

Life is like riding a bicycle. To keep your balance you must keep moving.

Tableau workshop

Een korte introductie voor de gemeente Enschede tot de wondere wereld van de datavisualisatie.

Door Tycho Horn

Workshop Tableau	3
1. Connectie maken met de data	3
2. Maak je eerste visualisatie!	4
3. Data samenvoegen	5
4. Koppel gegevens op eenzelfde sleutel	6
5. Werken met de kaart	7
6. Je gegevens tezamen op het dashboard	8

Workshop Tableau

Gemeente Enschede

1. Connectie maken met de data

https://github.com/Tychotych1/tableau_gem_enschede

Je wilt je dataset prepareren voor analyse. Zorg dat je de onderstaande punten heb gerealiseerd:

- Maak connectie met het bestand: Inwoners_Migrated_Data.csv
- Check of Tableau de juiste datatypen aan de kolommen heeft gekoppeld:
 - Verander de geografische rol van de kolom: **Gemeente** naar 'County'.
 - Zorg dat je alleen de buurten als naam overhoud.
 - Pas de kolomnaam 'Aantal' aan naar 'Aantal inwoners'.

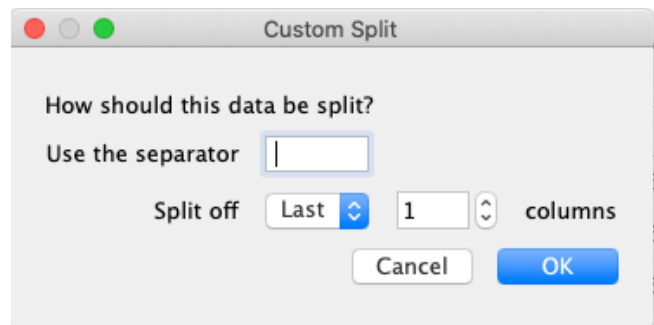
Stappen:

1. Gebruik de connector om een 'Text file' te importeren en selecteer het bestand Inwoners_Migrated_Data.csv.
2. Klik op het 'Abc' icoontje en verander via 'Geographic Role' de kolom naar 'County'.
3. Klik op het pijltje bovenaan de kolom 'Gebied' en klik daarna op 'Custom Split'.

Door een ' ' spatie als separator toe te voegen kunnen we bepalen op welke plekken de data wordt gesplitst.

Selecteer bij Split off 'Last', dit doen we omdat we alleen het laatste gedeelte van de rij willen gebruiken.

4. Verwijder de kolom 'Gebied' en pas de naam van de door te splitten gemaakte kolom aan naar 'Buurt'.



2. Maak je eerste visualisatie!

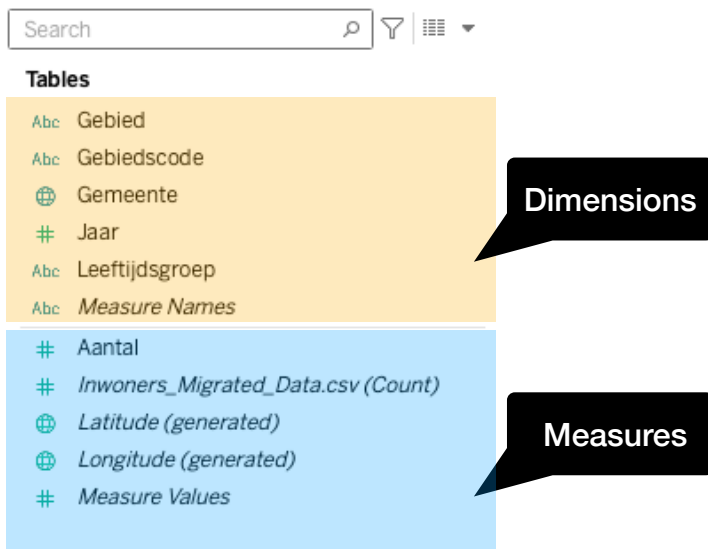


Tableau verdeelt de gegevens in twee hoofdtypen: **dimensions** en **measures**.

Dimensions zijn meestal de velden die niet kunnen worden geaggregeerd;

Measures, zoals de naam suggereert, zijn die velden die kunnen worden gemeten, geaggregeerd of gebruikt voor wiskundige bewerkingen.

1. Zorg dat 'Jaar' is geconverteerd naar Dimensions, je gaat hier immers niet mee rekenen. Je wilt je data namelijk kunnen wegzetten per jaar.
2. Maak een visualisatie die met een staafdiagram en **een filter per buurt** die het verloop laat zien van het aantal inwoners van de wijk 'Boddenkamp'. Denk hierbij goed aan de volgorde van de kolommen: *aantal inwoners per buurt per jaar*.
3. Voor de goede leesbaarheid kun je nog de assen omdraaien. Doe dit.

