

Instructie Voortraject en Prototypetraject

AIS-DW (DATWAHxx)

Instructie voortraject en instructie prototypetraject behorende bij de ANBC casus

In deze onderwijspublicatie is geen auteursrechtelijk werk opgenomen.
copyright HAN-AIM

Inhoudsopgave

Instructie Voortraject en Prototypetraject	1
AIS-DW (DATWAHxx)	1
1. Algemeen	3
2. Voortraject.	4
Opdracht 1: Stel het operationeel informatieplan op	4
Opdracht 2: Maak het meetplan voor de belangrijkste operationele KPI's	5
Opdracht 3: Beschrijf per geselecteerde operationele KPI de mogelijke maatregelen (Actieplan)	5
3. Prototypetraject	5
Opdracht 1: Bronanalyse	5
Opdracht 2: Functioneel ontwerp	6
Opdracht 3: Technisch ontwerp en implementatie	6
Opdracht 4: Ontsluiting	7
Opdracht 5: Testen	7
Opdracht F: Verslaglegging	7
Inleveren en demo	7
Bijlage A: Het LRS van het informatiesysteem	8
Bijlage B: Ontsluiting van KNMI data	9

1. Algemeen

Deze casus en opdrachtomschrijving is onderdeel van de beoordeling van de course DW van het semester AIS. Deze omschrijving heeft betrekking op **twee** beroepsproducten namelijk het voortraject en het prototypetraject. In de onderstaande tabel vind je de toetscodes en vakcodes van Osiris.

AIS-DW (DATWAHxx)		
Voortraject:	B_Casus	Casus
Prototypetraject:	B_Prototype	Prototype

Het voortraject betreft een casus met als output een advies waarbij je beargumenteerd een Balanced scorecard inricht met bijbehorende ksf-en en kpi's. Je werkt dit uit tot meetplannen en actieplannen.

Het prototypetraject betreft het ontwerp en ontwikkeling van een prototype en ondersteunende documentatie. Het prototypetraject bevat een analyse van de data, het ontwerp van een datavault, datamart. Daarnaast wordt er een prototype ontwikkeld van de kpi's uit het voortraject.

De casus beschrijving vindt je een apart document op Teams. Hieronder vindt je de instructies voor de beide beroepsproducten. Voor de ontwikkeling van het prototype heb je de uitkomst van het meetplan nodig, maar verder staan beide producten los van elkaar en kun je ze los van elkaar wel of niet halen, ze worden apart beoordeeld. Indien je alleen het prototype maakt of herkanst zal de docent je een aantal KPI's of een meetplan als input geven.

Voor beide toetsen moeten deelproducten en een totaalverslag ingeleverd worden. Die verslagen bestaan uit de verschillende producten die je maakt bij de opdrachten. Van het prototype moet ook een demonstratie gegeven worden.

Voor beide beroepsproducten geldt dat je deze als tweetal uitvoert, zodat je van elkaar kunt leren. Door het reviewen van elkaars werk kom je tot een beter eindresultaat. In de verslagen moet duidelijk aangegeven worden wat de persoonlijke bijdrage is geweest van elke student. Bij beide toetsen hoort mogelijk ook een individueel assessment indien er vragen zijn over de producten of de persoonlijke bijdrage.

Voor de oplevermomenten van de beroepsproducten en herkansingen, raadpleeg het toetsrooster.

Voor deze opdracht heb je de als uitgangspunt de opdrachtendatabase nodig. Een backup van deze database alsmede andere bestanden en vind je op Teams. Ook de casusbeschrijving vind je daar.

2. Voortraject.

Hoewel de processen ondersteund worden door applicaties en een database is dat voor het operationeel informatieplan niet zo van belang. Bij de keuze van uit te werken KPI's kunnen de al beschikbare gegevens wel een rol spelen.

Opdracht 1: Stel het operationeel informatieplan op

Stel het operationeel informatieplan op voor het bedrijf. Gebruik hiervoor de stappen zoals die hieronder beschreven zijn.

1A: Schrijf een geloofwaardige visie en missie voor ANBC op.

