

# Rapport Groep 1 – 13/03/2018

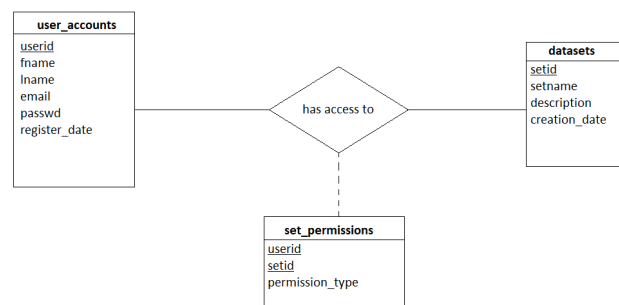
## 1. Inleiding

Het werk dat we reeds hebben gedaan heeft vooral te maken met de basis van het systeem. Dit is onder andere het gebruikerssysteem, het dataset systeem en de basis voor de transformaties. Wat het gebruikerssysteem betreft zijn gebruikers in staat om te registreren, in te loggen en hun gegevens te bekijken en te wijzigen. Er moet nog werk worden gemaakt van administrators die andere gebruikers kunnen beheren. Wat de datasets betreft zijn gebruikers in staat om datasets aan te maken en deze een naam en beschrijving te geven. Er zal in de toekomst nog een systeem komen waarmee gebruikers data kunnen inladen in de datasets. Wel hebben we reeds een aantal transformaties en een view-service geïmplementeerd. In de toekomst wordt de view-service uitgebreid door middel van grafieken, statistieken en de mogelijkheid om data te downloaden. De transform-service zal worden uitgebreid met extra transformaties en een historiek.

Wat de taakverdeling betreft is Luuk verantwoordelijk voor de layout en het framework van de website, Kasper is verantwoordelijk voor de Flask-gerelateerde code die de transform- en view-service met de frontend verbindt. Chris is verantwoordelijk voor het implementeren van de transformaties en het voorzien van toegang tot de tables vervat in de datasets. Atisha is verantwoordelijk voor het inladen van de files in de database, het onderhoud van de databases en hij helpt Chris met de transformaties.

## 2. Design

### Databank ontwerp



De databank is redelijk simpel, met maar 2 entiteiten: `user_accounts` en `datasets`. De relatie tussen deze twee is een many-to-many relatie, omdat een gebruiker aan meerdere datasets moet kunnen en een dataset moet door meerdere gebruiker ge-accessed worden.

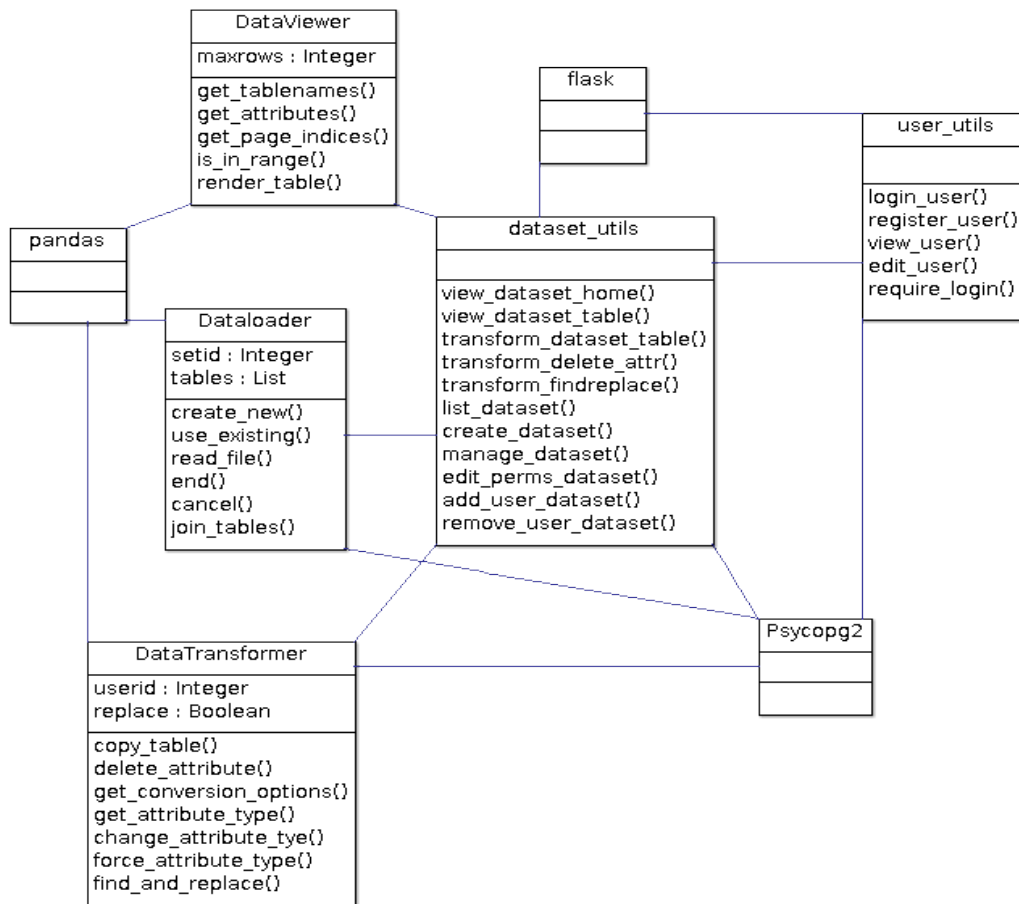
`Permission_type` in `set_permissions` is 'admin', 'write' of 'read'.

