Prototype data warehouse en business intelligence

The Cloud Consultants

# Versiebeheer

|  |  |  |  |
| --- | --- | --- | --- |
| Versie | Auteurs | Datum | Opmerking |
| 0.1 | A. van Dalen  S.A. Twardowski | 02-11-2020 | Eerste versie |
|  |  |  |  |

Inhoudsopgave

[Versiebeheer 1](#_Toc55242541)

[Introductie 3](#_Toc55242542)

[Data warehouse: definitie 4](#_Toc55242543)

[Karakteristieken 4](#_Toc55242544)

[Terminologieën 5](#_Toc55242545)

[OLTP en OLAP 5](#_Toc55242546)

[ETL 5](#_Toc55242547)

[Metadata 5](#_Toc55242548)

[Metadata repository 5](#_Toc55242549)

[Granularity 6](#_Toc55242550)

[Aggregate 6](#_Toc55242551)

[Data cube 6](#_Toc55242552)

[Data warehouse schema’s 7](#_Toc55242553)

[Star schema 7](#_Toc55242554)

[Snowflake schema 8](#_Toc55242555)

[Fact constellation schema 8](#_Toc55242556)

[Verschillen tussen een data warehouse en een operationele database 10](#_Toc55242557)

[Business intelligence (Wikipedia, 2020) 10](#_Toc55242558)

[Business intelligence proces 10](#_Toc55242559)

[Software 11](#_Toc55242560)

[Data warehouse software 11](#_Toc55242561)

[ETL software 12](#_Toc55242562)

[Datavisualisatie software 12](#_Toc55242563)

[Verwijzingen 13](#_Toc55242564)

# Introductie

Na het uitvoeren van onderzoek naar data warehousing en business intelligence wordt er een prototype gemaakt om aan te tonen welke waarde deze technieken kunnen bieden voor de organisatie. Er zal worden toegelicht welke keuzes zijn gemaakt voor de gebruikte data, de business intelligence vragen, het gekozen model wat hierbij past en de gebruikte software om het prototype te bouwen.

## Brondata

Om een goede voorstelling te maken van datawarehousing en business intelligence is het van belang om een realistische dataset te gebruiken die overeenkomt met een typisch bedrijf. De gekozen dataset voor dit prototype is de AdventureWorks dataset van Microsoft.

De AdventureWorks dataset is speciaal ontwikkeld voor leer- en prototype doeleinden en geeft een realistische voorstelling van een typische database welke binnen een bedrijf wordt gebruikt. De dataset bevat data over bijvoorbeeld personeel, winkels, klanten, producten, en orders. Tevens bevat de dataset data over verschillende jaren waardoor deze data goed is in te zetten voor een prototype van een datawarehouse.

De AdventureWorks dataset is in verschillende formaten te verkrijgen waaronder een operationele database en een datawarehouse. Voor het prototype maken wij gebruik van de operationele database als brondata. Deze data wordt met behulp van ETL tools getransformeerd en geïmporteerd in een nieuwe datawarehouse.

## Business intelligence vragen

Om een goed beeld te krijgen wat business intelligence voor de organisatie kan betekenen is het van belang de juiste informatie vragen te stellen. Wat van belang is bij het opzetten van de vragen is dat de brondata aansluit bij de te verkrijgen informatie. Wanneer de bron niet over de juiste data beschikt moet de vraag worden aangepast of moet er meer brondata worden verzameld welke aansluit bij de informatiebehoefte van de organisatie. Tevens moeten ook de vragen natuurlijk aansluiten bij de informatiebehoefte van de organisatie.

Om een goed beeld te geven wat business intelligence kan betekenen voor de organisatie hebben wij de volgende vragen opgesteld welke nieuwe inzichten kunnen geven aan het management:

***Vraag 1: Welke productcategorieën worden per periode het meest verkocht per locatie?***

Deze business vraag zal inzicht geven in welke producten populair zijn op bepaalde periodes en bepaalde locaties. Deze vraag bevat meerdere dimensies: productcategorie, tijd en plaats.

De tweede business vraag die we zullen behandelen is:

***Vraag 2: Wat is de omzet per winkel, per periode per categorie?***

Deze informatie kan inzicht geven in welke winkels bepaalde categorieën producten afnemen en hier bijvoorbeeld kortingsacties of marketing op aanpassen. Ook deze vraag bevat meerdere dimensies: winkel, tijd en productcategorie.

# Verwijzingen

Fontys University of Applied Sciences. (2020, 10 29). *Data Warehouse (DW) and Business Intelligence (BI) - what is it?* Opgehaald van fhict.infrastructure.com: https://fhict.instructure.com/courses/10205/pages/data-warehouse-dw-and-business-intelligence-bi-what-is-it?module\_item\_id=531842

Fontys University of Applied Sciences. (2020, 10 29). *Data Warehousing - Schemas*. Opgehaald van fhict.instructure.com: https://fhict.instructure.com/courses/10205/pages/data-warehousing-schemas?module\_item\_id=531846

Fontys University of Applied Sciences. (2020, 10 29). *Data Warehousing - Terminologies*. Opgehaald van fhict.infrastructure.com: https://fhict.instructure.com/courses/10205/pages/data-warehousing-terminologies?module\_item\_id=531844

Stichdata. (2020, 10 29). *OLTP and OLAP: a practical comparison*. Opgehaald van stichdata.com: https://www.stitchdata.com/resources/oltp-vs-olap/

Wikipedia. (2020, 10 29). *Aggregate (data warehouse)*. Opgehaald van wikipedia.com: https://en.wikipedia.org/wiki/Aggregate\_(data\_warehouse)

Wikipedia. (2020, 10 29). *Business intelligence*. Opgehaald van wikipedia.com: https://en.wikipedia.org/wiki/Business\_intelligence

Wikipedia. (2020, 10 3). *Data warehouse*. Opgehaald van Wikipedia.com: https://en.wikipedia.org/wiki/Data\_warehouse

Wikipedia. (2020, 10 29). *Granularity*. Opgehaald van wikipedia.com: https://en.wikipedia.org/wiki/Granularity

Wikipedia. (2020, 10 29). *Metadata*. Opgehaald van Wikipedia.com: https://en.wikipedia.org/wiki/Metadata