Stappen:

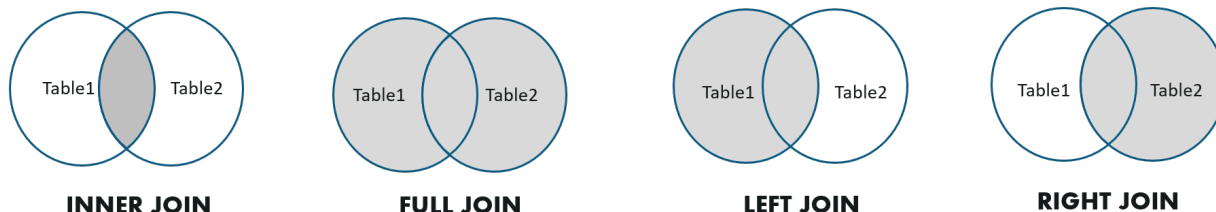
1. Klik op het pijltje rechts naast 'Jaar' en klik op 'Convert to Dimension'. Dit zorgt er voor dat we data nu per jaar kunnen visualiseren, dus eigenlijk als 'tijdseenheid'.
2. Sleep 'Buurt' naar Rows.
3. Sleep 'Jaar' naar Rows, wat zie je nu?
4. Sleep 'Aantal inwoners' naar Columns.
5. Je wilt kunnen filteren op basis van buurten, dus sleep je 'Buurt' naar Filters.
6. Klik op het pijltje naast je zojuist gemaakte filter en klik op 'Show Filter'.
7. Nu kun je uit de lijst van buurten 'Boddenkamp' selecteren.
8. Door op groene knopje te drukken genaamd 'Swap Rows and Columns' draai je de leesrichting van je visualisatie om. Dit kan soms handig zijn!



3. Data samenvoegen

Vaak heb je met meerdere dataset te maken als je je data wilt bereiken. Denk hierbij bijvoorbeeld aan KNMI-weerdata of CBS-gegevens. Maak het kan ook zo zijn dat je uit een database meerdere tabellen nodig hebt. Hierbij kan Tableau je goed helpen door gebruik te maken van: **Union en Joins**.

We gaan nu een oefening doen met **Joins** maar voordat we dat doen eerst even wat informatie:



INNER JOIN: Bevat voor elke rij waarden die overeenkomsten hebben in beide tabellen.

LEFT JOIN: Bevat voor elke rij alle waarden uit de linkertabel en overeenkomstige overeenkomsten uit de rechters tabel. Als een waarde in de linkertabel geen overeenkomstige overeenkomst heeft in de rechters tabel, ziet u een null-waarde in de join-resultaten.

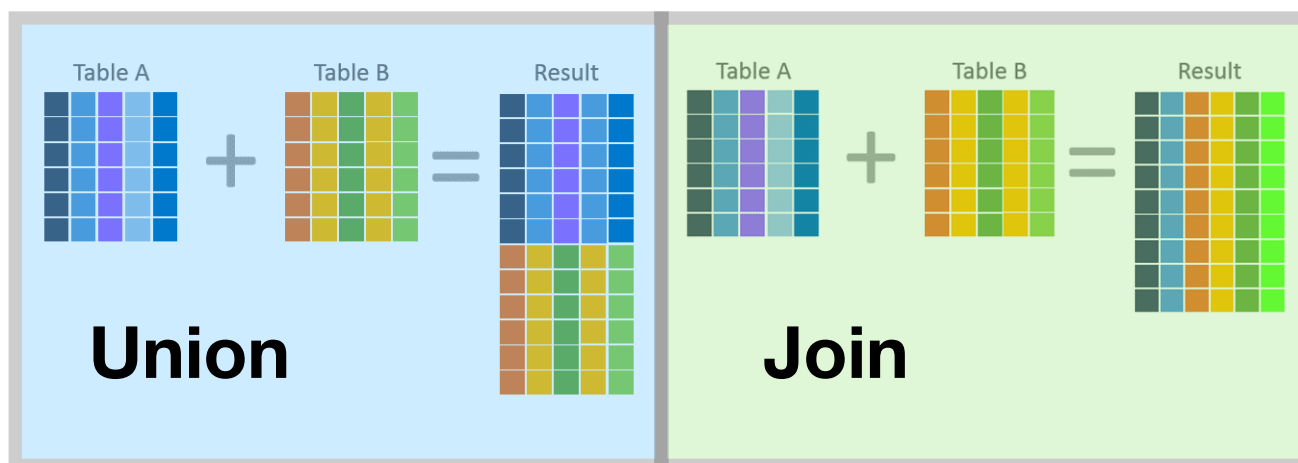
RIGHT JOIN: Bevat voor elke rij alle waarden uit de rechters tabel en overeenkomstige overeenkomsten uit de linkertabel. Als een waarde in de rechters tabel geen overeenkomstige overeenkomst heeft in de linkertabel, ziet u een null-waarde in de join-resultaten.

