

# Logboek Datavisualisatie

## Dataset kiezen

Keuze van de dataset: Emma, Matthias, Robin.

Meerdere datasets zijn voorgesteld. Ons probleem was dat veel daarvan gegenereerd waren d.m.v. AI (van Kaggle), niet steunend op echte data. Uiteindelijke dataset was:

[https://opendata.infrabel.be/explore/dataset/data\\_punctualite\\_typedetrain/table/?disjunctive.r el&sort=maand](https://opendata.infrabel.be/explore/dataset/data_punctualite_typedetrain/table/?disjunctive.r el&sort=maand)

## Ideeën verzamelen

Door: Emma, Matthias, Robin.

Ideeën van wat er getoond kan worden zijn genoteerd in een Google docs, een deeltje ervan:

- *soorten treinen vergelijken doorheen de jaren met hun vertraging*
  - *line chart*
  - *interactief: aan en uitzetten welke treinen op grafiek*
- *grafiek: per tijdseenheid (seizoen, maand, ...) verband aantal vertraagde treinen-aantal minuten tragen*
  - *totaal vergelijking: alle type treinen opgeteld per tijdseenheid vergelijken met de totaal som van vertraging voor die eenheid*  
*(x treinen hadden y minuten vertraging in tijdsperiode z)*
  - *per type trein zelfde vergelijking*
  - *interactief: kies soort trein (alles, IC, L, S, ...)*
- *evolutie stiptheid per seizoen*
  - *cumulatief: starten in het begin van het jaar met 0 vertraging en dan zo door tot eind jaar max vertraging en zien of het duidelijk is wanneer het omhoog schiet*
  - *per treinsoort een plot en eentje met alle treinen samen*

Verder zijn er ook nog enkele voorbeeld grafiekjes met valse data gemaakt om een idee te krijgen van hoe bepaalde soorten grafieken data visualiseren (te vinden op

[https://github.com/Datavisualiatie-UGent/project-dv25-6/blob/main/ideas/example\\_graphs.ipynb](https://github.com/Datavisualiatie-UGent/project-dv25-6/blob/main/ideas/example_graphs.ipynb)).

## Analyse gemaakt in Jupyter notebook

Analyse door: Robin, Emma.

In de directory: <https://github.com/Datavisualiatie-UGent/project-dv25-6/tree/main/analysis> zijn de originele analyses gemaakt van de dataset. Om eerst de data goed te bekijken,

kozen we voor Jupyter notebook, omdat we dit al lang kenden. Hier zijn versies van een grafiek geprobeerd met de data die we willen tonen.

## Verhaal

Verhaal door: Robin, Emma.

Vervolgens werden de beste grafieken uitgekozen en in een verhaal geschreven. Dit is te vinden op:

<https://github.com/Datavisualiatie-UGent/project-dv25-6/blob/main/ideas/story.ipynb>

## Opzetten Observable framework GitHub

Opzetten door: Matthias.

Data omzetten van een `` separated values naar een csv zodat Observable dat kan parsen.

Het Observable project opzetten en beginnen met de grafieken vanuit de Jupyter notebook vertalen naar Observable, vooral Observable Plots zijn hiervoor gebruikt.

## Observable grafieken

Interactieve grafieken gemaakt door: Robin, Emma.

Om de grafieken te maken werden Observable notebooks gebruikt om de uiteindelijke Jupyter code (van de verhaal sectie) om te zetten naar Observable JS. Verder is er in deze notebooks nog wat geëxperimenteerd, om betere voorstellingen te kiezen dan in de Jupyter notebook. Zo was de bubble chart helemaal niet duidelijk in Jupyter, maar werd deze met D3 wel tot een goede grafiek omgezet.

Het grootste deel was het vertalen van Jupyter Python naar Observable JS. Nadien zijn de Observable notebooks vertaald naar Observable Framework zodat de grafieken in de uiteindelijke website zichtbaar zijn. Hierbij was er wat vertaalwerk nodig aangezien het dialect in de Observable notebooks nog enige verschillen heeft t.o.v. JS (zie foto's).



# Finetuning

Afwerking door: Matthias, Robin, Emma.

Als laatste stap is de uiteindelijke tekst meerdere keren nagelezen. Telkens zijn er kleine verbeteringen uitgevoerd om de verwoording duidelijk te maken. De kleuren van sommige grafieken zijn uniform gemaakt en sleutelwoorden uit de tekst zijn in het vet gezet.

# Github Pages

door: Matthias

Het project bouwen en op github zetten zodat het interactief te gebruiken is zonder dat de gebruiker iets lokaal moet doen.