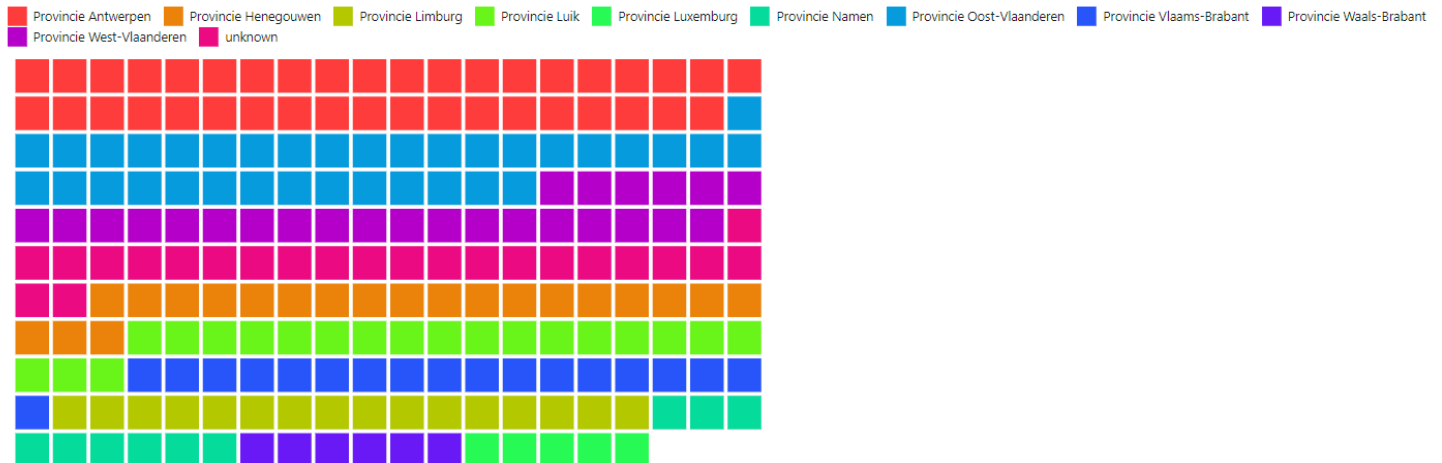
Finale versie

In de laatste iteratie van de kaart werd nu gewerkt met canvas en d3 zoom voor iets meer interactiviteit met de kaart. De legende werd voor betere zichtbaarheid uit de kaart gehaald. De datapunten komen nu ook iets groter in beeld, voordien waren “dodelijke gewonden” en “met doden” niet zo zichtbaar in de initiële rendering van de kaart met alle categorieën aangevinkt.



Waffle chart: ongevallen per provincie

Eerste versie van de waffle chart: Een gesorteerde weergave van alle provincies waar per 1000 slachtoffers een volledig vakje gekleurd werd. Het eerste probleem hierbij was dat de kleuren tussen opeenvolgende provincies niet altijd genoeg verschillen, en dat de legende aan de kleine kant was.

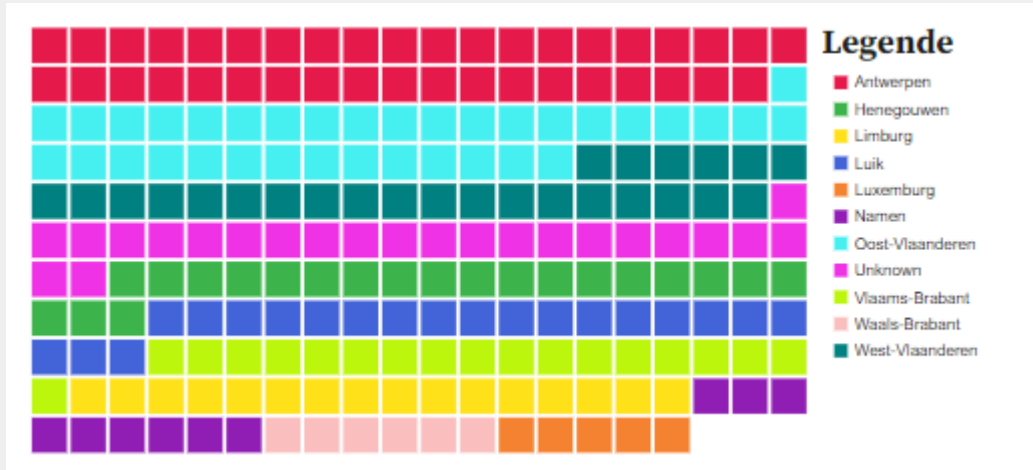


Een tweede versie heeft al een betere inkleuring waardoor de opeenvolgende provincies iets duidelijker te onderscheiden zijn. Een ander probleem was dat deze waffle chart geen rekening houdt met het aantal inwoners per provincie. We hebben besloten om daarvoor een bijkomende grafiek (ongevallen per capita) te maken en deze te behouden om absolute aantallen te kunnen vergelijken.