- Maak hierbij gebruik van het boek van Jager en van Slooten.
- Geef een toelichting waaruit blijkt dat de missie en visie op elkaar aansluiten.
- De visie en missie mag niet langer zijn dan een 1 A4.

1B: Stel een schematisch procesmodel op van de onderstaande operationele bedrijfsprocessen. Dit mag op een manier doen die je zelf praktisch vindt; via BPMN, data flow diagram, activity diagram, de methode van Kerklaan etc. Maak gebruik van de interviews die je met de directie hebt gehouden.

Geef per proces in ieder geval de input, proces, output en effect weer. Zoals besproken in de lessen van de eerste week.

Richt je op de volgende operationele bedrijfsprocessen:

- Het primaire proces(sen) rond **Afhalen en bezorgen**
 - o Dit proces is beschreven in de casus
- Het primaire proces rond **Inroosteren van de koeriers**
 - o Dit proces dien je zelf helder te krijgen via elicitatiegesprekken met de docent
 - o Denk hierbij ook aan foutief geleverde bestellingen
- Het ondersteunende proces(sen) rond **Inroosteren van de bemensing van de centrale en het roosteren van de bikers die paraat moeten zijn.**
 - o Dit proces dien je zelf helder te krijgen via elicitatiegesprekken met de docent

1B: Maak een performancemanagementsysteem gebaseerd op de balanced score card. Laat je inspireren door het boek van Jager en van Slooten.

- Stel een tabel op met mogelijke eisen en wensen (KSF-en) en/of PI's per proces en per actor (actoren tabel). Wees hierin enigszins selectief en baseer je op de problemen gesignaleerd in de casusbeschrijving en elicitatiegesprekken. Bij alle KSF-en en daarbij horende PI's geef je een korte toelichting. Beschrijf daarbij ook kort waarom ze van belang zijn. Neem deze toelichting niet op in je tabel maar doe dit bij voorkeur los. Zeker als een KSF bij meerdere processen en actoren voor komt, kan de tabel anders erg onoverzichtelijk en lang worden.
Dit is de brainstormfase, hierna ga je ordenen.
- Orden de KSFen en PI's onder input, proces, output en effect (concept Informatieplan). Beperk je tot de 8-12 belangrijkste KSF-en per proces en leg een relatie met de elicitatiegesprekken.
- Orden de PI's ook onder de KSF-en en maak de tabel compleet (geordend Informatieplan met per KSFen de PI's). Bepaal welke PI's kritisch zijn (KPI's).
- Laat de door jouw opstelde KSF-en en KPI's zien in een BSC. Is het gebalanceerd? Zo nee ga een paar stappen terug en vul aan.

De brondatabase kan je goede ideeën geven bij deze uitwerkingen, maar het verbeteren van de processen is uitgangspunt: het kan niet anders of je komt PI's tegen waarvoor de gegevens niet in de huidige database staan. Je hoeft in deze fase dus geen rekening te houden met gegevens die wel of niet beschikbaar zijn. Belangrijk is om te achterhalen welke stuurinformatie gewenst is.

Opdracht 2: Maak het meetplan voor de belangrijkste operationele KPI's

- Bespreek de gevonden KPI's met je docent. Selecteer met criteria de zes meest belovende KPI's.
- Zorg dat je in ieder geval vier KPI's hebt die met de gegevens uit het bronsysteem te realiseren zijn. Maak in ieder geval de omzet inzichtelijk
- Stel voor elk geselecteerde KPI een meetplan op. Maak hierbij gebruik van de voorbeelden uit de presentaties van de lessen en het uittreksel uit het boek van Kerklaan (beschikbaar op teams en behandeld in de les).
 - o Geef duidelijk aan welke gegevens je nodig hebt voor deze KPI's.
 - o Geef een voorbeeldgrafiek met een verantwoording van de gekozen lay-out. Geef aan waarom de gekozen lay-out/grafiektype de beste keus in deze situatie is.
- Zorg dat er één KPI is die inzicht geeft in de invloed van het weer op de omzet of verkoop. Geef aan of de benodigde gegevens beschikbaar zijn in het bronsysteem. Bij twijfel: overleg met de docent en stel eventueel later je keuze bij.

