

KULIAH CATATAN

7023T Lanjutan Sistem Database

sesi 05

Dimensi Modeling 1

HASIL PEMBELAJARAN

- Peserta diharapkan Mampu Memahami Tahapan Proses *pemodelan dimensi*
- Peserta diharapkan DAPAT menjelaskan Hal-Hal Yang Perlu dipersiapkan *hearts* Tahap *persiapan* Proses PADA *pemodelan dimensi*.
- Peserta diharapkan Mampu Memahami Tahap *pengembangan model iteratif* Proses PADA *pemodelan dimensi*.
- Peserta diharapkan DAPAT menjelaskan bagaimana Proses *ulasan* Dan *pengesahan* PADA Proses *pemodelan dimensi*.

GARIS MATERI (Sub-Topik):

1. *Ikhtisar proses*
2. *Persiapan*
3. *Pengembangan Model berulang*
4. *Review dan Validasi*

Ikhtisar proses

Proses desain *Model dimensi* dilakukan hearts Tiga Tahap Yaitu *persiapan*, *pengembangan model berulang*, Dan *ulasan dan validasi*. PADA Tahap *persiapan* Dibuat diagram *skema bintang* berdasarkan *matriks bus perusahaan*. Diagram Suami akan gandum menjelaskan Dari tabel Fakta, Atribut Dan metrik Yang akan digunakan sebagai *mengukur*. Diagram also Suami akan menjelaskan tabel dimensi Beserta atributnya. Selanjutnya PADA Tahap *Model iteratif pengembangan* dilakukan Pemetaan Yang menunjukkan tabel-tabel PADA Sumber Data Yang akan digunakan untuk review membentuk tabel Fakta maupun tabel dimensi PADA Sistem DW / BI. Proses Pengembangan Model dilakukan Beroperasi iteratif tabel demi tabel, DENGAN analisis Yang PADA rinci Semakin SETIAP iterasinya. Model Yang dihasilkan selanjutnya akan direview Dan divalidasi Oleh Perwakilan Dari pengguna Bisnis maupun IT untuk review meyakinkan desain Yang dihasilkan memenuhi Kebutuhan.

Persiapan

Tahap *persiapan* ATAU Persiapan untuk review *pemodelan dimensi* terdiri Dari beberapa fase, Yaitu:

- Mengidentifikasi tim Yang akan dibentuk Serta siapa partisipannya, seperti dijelaskan PADA tabel 1. Mayoritas Pekerjaan akan dilakukan Oleh tim inti, Yang terdiri Dari *data modeller*, *analisis bisnis* Dan ANGGOTA tim ETL.
- Melakukan tinjauan Terhadap Dokumen Kebutuhan Bisnis, Mulai dari *dimensi modeling*.
- Melakukan tinjauan Terhadap Sumber data, DAPAT dilakukan DENGAN mengeksekusi Cara *pertanyaan sederhana* ATAU memanfaatkan Perangkat Lunak *Data profil*. fase Suami bertujuan untuk review meningkatkan pemahaman Terhadap karakteristik Sumber data.
- Mempersiapkan Lingkungan yaang akan digunakan untuk review memodelkan dimensi. Sketsa Awal Dari Model DAPAT disajikan hearts Bentuk *spreadsheet* seperti diilustrasikan PADA tabel 2. Pendekatan Suami akan Model menghasilkan Penyanyi DAPAT dikembangkan Beroperasi iteratif Beroperasi Mudah.
- Menentukan Aturan penamaan tabel.

Tabel 1. Partisipan Dan peranannya hearts *pemodelan dimensi*

Peserta	Tujuan / Peran dalam Proses Modeling
Data modeler	Tanggung jawab utama desain, fasilitator
kuasa pengguna	kebutuhan bisnis, pakar sumber, definisi bisnis
analisis Bisnis	analisis bisnis dan pakar sumber, definisi bisnis
Data steward	Perjanjian drive pada nama perusahaan, definisi, dan aturan
<u>source pengembang sistem</u>	Sumber ahli, bisnis aturan DBA
	bimbingan desain, pembelajaran awal
ETL arsitek dan pengembang	pembelajaran awal
<u>arsitek BI dan pengembang pers</u>	syarat aplikasi BI, awal sopir belajar Bisnis atau
komite pengarah governance	Penamaan dan definisi bisnis masalah resolusi, validasi model