FULL JOIN: Bevat voor elke rij alle waarden uit beide tabellen. Als een waarde uit een van beide tabellen niet overeenkomt met de andere tabel, ziet u een null-waarde in de samenvoegresultaten.

Union vs Joins

Union; is een methode om gegevens te combineren door rijen van de ene tabel onderaan een andere tabel toe te voegen. U wilt bijvoorbeeld nieuwe transacties in de ene tabel toevoegen aan een lijst met eerdere transacties in een andere tabel. Zorg ervoor dat de tabellen die u samenvoegt hetzelfde aantal velden, dezelfde veldnamen en de velden van hetzelfde gegevenstype zijn

Join(s) is een methode om gerelateerde gegevens op gemeenschappelijke velden te combineren. Het resultaat van het combineren van gegevens met behulp van een join is een tabel die doorgaans horizontaal wordt uitgebreid door gegevensvelden toe te voegen.



4. Koppel gegevens op eenzelfde sleutel

1. Zorg dat de data `Inwoners_Migrated_Data.csv` wordt aangevuld met `CBS_buurten2020.shp`. Hierbij gebruik makend van een **Left Join**, dit omdat we NULL-waarden willen vermijden.
2. Schoon de dataset zo veel mogelijk op door alle kolommen te verbergen behalve:
 1. GM_buurt (hernoem dit naar Stad)
 - Buurt
 - Jaar
 - Leeftijdsgroep
 - Aantal inwoners

=Abc Calculation Buurt	# Inwoners_.... Jaar	Abc Inwoners_Migrated_Data.csv Leeftijdsgroep	# Inwoners_Migrated_Data.csv Aantal inwoners 
--	--	---	--

Als het goed is moet de dataset er nu zo uit zien.

3. Je wil graag gebruik maken van een externe databron. Dit kan echt van alles zijn en is vaak ook cruciaal om verbanden te ontdekken. In dit geval maken we gebruik van CBS-data, altijd handig en accuraat. Voeg aan de dataset **Inwoners_Migrated_Data** de volgende bestanden:

- 'Geregistreerde_criminaliteit_enschede'
- 'Geregistreerde_criminaliteit_nederland'.

Stappen:

1. Importeer het 'Spatial'-bestand **CBS_buurten2020.shp** en sleep dat naast `Inwoners_Migrated_Data`. Nu zal Tableau je vragen op welke sleutel je de twee bestanden met elkaar wil koppelen: gebruik hier **Gebiedscode** en **BU_code** voor aangezien deze bij beide datasets overeenkomend zijn, dus op gelinkt kan worden.
2. Klik nu op het pijltje rechtsboven in de kolommen die je niet meer nodig hebt en klik 'Hide'. Doe dit voor alle kolommen behalve de bovengenoemde.
3. Voeg de dataset over criminaliteit in Enschede toe en doe dit ook voor Nederland. Het zijn beide text-bestanden (csv) dus herhaal je de eerdere procedure van stap 1.1.
4. Let goed op dat je eerst **Enschede** en daarna pas **Nederland** toevoegt. Je relateert namelijk Enschede aan Nederland. Als je het andersom zou doe heb je geen sleutel om op te joinen.
5. Join 'Geregistreerde_criminaliteit_enschede' aan 'Stad' en 'Regio's'. Koppel daarna 'Geregistreerde_criminaliteit_nederland' hieraan met de sleutel 'Regio's' van de beide dataset.

 Inwoners_Migrated_Data



Als het goed is moet je flow er nu zo uit zien.