De datasets van de user worden opgeslagen in aparte postgresql schema's zodat die van elkaar gescheiden blijven. Hierdoor treden er geen problemen op in verband met tabelnamen of permissions.

De naam van zo'n dataset schema is de `setid`, die gevonden kan worden in de `datasets` tabel.

Per schema zal er ook een extra original (back-up) schema zijn die een de oorspronkelijke data bevat voor eventuele transformaties. In dit schema zal ook een log-table aanwezig zijn die elke aanpassing die gemaakt is bijhoudt.

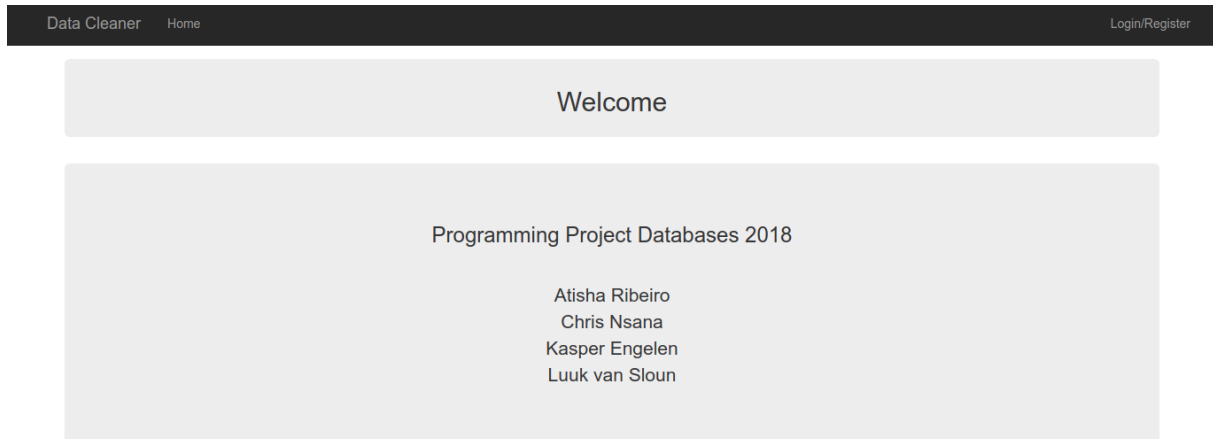
## Programma ontwerp



### 3. Manual

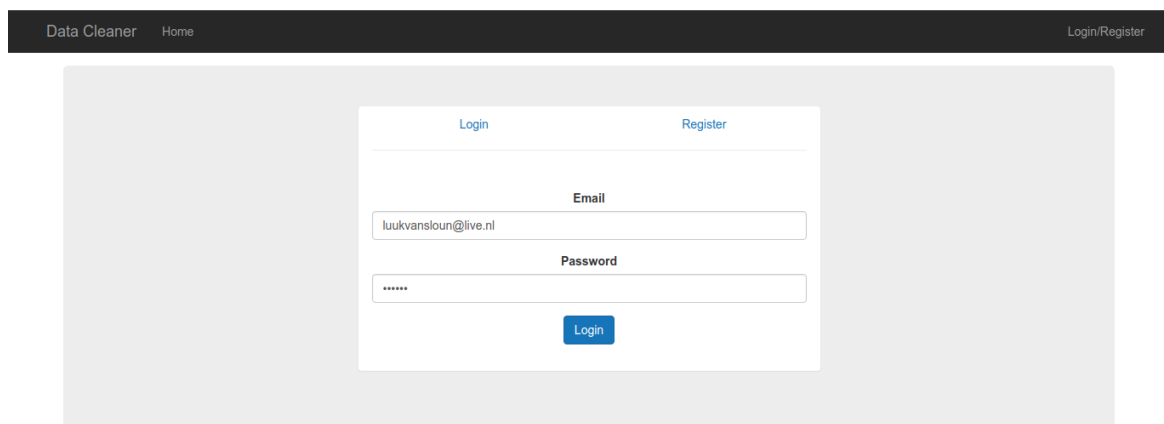
- Standaard home page