Tabel 2. Contoh sketsa Awal pemodelan dimensi

Atribut Nama Deskripsi		Nama alternatif	Nilai sampel
<u>Penawaran Khusus ID</u>	Sumber sistem kunci Penawaran		
Khusus Nama	Nama / deskripsi Penawaran Khusus	Nama promosi, Penawaran khusus deskripsi	Volume Diskon 11 sampai 14; Jatuh <u>diskon 2006</u>
<u>Persen diskon</u>	Persen item diskon		
Penawaran Khusus Type	Deskripsi dari jenis promosi, penawaran khusus atau diskon.	promosi Jenis	V olume Diskon; Produk dihentikan
Penawaran Khusus Kategori	Channel yang Promosi berlaku	promosi Kategori	Reseller; Pelanggan
Mulai tanggal	hari pertama promosi tersedia		2008/06/15
Tanggal akhir	hari terakhir promosi tersedia		2008/12/31
minimum Quantity	kuantitas minimum yang diperlukan <u>untuk memenuhi syarat untuk promosi</u>		0
maksimum Jumlah	kuantitas maksimum yang diizinkan di bawah promosi		BATAL

Pengembangan Model Awal dimulai model diagram DENGAN Yang *dimensi*. diagram Suami merepresentasikan tabel Fakta Dan Dimensi Proses untuk review Bisnis yang dipilih. diagram Suami kadang-kadang disebut sebagai *gelembung-chart*. SETIAP tabel Fakta Dari Proses Bisnis Tertentu Perlu digambarkan hearts diagram terpisah. Peletakan tabel-tabel dimensi disarankan Sesuai DENGAN Tingkat kepentingannya, Dimana dimensi Yang Penting diletakkan Bagian PADA differences seperti diilustrasikan PADA gambar 1.



Gambar 1. diagram contoh *level tinggi*

Pengembangan Model berulang

Proses *pengembangan model iteratif* terdiri Dari Empat Langkah, yakni:

1. Bisnis Memilih Proses - using perusahaan matriks data warehouse bus,

Dimana SETIAP baris Dari matriks merepresentasikan Proses bisnisnya. Proses Pemilihan Bisnis dilakukan Beroperasi iteratif, Satu Proses hearts Satu iterasi.

2. Mententukan gandum Dari tabel Fakta - butir DAPAT digunakan untuk review Menjawab

Pertanyaan "APA Yang direpresentasikan Oleh tiap baris Dari tabel Fakta?". Gandum menyatakan Tingkat data detail Dari Yang Disimpan PADA tabel Fakta. Semakin Tinggi

Tingkat rinci akan Semakin Baik. Melewatkan Langkah Suami merupakan Kesalahan desain Yang Paling Sering Terjadi.

3. Mengidentifikasi tabel-tabel dimensi - jumlah minimum Dari tabel dimensi

mengikuti gandum Yang dipilih PADA tabel Fakta. Dimensi Lain DAPAT ditambahkan kemudian, namun Perlu memiliki Tingkat gandum Yang sama DENGAN tabel Fakta.

4. Mengidentifikasi tabel Fakta - biasanya diturunkan Dari hasil temuan Pengukuran. * Semua

Fakta Harus memiliki Tingkat gandum Yang sama.

Memulai DENGAN tabel dimensi Adalah pendekatan Yang Paling Mudah, contohnya

dimensi *tanggal*. Lakukan desain Satu ATAU dua tabel untuk review SETIAP sesinya, Hal Penyanyi dilakukan Model PADA untuk review menghindarkan kelihangan rinci *level tinggi*.

Identifikasi Sumber Data Terbaik Yang akan digunakan untuk review Mengisi sasaran tabel, baik tabel Fakta maupun tabel-tabel dimensi. Mulai DENGAN daftar menyusun Dari SEMUA Kandidat Data Sumber; sebagian Kandidat mungkin Saja Berasal Dari Dokumen *kebutuhan bisnis*, Dan sisanya mungkin Berasal Dari Personil IT. Selanjutnya lakukan analisis Dan *profil* Terhadap SETIAP Kandidat. Kemudian Pilih Sumber Data Terbaik Yang akan hearts digunakan dan model yang dokumentasikan Alasan Pemilihan tersebut.

Table Name:

DimOrderInfo

Table Type:

Dimension

View Name:

OrderInfo

Description:

OrderInfo is the "junk" dimension that includes miscellaneous information about the Order transaction

Used in schemas:

Orders

Generate script?

