

PERANCANGAN DATA WAREHOUSE DAN PENERAPAN METODE ASSOCIATION UNTUK EKSTRAKSI DATA PENJUALAN PADA PT. CNI PALEMBANG

NurArifHidayat (Ndsarif@gmail.com), Nurtriana (Nurtriiian@gmail.com)

Mardiani, S.Si, M.T.I. (Mardiani@stmik-mdp.net)

Jurusan Teknik Informatika

STMIK MDP

Abstrak: Data warehouse merupakan sebuah cara yang digunakan untuk menampung data dalam jumlah besar, untuk membangun sebuah data warehouse dibutuhkan sebuah tools *microsoft SQL server 2008*, *microsoft SQL business intelligence development* dan *WEKA (Waikato Environment for Knowledge Analysis)*. Dengan menggunakan tools *microsoft SQL server 2008*, *microsoft SQL business intelligence development* dan *WEKA (Waikato Environment for Knowledge Analysis)* dapat dibangun sebuah data warehouse penjualan dengan mengumpulkan data-data penjualan meliputi *payment*, *customers id*, *gender*, *age*, *occupation*, *product id*, *product name*, *product cost*, dan *agency id* yang dimasukan kedalam sebuah *Ms. Excel* kemudian hasil dari *Ms. Excel* dimasukan ke dalam *database microsoft SQL server 2008*, setelah data masuk ke dalam *database* kemudian data tersebut di *load* kedalam *microsoft SQL business intelligence development* untuk dianalisis kemudian dimining menggunakan *WEKA (Waikato Environment for Knowledge Analysis)*. Hasil analisis data warehouse dan data mining yang dibuat dapat memberikan informasi yang lebih jelas untuk pihak PT. CNI dan mempermudah dalam menganalisis barang-barang yang berpasangan dan *association rule* dari pasangan barang tersebut.

Kata kunci: *Data warehouse*, *data mining*, *microsoft SQL business intelligence development*, *microsoft SQL server 2008*, *WEKA*, *database*, penjualan.

1 PENDAHULUAN

Perkembangan dunia bisnis yang sangat pesat, mendorong terbentuknya suatu timbunan data-data yang berukuran sangat besar. Informasi yang didapat dari data dalam bentuk tabel-tabel tersebut kurang dapat dimanfaatkan oleh pihak manajemen perusahaan dalam pengambilan keputusan untuk kemajuan perusahaan. Oleh karena itu perlu adanya aktivitas penggalian (ekstraksi) data yang masih tersembunyi untuk selanjutnya diolah menjadi pengetahuan yang bermanfaat dalam pengambilan keputusan.

PT.CNI Palembang memiliki volume produk dan transaksi yang cukup besar, perusahaan ini juga memiliki beberapa cabang

lokasi sehingga menyulitkan perusahaan untuk mengetahui produk mana yang lebih cepat laku dan kurang diminati pelanggan per daerah.

2 LANDASAN TEORI

2.1 Data Warehouse

Menurut Feri Sulianta dan Dominikus dalam buku *Data Mining Meramalkan Bisnis Perusahaan* (2010, hal. 32), *data warehouse* adalah kumpulan data dari berbagai sumber yang ditempatkan menjadi satu dalam tempat penyimpanan berukuran besar lalu diproses menjadi bentuk penyimpanan multi-dimensional dan didesain untuk *querying* dan *reporting*.

Karakteristik *Data Warehouse* :

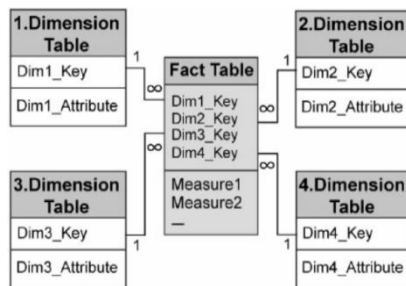
- Subject-oriented* (Berorientasi Subjek)
- Integrated* (Terintegrasi)
- Time Variant* (Rentang Waktu)
- Nonvolatile*

2.2 Dimensional Modelling

Beberapa konsep pemodelan data *Warehouse* pada *dimensionality modeling* yang dikenal pada umumnya, konsep-konsep tersebut adalah *star schema*, *snowflake* dan *fact constellation schema*.