Op de standaard home page kan men enkel inloggen en registreren. Een welkomsttekst zal later toegevoegd worden.



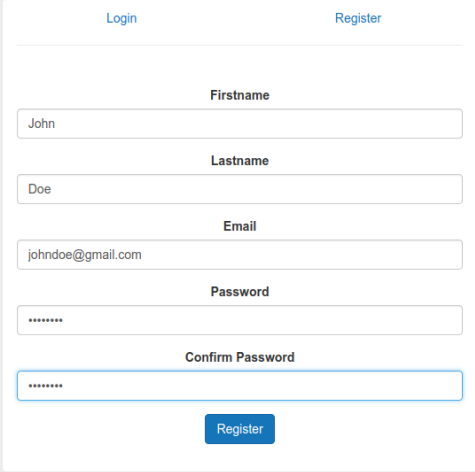
- Login Pagina

Hier kunnen gebruikers inloggen. Wanneer men een foutief wachtwoord invoert of een emailadres dat niet bestaat, zal een flash message ze hiervan op de hoogte brengen.



- Registratie Pagina

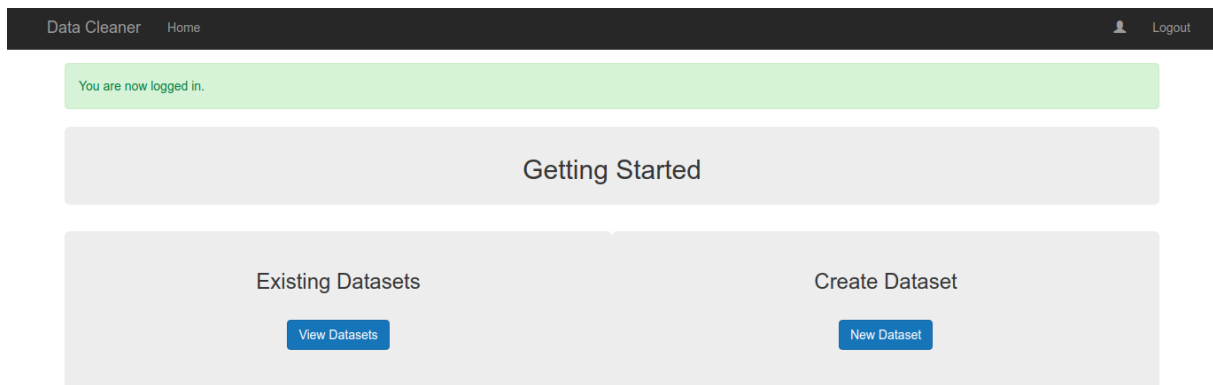
Nieuwe gebruikers kunnen zich hier registreren. Ieder invoerveld heeft de restrictie dat het verplicht ingevuld moet worden. Daarnaast wordt er gecontroleerd of het wachtwoord minimaal 6 en maximaal 50 tekens bevat en of de ingevoerde wachtwoorden overeenkomen.



