Projectplan

**Groep: 4.5**

**Groepsgenoten: Akasha Wilms (18104355),**   
**Mehmet ali Yildiz (21108730),**   
**Stan Runge (21092141),**   
**Bram Steenkamp (21113653),**   
**Tugay turk Türk (21122377).**

# Inhoudsopgave

Stepping Stones + canvas processen

Kwaliteitscriteria

Canvas afbeelding

Diagnose

1. Doelstelling

Doel

1. Aanpak (methoden en technieken)
2. Planning + deliverables
3. Benodigde middelen

Resultaat

1. Beoogde resultaten

Stepping Stones + canvas processen

Diagnose

**Analyseproces**

- Domeinanalyse

- Kostenraming maken

- Kwaliteitsonderzoek datamateriaal

Doel

**Ontwerpproces**

- Business-assets en KPI's in kaart brengen

**Implementatieproces**

- Verzamelen datamateriaal (scrapen en aangeleverde data)

- Eventueel herstellen database

- Filteren en sorteren van data

- samenvoegen databronnen (ETL)

Resultaat

**Maakproces**

- Machine learning toepassen

- omgeving dashboard realiseren

- dashboard realiseren

- Een rapport samenstellen dat de betrouwbaarheid van het dashboard verklaart

Kwaliteitscriteria

**Analyseproces**

- Domeinanalyse

*Kwaliteitscriteria:*

De domeinanalyse voldoet wanneer er een duidelijk beeld is van de opdracht van de opdrachtgever. In de domeinanalyse staan een duidelijke aanleiding en doelstelling. In de domeinanalyse wordt er geschetst wat de eindgebruiker ermee wil bereiken. Er behoort ubiquitous language verklaard te worden. Er behoort een domeinmodel gemaakt te zijn.

- Kostenraming maken

*Kwaliteitscriteria:*

De kostenraming voldoet wanneer er een concreet kostenplaatje is gemaakt van de hosting van het datalandschap.

- Kwaliteitsonderzoek datamateriaal

*Kwaliteitscriteria:*

Dit proces voldoet aan de eisen wanneer er een helder rapport is opgesteld die de kwaliteit van het gegeven datamateriaal opheldert.

**Ontwerpproces**

- Business-assets en KPI's (Kritieke Prestatie Indicator) in kaart brengen

*Kwaliteitscriteria:*

Dit proces voldoet aan de eisen wanneer de business-assets in kaart zijn gebracht. Ook voldoet dit proces aan de eisen wanneer de KPI's in kaart zijn gebracht.

**Implementatieproces**

- Verzamelen datamateriaal (scrapen en aangeleverde data)

*Kwaliteitscriteria:*

Dit proces voldoet aan de eisen op het moment dat er voldoende data uit secundaire databronnen verzameld om representatieve resultaten te kunnen tonen.

- Eventueel herstellen database

*Kwaliteitscriteria:*

Indien blijkt uit het kwaliteitsonderzoek dat de data uit het gegeven datamateriaal in ieder geval voor een deel onjuist blijkt te zijn, kan dit proces van start gaan. Het herstellen van de database voldoet dan aan de eisen in het geval dat lege velden representatief gevuld worden en incorrecte data gecorrigeerd is.

- Filteren en sorteren van data

*Kwaliteitscriteria:*

Ter voorbereiding op het analyseren van de data behoort de data eerst juist gefilterd en gesorteerd te zin om aan de kwaliteitscriteria te voldoen van dit proces.

- Samenvoegen (ETL)

*Kwaliteitscriteria:*

Om te voldoen aan de kwaliteitscriteria van dit proces behoren de databases samengevoegd te worden om de benodigde data te kunnen analyseren naar wens van de opdrachtgever.

**Maakproces**

- Machine learning toepassen

*Kwaliteitscriteria:*

Machine learning moet toegepast zijn om verwachte aankoopcombinaties te bepalen en deze moeten vervolgens inzichtelijk gemaakt kunnen worden.

- Omgeving dashboard realiseren

*Kwaliteitscriteria:*

De omgeving staat idealiter gedeloyed in een webomgeving en voldoet wanneer deze makkelijk bereikbaar en bruikbaar is.

- Dashboard realiseren

*Kwaliteitscriteria:*

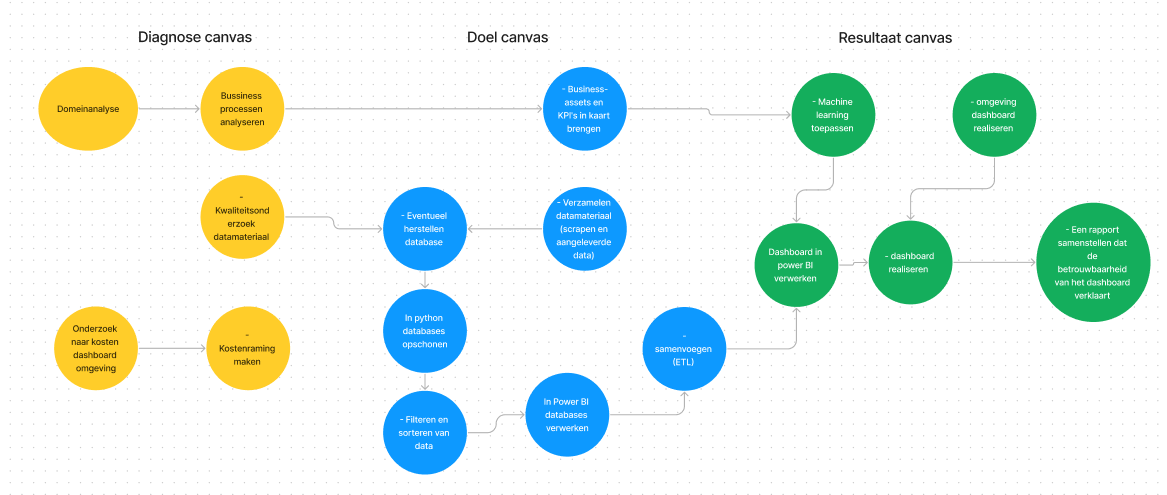
Het dashboard voldoet wanneer het dashboard interactief is, het geheel werkt en live is. Op het dashboard staat relevante sturingsinformatie voor OutdoorFusion gepresenteerd inclusief Machine Learning uitkomsten. Het liefst gedeployed in een webomgeving. Als bronnen voor dit dashboard is de aangeleverde data gebruikt maar ook op het web gevonden data. Het dashboard is zeer breed inzetbaar en richt zich op een breed scala aan management vragen en KPI’s.