Opdracht 3: Beschrijf per geselecteerde operationele KPI de mogelijke maatregelen (Actieplan)

Beschrijf per geselecteerde KPI zo compleet mogelijk de maatregelen die ondernomen kunnen worden als de KPI onvoldoende scoort.

Opdracht 4: Conclusie en Inleveren voortraject.

- Maak van de voorgaande opdrachten een net verslag. Voeg aan het verslag een heldere, op de voorgaande opdrachten gebaseerde conclusie en advies toe (maximaal een half A4).
- Het verslag is maximaal 15 A4 groot.
- Zie iSAS voor het inlevermoment en lever je verslag in.
- Ga verder met het volgende deel van de casus (prototypetraject)

3. Prototypetraject

Voor het ontwerpen en realiseren van het prototype maak je onderstaande opdrachten. Je hoeft ze niet helemaal in deze volgorde te doorlopen. Het is zelfs aan te raden om FO, TO en ontsluiting **iteratief** te doorlopen om te komen tot een beter aaneensluitend en consistent geheel.

Opdracht 1: Bronanalyse

- Bestudeer het bronsysteem, controleer of de betekenis van de kolommen duidelijk is, controleer en bespreek de *kwaliteit* van de gegevens. Bekijk of de gegevens *volledig* zijn en *geldig*. Leg je bevindingen vast.
 - o maak gebruik van de in lessen, of eerder, besproken begrippen op het gebied van datakwaliteit.
- Kijk ook naar de kwantiteit van de data.
 - o Hoeveel consignments zijn er?
 - o Hoeveel parcels zijn er?
 - o In welke tijdsperiode speelt dit zich af.
 - o Hoeveel medewerkers zijn er?
- Kijk naar de semantiek. Wat is de betekenis van parcel en consignment. Wat betekent dit voor mijn kpi's? Wordt er gebruik gemaakt van statusvelden? Wat is hiervan de betekenis en speelt het een rol bij de realisatie van het dashboard? Zijn eventuele meetwaarden additief?

- Maak een selectie op de KNMI site van de dataset die je nodig hebt en doe daarmee hetzelfde. Tips over de ontsluiting van KNMI data vind je in bijlage B.

Opdracht 2: Functioneel ontwerp

Maak een functioneel ontwerp (FO) voor een dashboard gebaseerd op kubussen en sterren.

In dit functioneel ontwerp komen aan de orde:

- Een samenvatting van de te realiseren KPI's uit je meetplan;
- Een bulplot van de functionele hiërarchieën voor de eindgebruiker;
- Wat is de functionele invloed van je data-analyse op de meetplannen die hebt gemaakt? Scherp de definitie van je KPI's aan.
- Andere aanvullende functionele requirements indien van toepassing;
- Schermontwerpen voor je dashboard. Geef aan op welke manier of manieren je de KPI's gaat visualiseren (soms wil je er één op meerdere manieren visualiseren en soms kun je er een aantal combineren), welke filters, knoppen en hiërarchieën de gebruiker tot zijn beschikking krijgt en wat de flow wordt in het dashboard. Ook dit is een uitbreiding en aanvulling op de visualisaties in het meetplan.
- Onderbouw de keuzes die je tijdens het opstellen van het FO gemaakt hebt.

Opdracht 3: Technisch ontwerp en implementatie

Let op: voordat je verder gaat moet je de casus Relational to Dimensional hebben afgerond. De realisatie van de ETL die hieronder van je gevraagd wordt is zonder praktische ervaring en inzicht in een goede werkwijze zoals in het practicum Relational to Dimensional veel meer tijd kosten.