Y

Target										Source							
Column Name	Description	Datatype	Size	Key?	FK To	NULL?	Default Value	Unknown Member	Example Values	SCD Type	Source System	Source Schema	Source Table	Source Field Name	Source Datatype	ETL Rules	Comments
OrderInfoKey	Surrogate primary key	smallint			PK ID	N		-1	1, 2, 3, 4...		ETL Process					Standard surrogate key	
BKSalesReasonID	Sales reason ID from source system	smallint				N		-1			O/EI	Sales	SalesReason	SalesReasonID	int	Convert to char, left-pad with zero, R for reseller row.	We need to insert a single row for Reseller
Channel	Sales channel	char	8					Unknown	Reseller, Internet, Field Sales	1	O/EI	Sales	SalesReason	Derived		"Internet" for real sales reasons, "Reseller" for reseller row.	
SalesReason	Reason for the sale, as reported by the customer	varchar	30					Unknown		1	O/EI	Sales	SalesReason	Name	nvarchar(50)	Convert to varchar, "Reseller" for reseller row.	
SalesReasonType	Type of sales reason	char	10					Unknown	Marketing, Promotion, Other	1	O/EI	Sales	SalesReason	ReasonType	nvarchar(50)	Convert to varchar, "Reseller" for reseller row.	
AuditKey	What process loaded this row?	int			FK Audit Dim	N		-1		1	Derived					Populated by ETL system using standard technique	

Comments

Order_Info is a "junk" dimension with only a handful of rows based on "Channel" and "Sales Reason". We currently have only three channels and sales reasons only for field sales and internet sales. We can eliminate a dimension by combining these two.

Gambar 2. Model dimensi contoh desain rinci

Identifikasi dimensi-dimensi Yang DAPAT Bersama digunakan (*dimensi sesuai*)

Serta tabel Fakta Dasar Dan turunannya. Tabel Fakta aditif DAPAT dihasilkan Dari perhitungan PADA tabel Fakta lain untuk review baris Yang sama. Apabila kitd memiliki Banyak tabel Fakta turunan, disarankan untuk review mendokumentasikannya PADA Sebuah worksheet. Rancangan detail

Dari SETIAP tabel Fakta Dan dimensi Perlu didokumentasikan hearts *desain rinci worksheet* Seperti Yang diperlihatkan PADA Gambar 2. Proses desain Beroperasi rinci meningkatkan pemahaman Bisnis kitd Terhadap Proses Dan DAPAT Menjadi PANDUAN Saat diperlukan perubahan Berlangganan Proses Bisnis, tabel Fakta, maupun tabel dimensi.

Review dan Validasi

Model dimensi Perlu direview Oleh beberapa parties DENGAN Sudut pandang Dan keahliannya masing-masing. Parties Yang Pertama kali disarankan untuk review melakukan tinjauan Adalah Yang Berasal Dari DBA (database administrator) ATAU *pengembang sistem* Dari Departemen IT. Mereka memiliki pemahaman yang Baik Terhadap Sistem transaksional namun biasanya TIDAK memiliki Pengetahuan Yang mencukupi Mengenai *Model dimensi*. Oleh KARENA ITU Perlu diberikan tutorial Singkat ditunjukan kepada mereka, Proses agar tinjauan DAPAT Berjalan Sesuai DENGAN Harapan. Tinjau berikutnya disarankan dilakukan Oleh pengguna Dari Kalangan Bisnis yang TIDAK terlibat Langsung DENGAN Proses desain. Ulasan Suami also merupakan SESI edukatif Bagi mereka, DENGAN Cara menunjukkan beberapa contoh sederhana Yang DAPAT mengilustrasikan kemampuan analitik Model Dari *dimensi*. Tujuannya Adalah Untuk melakukan Verifikasi bahwa model yang DAPAT Menjawab Pertanyaan-Pertanyaan Berlangganan Proses Bisnis.

Penghasilan kena pajak desain direview Dan divalidasi, Langkah berikutnya Adalah mendokumentasikan desain tersebut. Dokumen rinci desain DAPAT akhir disajikan hearts Bentuk *lembar kerja* (Seperti Yang Sudah diilustrasikan PADA gambar 2) Yang DAPAT Menjadi dasar dasar Dokumen untuk review *Sumber-to-sasaran pemetaan*, Yang menjelaskan Dari mana SETIAP Kolom PADA Target tabel memperoleh Sumber datanya. *Sumber-to-sasaran pemetaan* merupakan shalat Dokumen Yang itub Bagi tim ETL. Beberapa ASPEK Penting Dari Proses *dimensi pemodelan* dirangkum PADA gambar 3.