2.2.1 Star Schema (Skema Bintang)

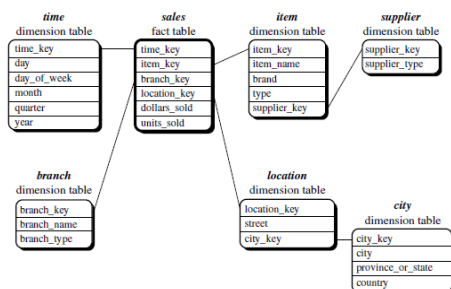
Skema bintang adalah sebuah logical struktur yang mempunyai sebuah table fakta berisi data terbaru di tengah, yang dikelilingi table dimensi yang berisi data referensi.



Gambar 1: Star Schema

2.2.2 SnowflakeSchema

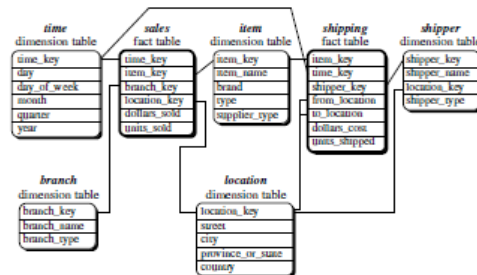
Skema kepingan salju adalah metode normalisasi table dimensi dalam skema bintang ketika anda benar-benar menormalkan semua table dimensi, struktur yang dihasilkan menyerupai kepingan salju dengan table fakta ditengah. Dapat dilihat pada gambar 2.5 (Ponniah 2011, h.259).



Gambar 2: Snowflake Schema

2.2.3 Fact Constellation Schema

Fact constellation schema adalah skema multi dimensional yang berisikan lebih dari satu tabel fakta yang saling berbagi tabel dimensi.



Gambar 3: Constellation Schema

2.3 Data Mining

Menurut Vercellis (2009, h.77) *Data mining* adalah sebuah proses berulang bertujuan untuk menganalisa *database* yang besar dengan tujuan penggalian informasi dan pengetahuan yang berkaitan dengan pengambilan keputusan dan pemecahan masalah.

2.4 ETL (Extract, Transfrom, Load)

a. Extraction

Menurut Golfarelli dan Rizzi (2009), *Extraction* adalah pengambilan data yang *relevant* atau berkaitan dari sumber data. *Extract* merupakan proses yang pertama kali dilakukan dalam pengisian data *warehouse*.

b. Cleansing

Proses *cleansing* sangatlah penting dalam sistem data *warehouse* untuk meningkatkan kualitas data, karena normalnya data dari sumber data berkualitas rendah.

c. Transformation

Transformation mengubah format data dari sumber data operasional menjadi format data *warehouse* yang lebih spesifik.

2.5 Algoritma Apriori

Analisis asosiasi atau *association rule mining* adalah teknik *data mining* untuk menemukan aturan asosiatif antara suatu kombinasi *item*. Contoh aturan asosiatif dari analisis pembelian disuatu pasar swalayan adalah dapat diketahuinya berapa besar kemungkinan seorang pelanggan membeli roti bersamaan dengan susu.

Adapun tahap mencari kombinasi *item* yang memenuhi syarat minimum *support* dari nilai *support* dalam *database*. Nilai *support* sebuah *item* diperoleh dengan rumus berikut :

$$Support(A) = \frac{JumlahtransaksimengandungA}{TotalTransaksi}$$

Sementara itu, nilai *support* dari 2 *item* diperoleh dari rumus 2 berikut:

$$\begin{aligned} Support(A, B) &= P(A \cap B) \\ Support(A, B) &= \frac{\sum TransaksimengandungAdanB}{\sum Transaksi} \end{aligned}$$

Aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif $A \rightarrow B$.