5. Werken met de kaart

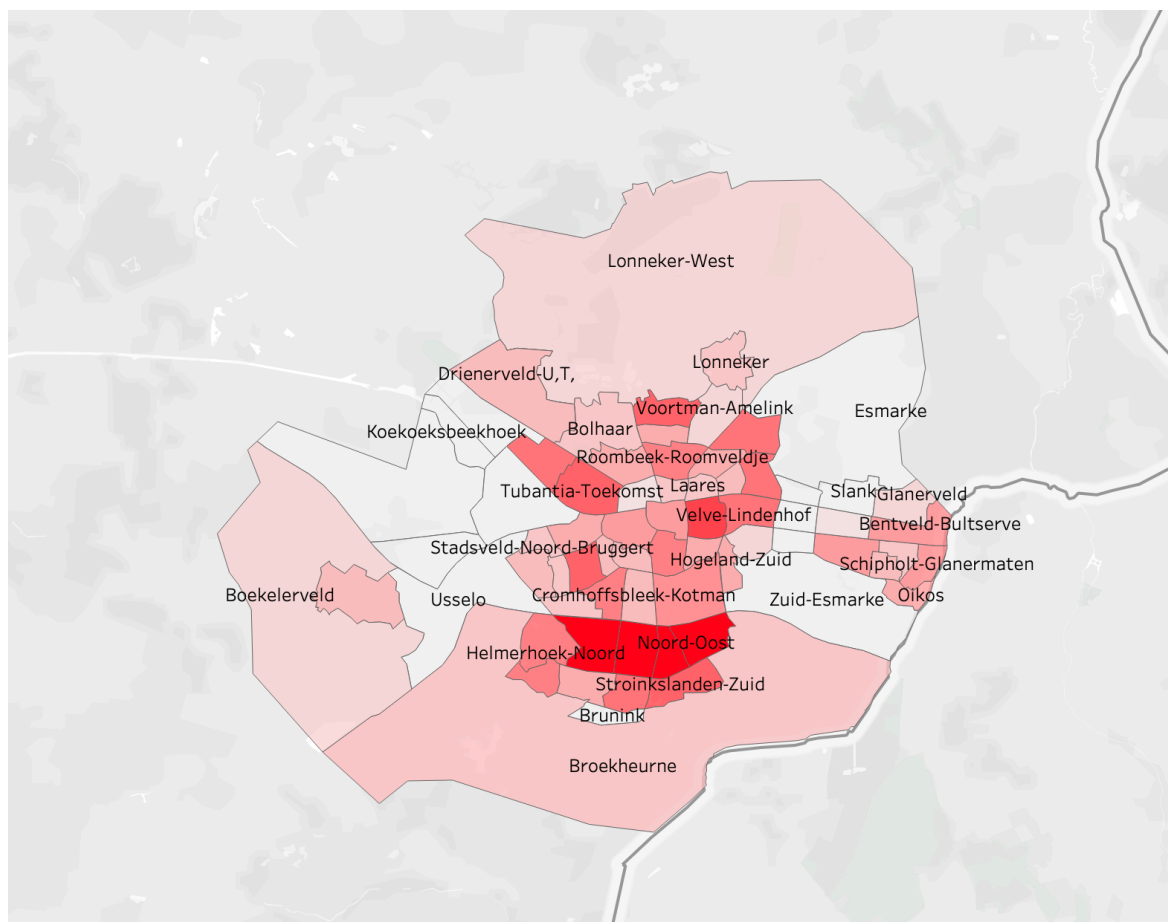
Vaak is het handig om gegevens te visualiseren op een kaart (van Nederland). Hiervoor kent Tableau heel veel mooie handigheden met geologische data (lat / long).

Wij hebben eerder de spatiale-gegevens gejoind aan de inwoners-dataset. Laten we eens kijken wat we daarmee kunnen!

1. Zorg dat je de kaart van Nederland vult met de gemeentes van Enschede.
2. Laat de labels van de buurtnamen zien op de kaart.
3. Laat met een kleurverloop heet aantal inwoners in de buurten zien.
4. Maak het mogelijk om te filteren op 'Aantal inwoners' en 'Jaar'.

Stappen:

1. Sleep 'Longitude' naar Columns en 'Latitude' naar Rows. Hierdoor krijg je de kaart te zien. Anders selecteer je die bij het Marks-menu.
2. Selecteer nu uit de CBS_buurten dataset: 'Geometry'. Nu zul je alle gemeentes van Nederland zien.
3. Aangezien we alleen geïnteresseerd zijn in de gemeentes van Enschede voegen we een filter toe op basis van 'Buurt'. Sleep daarvoor 'Buurt' en 'Jaar' van de Inwoners_Migrated_Data-dataset naar Filter. Kies bij Jaar 'Use all'.
4. In het Marks-menu zie je meerdere opties, waaronder Label. Dit gebruik je om teksten toe te voegen aan de visualisatie. Sleep 'Buurt' naar Label om de buurtnamen op de kaart weer te geven.
5. In datzelfde menu zie je ook Color. Sleep 'Aantal Inwoners' naar Color om de kleuren te baseren op het aantal inwoners per buurt.
6. Dubbelklik op de zojuist gemaakte kleurenpalet en je kunt de kleuren naar gelieven aanpassen