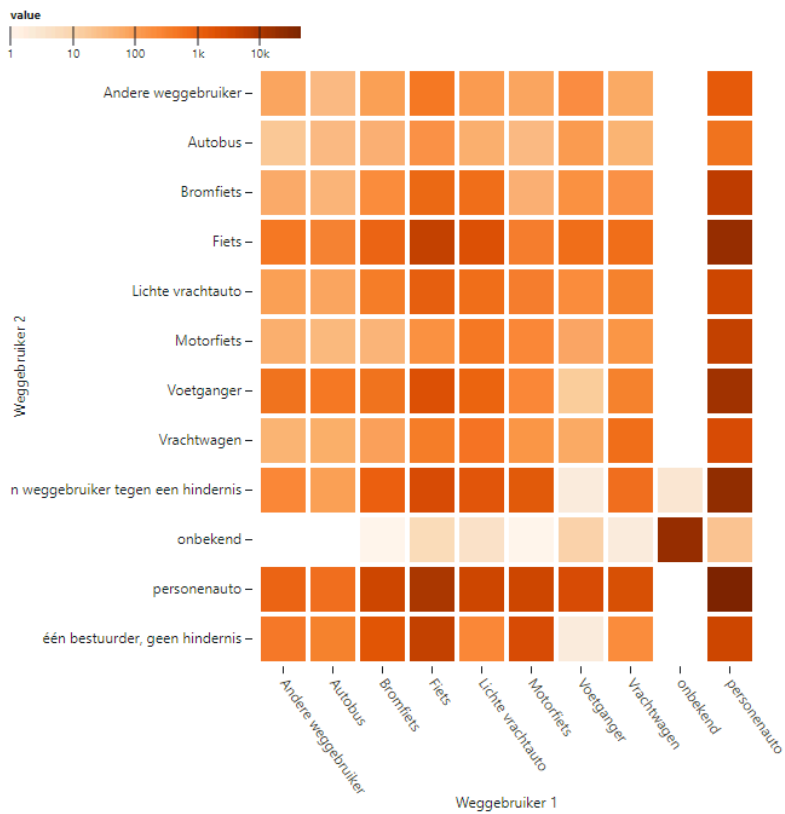


Finale versie

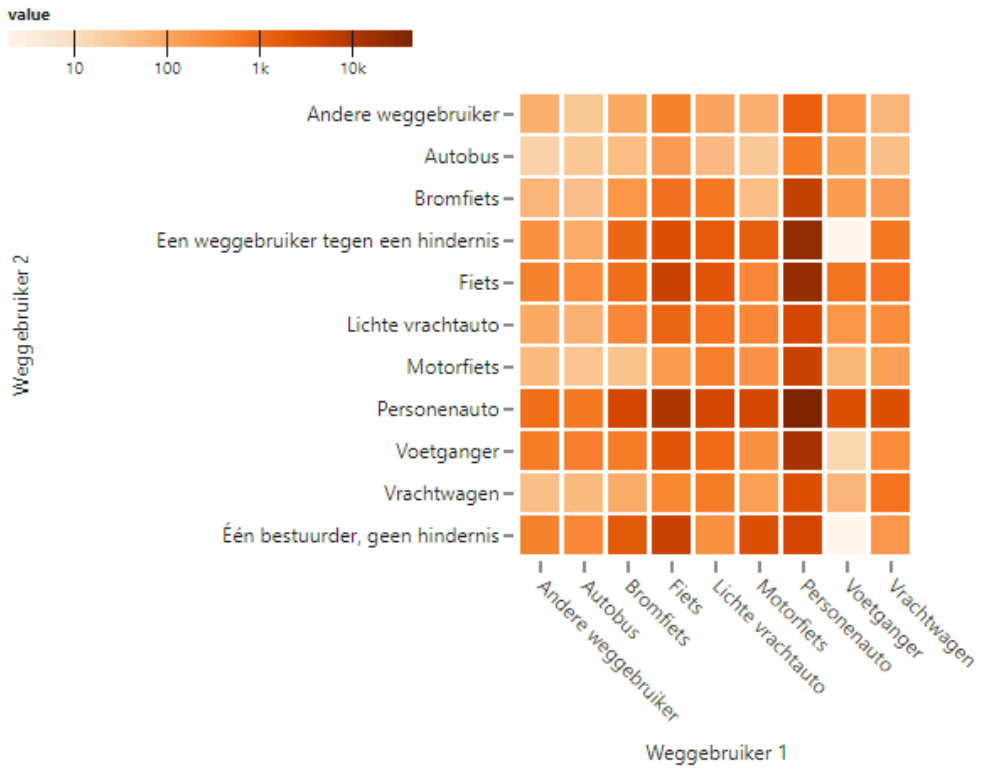
Na de legende nog wat mooier te maken en de 1000 ongevallen per blokje boven in de tekst te vermelden, is dit de finale versie van de waffle chart.