- Een rapport samenstellen dat de betrouwbaarheid van het dashboard verklaart

*Kwaliteitscriteria:*

Het rapport omtrent het dashboard en zijn data voldoet wanneer de keuzes voor het dashboard worden toegelicht en alles inhoudelijk wordt verantwoord.

CANVAS afbeelding



Toelichting canvas

In dit project zullen wij eerst analyserende taken uitvoeren die slaan op de diagnose van de stepping stones methodiek. Eerst zullen wij het domein met de business processen en het gegeven datamateriaal analyseren. Vervolgens gaan wij in kaart brengen hoe wij het dashboard gaan realiseren door externe data te verkrijgen, de gegeven data opschonen en alle databronnen verwerken in één omgeving. Tot slot gaan wij machine learning toepassen op het datamateriaal, waarbij wij het dashboard en zijn omgeving kunnen realiseren. Hierop gaan wij nog een rapport leveren op de kwaliteit van het dashboard en hoe betrouwbaar het is.

Doelstelling

Opdracht:

* Een werkend, interactief, live dashboard waar alle relevante sturingsinformatie voor OutdoorFusion gepresenteerd wordt inclusief Machine Learning uitkomsten
  + Deze is het liefst gedeployed in een webomgeving
  + Bron voor het dashboard zijn
    - Primair: de aangeleverde data,
    - Secundair: (on)gestructureerde publiek beschikbare (internet)bronnen
  + Het dashboard moet zeer breed inzetbaar zijn en zich richten op allerlei management/sturingsvragen/ KPI's
* Een voortdurende overdraagbaarheid van het project
* Een rapport waaruit blijkt dat de dashboards functioneel en met betrekking tot performance adequaat werken
* Een procesvoering waarbij rekening gehouden wordt met mogelijke tussentijdse uitbreiding van de scope van het data-landschap
* Een raming van de kosten voor de hosting van het datalandschap

Aanpak (methoden en technieken)

Dataverzameling methodieken

* Data van externe en publieke bronnen gaan we verzamelen door middel van scraping scripts die we maken die websites gaan bezoeken en data verzamelen.
* De aangeleverde 3 databases gaan wij opschonen, filteren en sorteren (ETL), waarna zetten wij de data in 1 grote SQL-database zodat wij dit makkelijk kunnen gebruiken in ons dashboard.

# Planning & deliverables

## Planning

### Week 1

* Projectplan
  + Stepping stones
    - Diagnose
    - Doel
    - Resultaat
* Domeinanalyse

### Week 2

Eerst businessvragen opstellen en opgeleverde data opschonen en integreren voordat we externe data verzamelen, opschonen en integreren.

#### Businessvragen opstellen (Akasha, Bram| Woensdag)

* Business-assets in kaart brengen (Akasha)
* KPI’s in kaart brengen (Akasha)
* Business-vragen opstellen (Bram)

#### Opgeleverde data opschonen en integreren (Stan, Tugay, Bram| Woensdag)

* SQL-server en database opstellen (Stan | Maandag)
  + Anders bestand delen met elkaar
    - SharePoint/OneDrive mogelijk?
* Data van databases integreren
* Opgeleverde data kwaliteit controleren en opschonen

#### Externe daten verzamelen, opschonen en integreren (Mehmet | Vrijdag)

* Data verzamelen van externe bronnen
* Externe data opschonen
* Externe data integreren

#### Overig

* GitHub-repo opstellen en delen met onszelf + Henry Robben

### Week 3

* Eventueel herstellen database
* Filteren en sorteren van data
* Kostenraming
* Voorbereiden voor presentatie

### Week 4

* Machine learning implementeren
* Samenvoegen databronnen (ETL)
* Omgeving dashboard realiseren
* Dashboard realiseren

### Week 5

* Producten aanvullen waar nodig
* Rapport over kwaliteit en betrouwbaarheid dashboard

Deliverables

* Dashboard (grootste)
  + Data verzameling
    - Geleverde databases
    - Data van andere sources
  + Data warehouse
  + Machine learning
* Kostenraming
* Rapport over kwaliteit en betrouwbaarheid dashboard

Benodigde middelen

Data:

Opgeleverde data

Verzamelde data door scraping (van het web)

Software:

Microsoft Power BI

Microsoft SQL server management

Beoogde resultaten

Het dashboard voldoet wanneer het dashboard interactief is, het geheel werkt en live is. Op het dashboard staat relevante sturingsinformatie voor OutdoorFusion gepresenteerd inclusief Machine Learning uitkomsten. Het liefst gedeployed in een webomgeving. Als bronnen voor dit dashboard is de aangeleverde data gebruikt maar ook op het web gevonden data. Het dashboard is zeer breed inzetbaar en richt zich op een breed scala aan management vragen en KPI’s.