Nilai *confidence* dari aturan A B di peroleh dari rumus beikut :

$$\begin{aligned} Confidence &= P(B|A) \\ &= \frac{\sum TransaksimengandungAdanB}{\sum TransaksimengandungA} \end{aligned}$$

2.6 Weka (Waikato Environment for Knowledge Analysis)

Menurut (Feri Sulianta, hal.63) *Weka* merupakan aplikasi *datamining* yang berbasis *open source* (GPL) dan *berengine Java*. Pengertian *Weka* menurut situs resminya adalah sebagai berikut : *Weka* adalah koleksi dari algoritma pembelajaran mesin untuk tugas-tugas *data mining*. Algoritma dapat diterapkan secara langsung ke data set atau disebut kode *Java*. *Weka* berisi alat untuk data *pre-processing*,

classification, regresstion, clustering, association rules, and visualization. Hal ini juga cocok untuk mengembangkan skema pembelajaran mesin baru.

2.7 Nine-step Methodology

Nine-step Methodology (Connolly dan Begg, 2005, h.1187). Kesembilan tahap itu yaitu:

1. Pemilihan Proses
2. Pemilihan *Grain*
3. Identifikasi dan penyesuaian
4. Pemilihan Fakta
5. Penyimpanan *pre-calculation* di tabel
6. Memastikan tabel dimensi
7. Pemilihan durasi *database*
8. Melacak perubahan dari dimensi secara perlahan
9. Penentuan prioritas dan model *query*

2.8 CRISP-DM (Cross Industry Standard Process for Data Mining)

Dalam penerapan *Data Mining* digunakan metodologi *CRISP-DM* (*Cross Industry Standard Process for Data Mining*) (Olson & Delen, 2008).

1. *Business understanding*
2. *Data understanding*
3. *Data preparation*
4. *Modeling*
5. *Evaluation*
6. *Deployment*

3 PERANCANGAN DATA WAREHOUSE

3.1 Perancangan Data Warehouse

Perancangan *Data Warehouse* pada PT. CNI Palembang menggunakan metodologi sembilan tahapan (*nine-step methodology*) atau biasa disebut dengan metode Ralph Kimbal, metode ini meliputi :

a. Pemilihan Proses

Proses ini meliputi pendataan tiap-tiap pelanggan yang membeli produk CNI dari setiap agen-agen CNI. Adapun data – data atribut yang digunakan antara lain : Tanggal transaksi penjualan, kelompok transaksi, ID

pelanggan, kelompok kelamin, umur dan pekerjaan dari pelanggan, ID agen, nama agen, kota penjualan agen, ID produk, nama produk, harga produk, kelompok produk, tingkatan kelompok produk, jumlah barang yang dibeli, jumlah pembayaran, pajak pertambahan nilai dan total pembayaran dari transaksi pembelian yang dilakukan.

b. Pemilihan Grain

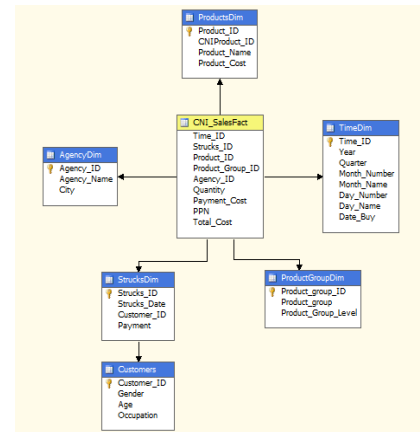
Grain merupakan proses untuk menentukan apa yang digambarkan oleh *record* di dalam tabel fakta. Terdapat empat perhitungan dalam tabel fakta yang akan dibuat di dalam *data warehouse* CNI ini yaitu *Quantity* atau jumlah barang, *payment cost* atau jumlah pembayaran tanpa PPN, adapun PPN atau Pajak Pertambahan Nilai dan *total cost* atau total pembayaran. Berikut adalah *grain* yang ada dalam perancangan *Data Warehouse* pada CNI Palembang.

c. Identifikasi dan Penyesuaian Dimensi

Pada tahap ini dilakukan identifikasi dan penyesuaian dimensi yang terkait dengan fakta yang ditampilkan dalam bentuk matriks. Dimensi yang digunakan dalam tabel fakta penjualan pada CNI Palembang adalah dimensi *time* (waktu), dimensi *strucks* (transaksi) dimensi *agency* (agen penjualan), dimensi *product* dan dimensi *product group*.

d. Pemilihan Fakta

Pada tahap ini dilakukan pemilihan fakta yang akan digunakan pada tabel fakta penjualan CNI. Adapun kolom – kolom pada tabel fakta penjualan CNI meliputi : *Time_ID*, *Strucks_ID*, *Product_ID*, *Product_Group_ID*, *Agency_ID*, *Quantity*, *Payment_Cost*, *PPN*, dan *Total_Cost*.