The registration form is centered on a light gray background. It has a white background and rounded corners. At the top, there are two links: "Login" and "Register". Below these are five input fields: "Firstname" (containing "John"), "Lastname" (containing "Doe"), "Email" (containing "johndoe@gmail.com"), "Password" (containing "\*\*\*\*\*"), and "Confirm Password" (containing "\*\*\*\*\*"). A blue "Register" button is at the bottom.

- Home Pagina wanneer ingelogd

Wanneer een gebruiker ingelogd is, zal de home pagina veranderen ten op zichte van de standaard home pagina. Hier kan men kiezen om een nieuwe dataset aan te maken of de lijst van datasets te beschouwen die de gebruiker al gemaakt heeft of waar hij/zij rechten toe heeft. Een flash message geeft aan dat het inloggen succesvol was.



The home page for a logged-in user features a dark gray navigation bar at the top with "Data Cleaner" and "Home" on the left, and a user icon and "Logout" on the right. Below the navigation bar is a green flash message that says "You are now logged in.". The main content area has a light gray background. It starts with a section titled "Getting Started". Below this are two side-by-side cards. The left card is titled "Existing Datasets" and has a blue "View Datasets" button. The right card is titled "Create Dataset" and has a blue "New Dataset" button.

- Create Dataset Pagina

Hier kunnen gebruikers een nieuwe dataset aanmaken en daarbij een bestand uploaden. Dit mag een csv-, zip- of een dump-bestand. (De file upload moet nog geïmplementeerd worden) Net als bij het registratieformulier dienen beide velden ingevuld te worden.

Data Cleaner
Home
Logout

Create Dataset

Dataset Name

Dataset name

Test\_set

Description

Test123

Create

Upload File

Select Files

Allowed extensions: .csv, .zip, .dump

- Dataset List Pagina

Op deze pagina kan de gebruiker zijn/haar reeds aangemaakte datasets bekijken. Ook staan hier de datasets tussen waar ze rechten toe hebben.

Data Cleaner
Home
Logout

Datasets

Dataset	Description
<a href="#">Test</a>	1234
<a href="#">testname</a>	zippy
<a href="#">zip</a>	testtest

- Dataset View Pagina

Op deze pagina zal de gebruiker de tabellen zien die in de dataset zitten. De ‘view’- en ‘transform’-links brengen ze naar hun respectievelijke pagina.

Data Cleaner
Home
Logout

Dataset: zip
testtest

Manage
Permissions

Table		
test_csv	<a href="#">View</a>	<a href="#">Transform</a>

- Dataset Table View Pagina

Hier kan men de daadwerkelijke tabel beschouwen. De werking van pagina-indices spreken voor zich. Wanneer men in de forward-balk een paginanummer invoert, zal de desbetreffende pagina van de tabel weergegeven worden.

Data CleanerHome

Logout

Transform

Dataset: zip  
testtest

ManagePermissions

1234567891011...1000

Search Page No

region	country	item_type	sales_channel	order_priority	order_date	order_id	units_sold	unit_price	unit_cost	total_revenue	total_cost	total_profit
Sub-Saharan Africa	Namibia	Household	Offline	M	8/31/2015	897751939	3604	668.27	502.54	2408445.08	1811154.16	597290.92
Asia	Indonesia	Meat	Online	H	8/20/2010	472974574	2542	421.89	364.69	1072444.38	927041.98	145402.40
Sub-Saharan Africa	Djibouti	Household	Online	M	2/3/2011	854331052	4398	668.27	502.54	2939051.46	2210170.92	728880.54
Sub-Saharan Africa												

- Dataset Transform Table Pagina

Hier kan men de transformaties uitvoeren. Voor alsnog zijn dit enkel de Find & Replace en de Delete Attribute functies. Bij Find & Replace kiest met een kolom/attribuut waarin gezocht moet worden naar de zoekterm die de gebruiker ingeeft. Indien de zoekterm zich daadwerkelijk in die kolom bevindt, zal deze vervangen worden door de ingevoerde replacement. Bij Delete Attribute kiest men simpelweg een kolom/attribuut dat verwijderd dient te worden. Later zullen aan deze pagina ook de andere transformaties toegevoegd worden.

