

Stephan Kleijn 21052468, Joeri Lakens 2085692,

Youri van der Lans 21023875, Britney Paul 21080887, Maikel Hussein 19165471

Data Engineering & Data Science

De Haagse Hogeschool

09 Juni 2023

Conclusie en Redenering

[Hosting 1](#_Toc303737203)

[Database: supabase 1](#_Toc1457058617)

[Hosten Azure 1](#_Toc522908552)

[Github 1](#_Toc37350073)

[WebApplicatie 1](#_Toc1955893503)

[Django 1](#_Toc756423589)

[Keuze in datagroepen 1](#_Toc1678617085)

[Keuze Diagrammen 2](#_Toc493750675)

[Design keuze 2](#_Toc199555582)

[Hulpmiddelen 2](#_Toc1721409441)

[Draw.io 2](#_Toc496267461)

[Trello 2](#_Toc789440982)

[Discord 2](#_Toc693903115)

[OneDrive 2](#_Toc1569410021)

# Hosting

## Database: supabase

Supabase heeft een simpele API die wij zonder grote problemen hebben kunnen toevoegen. Ook hoef je niet voor elke kleine verandering een query te maken maar kan dit gewoon via de gebruiksvriendelijke interface. Voor query’s die wel gedaan moeten worden zijn er standaard scripts. Ook is Supabase gratis waardoor er geen kosten aan overgehouden worden.

## Hosten Azure

Wij zijn begonnen met kijken naar Azure omdat wij hier credits voor hadden. Na een beter onderzoek is gebleken dat Github en Azure makkelijk te integreren zijn met elkaar en er niet al te veel handelingen in de terminal nodig zijn. Later is ook gebleken dat er makkelijk gewisseld kan worden van server. Dit aan de hand van onze applicatie die constant een timeout error gaf, wat na de wissel was opgelost. Verder houdt Azure de application en container logs bij die worden samengevat zodat je als gebruiker sneller weet wat er aan de hand is met je webapplicatie en of container.

## Github

Git(hub) is ingezet voor het versiebeheer. Door het gebruik van Github desktop en github.com is de implementatie van git een erg gebruiksvriendelijke en snelle manier om versies te onderhouden. Het gebruik van branches is een essentiële feature in de werkflow van het project, hierdoor blijft de main branch schoon en kan deze makkelijk gedeployed worden op Azure. De code is aan het eind van het project ook makkelijk in te zien voor elke member van de repository voor eventuele nacontrole.

# WebApplicatie

## Django

Python is makkelijk om op te pakken als beginnend programmeur. Samen met de vele toepassingsmogelijkheden voor data-analyse kan je al vrij snel beginnen met data inlezen en hierbij grafieken maken. Door het Django framework te gebruiken is de backend ook in Python. Django is top voor het bouwen en genereren van webpagina’s. Zo kan je direct via de admin pagina een extra pagina toevoegen waar je weer verder op kan bouwen.

## Keuze in datagroepen

Door tijdens de domeinanalyse inzicht te krijgen op het bedrijfsproces is er met inzicht een hoofdvraag opgesteld tijdens het maken van het plan van aanpak. De hoofdvraag luidt als volgt: Welke producten worden het meest verkocht per productgroep? Ons doel met deze hoofdvraag is een marktinzicht te creëren voor OutdoorFushion zodat business analisten aan de hand van de dashboarden een ondernemingsplan kunnen maken. Omdat er meerdere databronnen zijn aangeleverd is er ook per bron een deelvraag gesteld die betrekking heeft op de belangrijkste waarde van die database. Samen vormen de uitkomsten van de deelvragen een antwoord op de hoofdvraag.

## Keuze Diagrammen

Alle Diagrammen zijn gemaakt met behulp van Plotly. In javascript biedt Plotly veel customization opties. De diagrammen zijn ook interactief en kunnen bijvoorbeeld in schaal live worden aangepast.

In het dashboard zijn veel staafdiagrammen te vinden. Staafdiagrammen geven een overzichtelijke vergelijking tussen datapunten, hierdoor kan bijvoorbeeld snel worden ingezien of een waarde over een periode afneemt of welk product het meeste winst oplevert.

## Design keuze

Door de achtergrond van de website grijs (#212121) achtergrond te maken tonen de diagrammen beter omdat deze lichter gekleurd zijn. De website is niet helemaal zwart omdat dit minder prettig is voor de ogen van de gebruiker.

## ETL

De webapplicatie is voorzien van een updateknop op de home pagina. Deze knop zorgt voor het opnieuw runnen van de ETL-scripts.

De ETL-scripts werken als volgt:

* Haal data op uit database. We hebben de relevante tabellen uit de geleverde datasets naar Supabase gecopieerd. Dit is het beginpunt van onze datastraat
* Voer ETL operaties uit op de opgehaalde data
* Schrijf de resultaten weg naar de F-tabellen in de Supabase. ‘F’ voor Final. Dit zijn de tabellen die gebruikt worden voor de visualisaties

# Hulpmiddelen

## Draw.io

Diagrams.net is een gratis en open source platformonafhankelijke software voor het tekenen van grafieken. De interface kan worden gebruikt om diagrammen te maken, zoals stroomdiagrammen, wireframes, UML-diagrammen, organigrammen en netwerkdiagrammen. Door deze breed inzetbare tool aan te passen naar de wensen van het project kunnen er veel diverse documenten worden opgebouwd binnen deze tool. Er kan live gewerkt worden door meerdere projectleden tegelijkertijd in hetzelfde document.

## Trello

We hebben besloten om Trello te gebruiken om onze taken bij te houden. Met Trello kunnen we taken toewijzen aan individuele teamleden, zodat iedereen duidelijk weet wie verantwoordelijk is voor welke taak. Dit zal ons helpen om de werklast efficiënt te verdelen en ervoor te zorgen dat alle taken worden uitgevoerd. Daarnaast biedt Trello ook een overzichtelijke weergave van de voortgang van het project, waardoor we gemakkelijk kunnen zien hoe ver we zijn gevorderd en of er nog actie moet worden ondernomen. Het gebruik van Trello had ons geholpen om georganiseerd te blijven en effectief samen te werken.

## Discord

Discord is gekozen als primair communicatiemiddel. Door gebruik van de app en de desktopversie is het wisselen tussen werkstations snel en zonder moeite. De app geeft notificaties van nieuwe berichten in de chatrooms waardoor iedereen snel updates over het project in kan zien. Vergaderingen in de voicerooms worden ondersteund door de mogelijkheid om je scherm te delen, hierdoor kan iedereen meekijken op je scherm als je een vraag of probleem hebt. Discord heeft ook de mogelijkheden om (kleine) bestanden te delen zodat er op eenvoudige wijze bestanden gedeeld kunnen worden.

## OneDrive

OneDrive wordt gebruikt voor het makkelijk delen en maken van documenten. Omdat de school accounts van Microsoft zijn wordt OneDrive gebruikt boven Google Drive. Door deze Microsoftaccounts is het delen en vinden van andere studenten ook simpeler. Via OneDrive kan er met meerdere personen aan één document worden gewerkt.