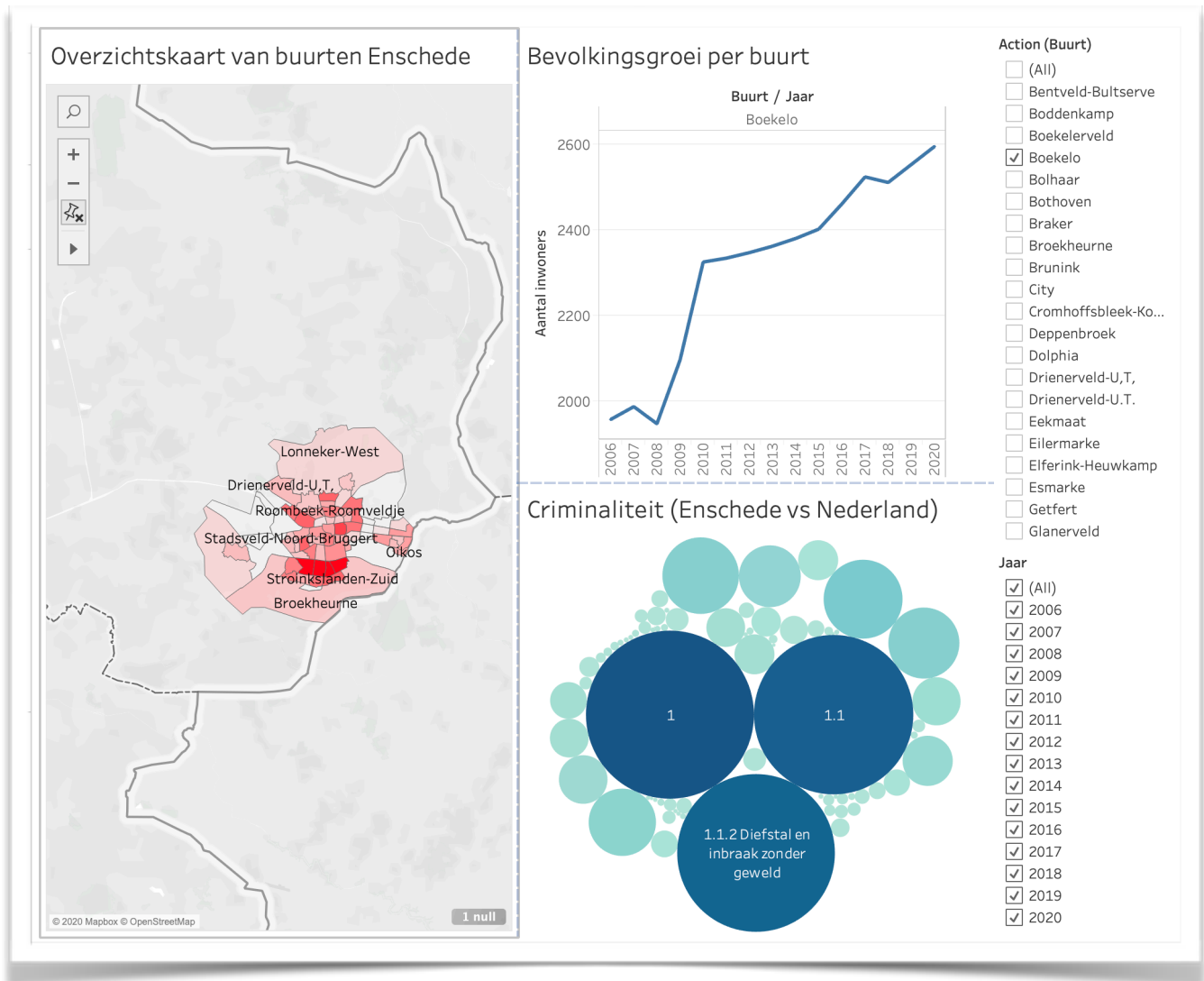
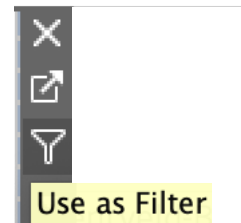


Mijn visualisatie van de inwonersaantallen per buurt in Enschede

6. Je gegevens tezamen op het dashboard

Je kunt snel en gemakkelijk je Sheets op het dashboard slepen. Hierbij kun je handig gebruik maken van de Filter optie 'Use as Filter'. Hierdoor zal de hele dashboard mee worden gefilterd.

Maak voor jezelf een dashboard en presenteer dit aan de groep.



Een voorbeeld van een dergelijk dashboard