Dataset: zip  
testtest

Permissions

1234567891011...1000

Search Page No

AttributeSearchReplacement


regionFind & Replace

AttributeDelete Attribute

region	country	item_type	sales_channel	order_priority	order_date	order_id	units_sold	unit_price	unit_cost	total_revenue	total_cost	total_profit
Sub-Saharan Africa	Namibia	Household	Offline	M	8/31/2015	897751939	3604	668.27	502.54	2408445.08	1811154.16	597290.92
Asia	Indonesia	Meat	Online	H	8/20/2010	472974574	2542	421.89	364.69	1072444.38	927041.98	145402.40

- Dataset Manage Pagina

Hier kan men de naam en de beschrijving van de dataset aanpassen.

[Data Cleaner](#) [Home](#)  [Logout](#)

Manage Dataset

Dataset name

zip


Description

testtest

Save Changes

- Dataset Permissions Pagina

Op deze pagina kan een admin gebruikers toevoegen met bepaalde rechten tot de dataset. Aan de hand van het emailadres van een gebruiker worden krijgt deze rechten tot de dataset. De mogelijke rechten zijn 'admin', 'write' en 'read'. Een admin kan hier tevens gebruikers die toegang hebben tot de dataset hun rechten ontnemen door ze te verwijderen.

[Data Cleaner](#) [Home](#)  [Logout](#)

Permissions

Add/Remove User

Email

Permission Type

Admin

Add User

Email

john.doe@gmail.com

Permission

read

✕

## 4. Status

- **[KLAAR]** Bestuderen van verschillende modules en technologieën en selecteren voor ons project {Atisha, Chris, Kasper en Luuk}
- **[KLAAR]** Database design (tables en relaties tussen tables utiwerken en implementeren in SQL) {Atisha en Chris}
- **[KLAAR]** Back-end van het login-systeem implementeren ( werken met user tables etc.) {Kasper en Luuk}
- **[KLAAR]** Front-end van het login-systeem (gebruiker op de site laten registreren & inloggen){Kasper en Luuk}
- **[KLAAR]** Inladen van CSV bestanden in het systeem {Atisha}
- **[KLAAR]** Implementeren v. type-gerelateerde transformaties (type verandering, find-and-replace etc.) {Chris}
- **[BEZIG]** Inladen van SQL bestanden in het systeem {Atisha}
- **[KLAAR]** Gebruikergegevens kunnen zien en wijzigen op de website {Kasper en Luuk}
- **[KLAAR]** Back-end van de view-service biedt een basis view aan {Chris}
- **[KLAAR]** Front-end van de basis view-service uitbouwen {Kasper en Luuk}
- **[KLAAR]** Front-end van de transform-service uitbouwen {Kasper en Luuk}
- **[BEZIG]** Back-up systeem voor oorspronkelijke, ruwe data uitbouwen {Atisha}
- **[BEZIG]** Historiek systeem voor de datasets en hun transformaties uitbouwen {Chris}
- **[BEZIG]** De gebruiker rijen laten verwijderen aan de hand van predicaten {Chris}
- **[GEPLAND]** Geavanceerde view-service functionaliteit ( grafieken en statistieken per kolom) {Atisha}
- **[GEPLAND]** Transformaties van attributen met numerieke types implementeren {Chris}
- **[GEPLAND]** Geavanceerde view functionaliteiten verwerken in de front-end {Kasper en Luuk}
- **[GEPLAND]** Overige transformaties implementeren (one-hot-encoding, datum/tijd etc.) {Chris}
- **[GEPLAND]** De website aantrekkelijker en gebruiksvriendelijker maken {Kasper en Luuk}
- **[GEPLAND]** Mogelijkheid voor gebruiker om eigen SQL queries in te geven realiseren {Chris}
- **[GEPLAND]** Webclient responsive maken {Luuk}
- **[GEPLAND]** Administrator gebruiker {Kasper en Luuk}
- **[GEPLAND]** File upload (CSV, SQL en ZIP files) {Kasper en Luuk}
- **[GEPLAND]** Transformatie – Ontbrekende data invullen {Chris}
- **[GEPLAND]** Tabel kunnen downloaden naar CSV {Chris en Atisha}
- **[GEPLAND]** Systeem automatisch opschonen {Chris en Atisha}