Terdapat beberapa teknik selain e *pemodelan hubungan ntity* Dan *dimensi pemodelan* Data memodelkan Untuk. Alternatifnya seperti *subjek pemodelan*, *pemodelan domain*, Bahkan *kualifikasi matriks / pemodelan*, *pemodelan transisi negara*, Dan *pemodelan arus informasi*. *pemodelan kualifikasi Fakta* teknik merupakan Yang berguna untu get Kebutuhan *granularity / pelaporan*. Masing-masing model yang memfokuskan PADA JENIS information Yang BERBEDA. Terdapat Lebih Dari Satu pendekatan untuk review get Kebutuhan Bisnis, Dan SETIAP

Mengelola Usaha dan Mengurangi Risiko

Ikuti praktik terbaik ini untuk menjaga upaya pemodelan dimensi dan risiko di cek:

- Menjamin pemodel data yang memimpin proses desain adalah modeler dimensi ahli. Jika tidak, pertimbangkan melengkapi dengan sumber daya luar.
- Setiap anggota tim desain harus secara menyeluruh dan benar-benar memahami kebutuhan bisnis.
- Pastikan untuk menyertakan kuasa pengguna sebagai bagian dari tim desain.
- Terus menyelidiki sumber data yang diusulkan dengan data Anda profil alat untuk menjamin data yang dibutuhkan untuk mendukung model data yang diusulkan tersedia.

menjamin Kualitas

langkah kunci untuk menjamin kualitas meliputi:

- Patuhi pemodelan praktik terbaik dimensi.
- Bersikeras partisipasi aktif dari kuasa pengguna dalam proses desain.
- Melakukan profiling data yang ekstensif.
- Jangan melewatkan IT dan pengguna sesi review.

Peran kunci

peran kunci untuk merancang model dimensi meliputi:

- Data modeler memimpin upaya pemodelan dimensi.
- Analis bisnis, kuasa pengguna, dan pengembang aplikasi BI mewakili kebutuhan analitik pengguna bisnis.
- pelayan Data membantu mendorong kesepakatan organisasi pada nama model dimensi ini, definisi, dan aturan bisnis.
- ahli source sistem membawa pengetahuan tentang sistem operasional.
- Tim ETL belajar tentang sumber-sumber, target, dan mendapat rasa angkat berat mereka harus lakukan untuk mengkonversi dari satu ke yang lain.
- pihak yang berkepentingan dalam TI dan komunitas bisnis yang lebih luas akan meninjau dan memberikan umpan balik pada desain.
- Vendor database dan akses data alat memberikan prinsip-prinsip desain database untuk mengoptimalkan produk mereka.

Deliverables kunci

iriman kunci untuk merancang model dimensi meliputi:

- tingkat tinggi diagram model
- Atribut dan metrik daftar
- Rinci dimensi desain lembar kerja
- daftar masalah

Gambar 3. Cetak biru untuk untuk review tindakan Proses *pengolahan dimensi*



People
Innovation
Excellence

SIMPULAN

- Penentuan *gandum* Dari tabel Fakta merupakan shalat Satu Kesalahan Yang memucat Sering Proses Saat dilakukan *pemodelan dimensi*.
- Salah Satu Dokumen Penting Yang Dari dihasilkan Proses *pemodelan dimensi* Adalah *Sumber-to-sasaran pemetaan* Yang akan Menjadi PANDUAN Bagi tim ETL hearts tugasnya melaksanakan.
- Perlu diadakan tutorial Singkat Mengenai *Model dimensi* ditunjukan kepada para pengguna Dari Kalangan Bisnis maupun Personil IT agar Proses *ulasan* Dan *pengesahan* DAPAT Berjalan Sesuai DENGAN tujuan.

DAFTAR PUSTAKA

1. Kimball, R. (2008). *Data Warehouse Lifecycle Toolkit*. John Wiley & Sons.
2. Kimball, R., & Ross, M. (2011). *Data Warehouse Toolkit: Panduan Lengkap untuk Modeling dimensi*. John Wiley & Sons.
3. Inmon, WH (2005). *Membangun Data Warehouse*. John Wiley & anak.