Maak een technisch ontwerp (TO) voor dit meetsysteem en realiseer de genoemde punten. In dit technisch ontwerp komen aan de orde:

- Datawarehouse-architectuur. Denk hierbij aan: welke lagen zijn er, welke bronsystemen, welke ETL processen zijn er nodig. Indien er meerdere mogelijkheden zijn, weeg die dan tegen elkaar af. Neem ook de tooling op waarmee je de gegevens wilt ontsluiten;
- Ontwerp een Data Vault model voor het DWH waarin je *alle* informatie uit het eigen bronsysteem op kunt slaan en *alleen* de data van het KNMI die je daadwerkelijk gaat gebruiken. Er zijn meerdere oplossingen mogelijk, op sommige punten zal je een keuze moeten maken, licht die keuzes toe. Geef het ERD (Entity Relationship Diagram) van een Data Vault als CDM of PDM in Power Designer;
- De fysieke implementatie van de Data Vault (SQL script opnemen in bijlage)
- Een beschrijving van de transformatie van het conceptuele model naar een sterschema of sterschema's en een ERD van het resultaat;
- Een toelichting op het sterschema. Leg uit wat de relatie is tussen je KPI's en de meetwaarden, en welke berekeningen nodig zijn op de data om tot de meetwaarden te komen. Deze berekeningen heb je nodig tijdens het ETL proces. Leg uit wat de granulariteit (het detailniveau) is van je ster of sterren. Leg ook de keuze van de dimensies uit. Laat daarbij merken dat je de ontwerpafwegingen maakt zoals aangegeven door van der Lek. Onderbouw kortom je keuzes;
- De fysieke implementatie van de sterren (SQL script opnemen in de bijlage);
- Documentatie van de ETL processen: uit welke stappen moeten deze processen bestaan en wat is de volgorde? Welke berekeningen moeten hierin plaats vinden? Let op, dit zijn er dus minimaal drie: van de originele bron naar DWH, van NoSQL bron naar DWH, van DWH naar ster(ren). Beschrijf per stap hoe deze gerealiseerd is. Licht complexere constructies toe en neem relevante (voorbeelden van) SQL op in een bijlage en verwijst hier naar.

Tip: kijk bij de realisatie van de DV en de sterren even naar de collation van de brondatabase en gebruik dezelfde om conflicten te voorkomen tijdens de realisatie van het ETL proces.

Bijvoorbeeld:

```
CREATE DATABASE [NorthwindDV]  
COLLATE Latin1_General_CI_AS;
```

Opdracht 4: Ontsluiting

- Ontsluit de data mart(s) met Power BI en een andere dashboard tool naar keuze bv Excel, Tableau of QlikView.
- Maak screenshots waarmee je laat zien dat de realisatie conform het FO is (of niet) en neem die screenshots op in het verslag (zie volgende punt).
- Omschrijf de realisatie van het dashboard en eventuele afwijkingen van het FO. Als er afwijkingen zijn, leg dan kort uit waarom.

Opdracht 5: Testen

- In opdracht 1 heb je de kwantitatieve eigenschappen van de database onderzocht. Toon aan dat deze waarden kloppen met de waarden uit de datamart, dashboard. Het zou gek zijn als deze waarden niet overeenkomen. Laat de gevonden waarden zien in een tabel ondersteund met de screenshots.
- Controleer van twee KPI's de uitkomsten in het dashboard door deze rechtstreeks met SQL te berekenen op de **brondatabase**.
- Komen jouw waarden niet overeen? Zoek uit waardoor dit komt en pas de ETL aan.