We hadden het idee om een heatmap te maken die illustreert hoe vaak welke combinaties van weggebruikers betrokken zijn bij ongevallen. Onze eerste versie toonde dit al mooi, maar gebruikte nog de rauwe data.

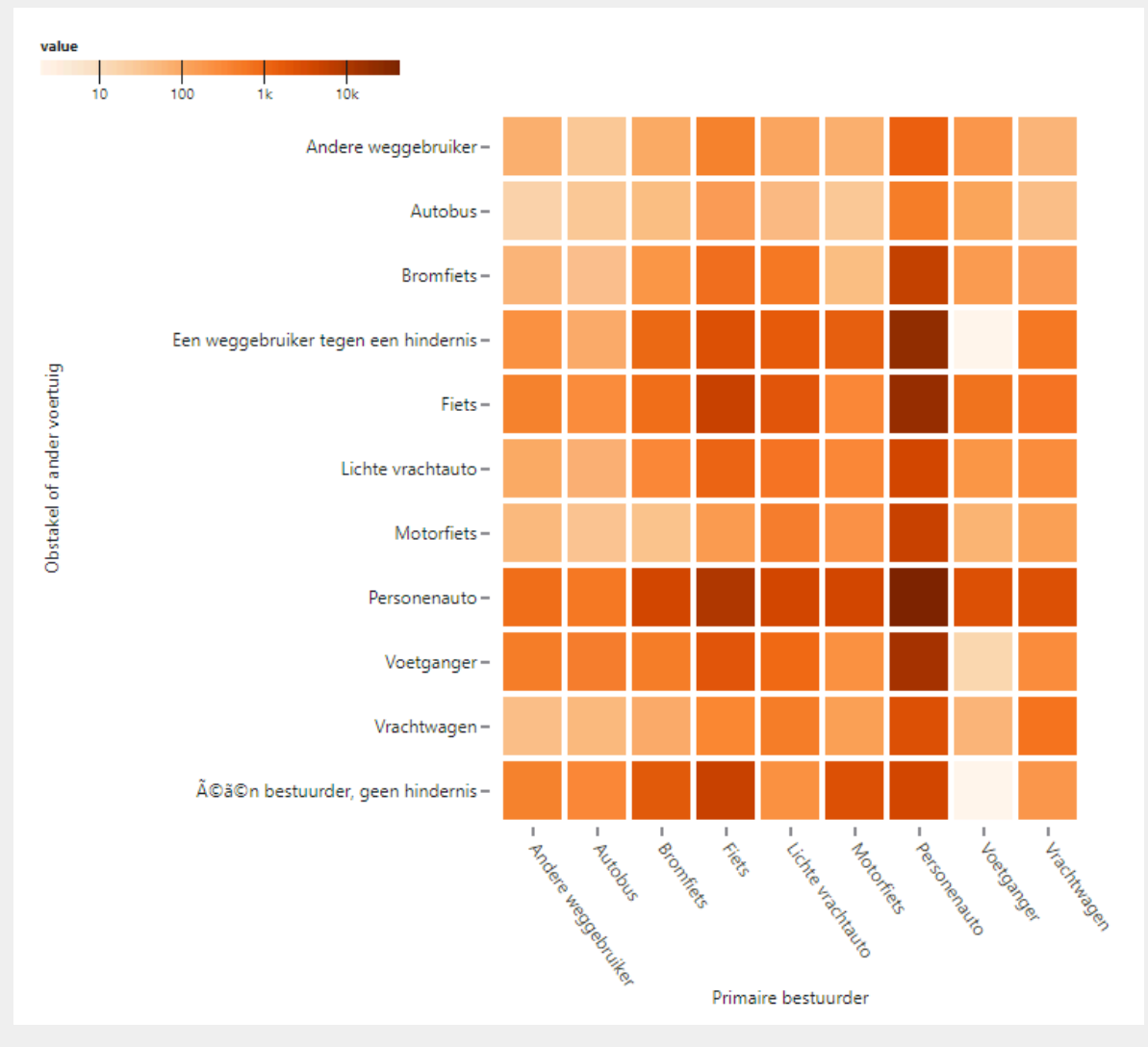


Als tweede iteratie hebben we de data opgekuist, waarbij we de datapunten met “onbekend” als waarde hebben weggelaten. Dit verwijdert de vreemde lege kolom en rij in de eerste iteratie.



Finale versie

Als laatste verbetering hebben we de dimensies van de heatmap aangepast, en hebben we de labels opgekuist zodat alles overzichtelijk en duidelijk is.



Ongevallen per capita

Finale versie

Bij deze grafiek was meteen duidelijk dat we voor een bar chart wouden gaan, gesorteerd op het totaal aantal ongevallen per capita per provincie. Dit illustreert mooi wat we wouden tonen, namelijk welke gemeenten er de slechtste leerlingen in de klas waren op vlak van ongevallen.