Opdracht F: Verslaglegging

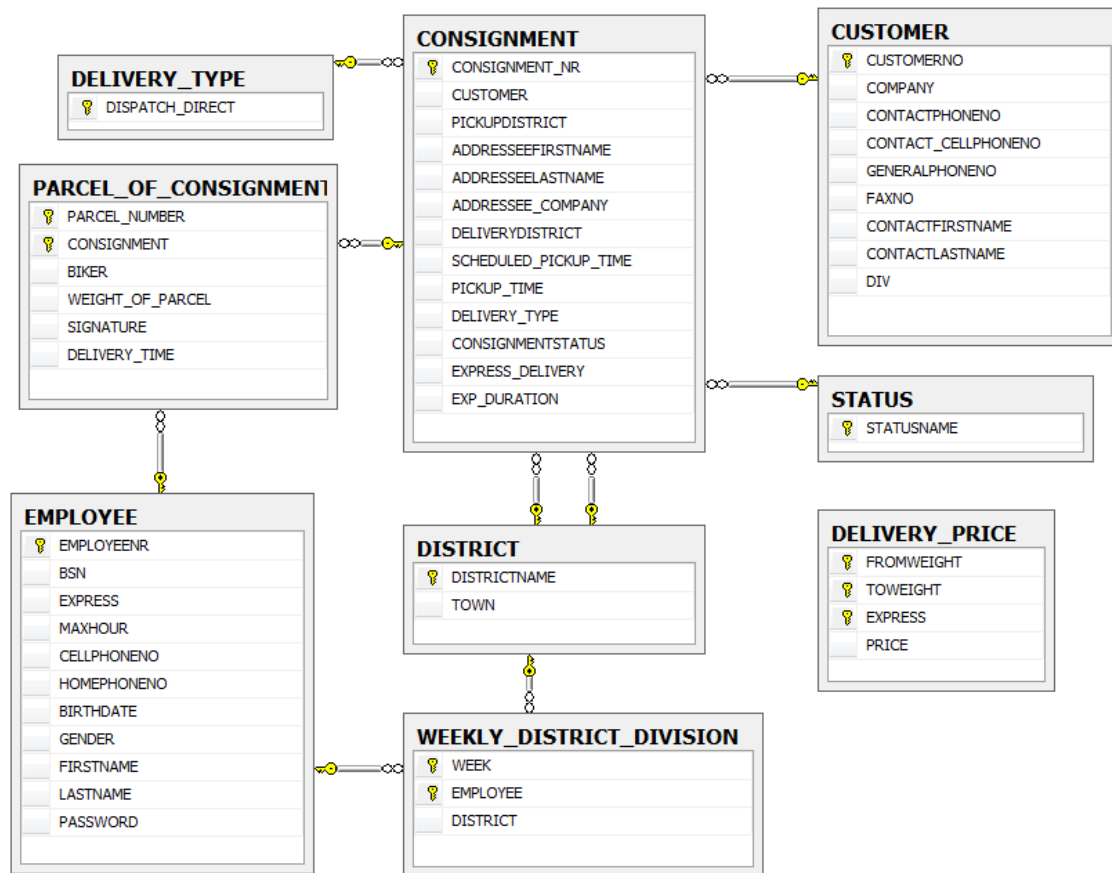
- Verwerk alle opdrachten tot één bondig verslag, waar je ze in volgorde, puntsgewijs, opneemt.

Inleveren en demo

Zie iSAS voor het inlevermoment. Je levert je verslag in en daarnaast de databases met de data vault en de ster, en de projecten van het ETL proces.

Geef een demo van je oplossing in week 8 of 9, maak hiervoor een afspraak met de docent.

Bijlage A: Het LRS van het informatiesysteem



Bijlage B: Ontsluiting van KNMI data

KNMI datasets kunnen op verschillende manieren ontsloten worden. Kies één van onderstaande opties:

1. Je download een dataset als tekstbestand en leest dat in in een database in SQL Server. Schrijf een query FOR JSON waarmee je de gegevens die je nodig hebt in een JSON file zet. Ontsluit de JSON file met SSIS. Daarvoor zijn verschillende mogelijkheden. Onderzoek die en kies er één. Je hoeft hier géén onderzoeksplan of -verslag voor op te stellen, hou het praktisch. Onderbouw alleen je keuze kort in je verslag. Je kunt hierbij gebruik maken van de klimaatdata: [KNMI - Klimatologie](#), bijvoorbeeld uurwaarden van weerstations: [KNMI - Uurgegevens van het weer in Nederland](#).
2. Je maakt gebruik van een API voor het open dataplatform. Zie voor de datasets en instructies over het gebruik van API's het KNMI Open Dataplatform: [Open Data | KNMI Dataplatform](#). Let wel op, de datasets op dit platform zijn ruwe, onbewerkte meetgegevens en het is veel lastiger om hier een bruikbare dataset te vinden dan bij de link genoemd onder optie 1.