Gambar 4: Fakta Skema Snowflake

e. Penyimpanan Pre Calculation Ditabel Fakta

Dalam table fakta terdapat data yang merupakan kalkulasi awal. Hasil dari kalkulasi awal ini disimpan dalam tabel-tabel fakta. Adapun kalkulasi awal yang disimpan pada table CNI_SalesFact.

f. Memastikan Tabel Dimensi

Dalam tahap ini, kembali pada tabel dimensi dan menambahkan gambaran teks terhadap dimensi yang memungkinkan. Gambaran teks harus mudah digunakan dan dimengerti oleh user.

g. Pemilihan Durasi Database

Data yang dimasukan kedalam *Data Warehouse* ialah data transaksi penjualan produk pada PT.CNI Palembang selama tiga tahun.

h. Melacak Perubahan dari Dimensi Secara Perlahan

Dimensi-dimensi dan atribut yang mungkin dapat berubah dapat dilihat pada tabel.

Tabel 1: Kolom Dimensi yang Dapat Berubah

Nama Dimensi	Atribut yang Mungkin Berubah
AgencyDim	Agency_Name City
ProductDim	Product_Cost Product_Name
StrucksDim	Age Occupation

i. Penentuan Prioritas dan Model *Query*

Dalam tahapan ini yang dilakukan adalah mempertimbangkan pengaruh pada perancangan fisik, seperti keberadaan dari ringkasan (*summaries*) dan penjumlahan (*aggregate*). Dalam perancangan *data warehouse*, kapasitas media penyimpanan merupakan salah satu aspek yang perlu dipertimbangkan.

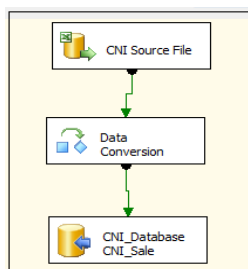
3.2 Extracting Data

Sumber data yang digunakan dalam perancangan *data warehouse* dan sumber data yang digunakan dalam perancangan *data warehouse* pada CNI Palembang ialah data *MS. Excel*.

Tabel 2: Sumber Data Penjualan CNI dalam Bentuk *Ms. Excel*

Payment	Customer_ID	Gender	Age	Occupation	Product id	Product Name	Product Cost	Agency_ID
Cash	CNIC135	P	42	Pegawai Swasta	HR025	CNI VITASIGI F RASA APEL	42000	CNIA221
Cash	CNIC135	P	42	Pegawai Swasta	N_07	PHYTO LITE SACHET	163000	CNIA221
Cash	CNIC135	P	42	Pegawai Swasta	HR024	CNI VITASIGI F RASA STRAWBERRY	42000	CNIA221
Cash	CNIC135	P	42	Pegawai Swasta	PC091	FIRE SENS NEW LABEL	23000	CNIA221
Cash	CNIC132	L	32	Wiraswasta	FD 385	SEHAT KECAK MANIS	14000	CNIA221
Cash	CNIC132	L	32	Wiraswasta	N_06	NARISH SACHET	175000	CNIA221
Cash	CNIC128	L	27	Wiraswasta	FD 05	PROGENE	610000	CNIA221

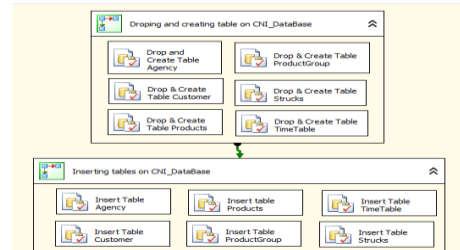
Kemudian data yang telah dikonversi akan dimasukkan kedalam Tabel CNI_Sale. Berikut gambar dari proses *Extracting Data*.



Gambar 5: Proses Ekstrak Data *Ms. Excel*

3.3 Transformation

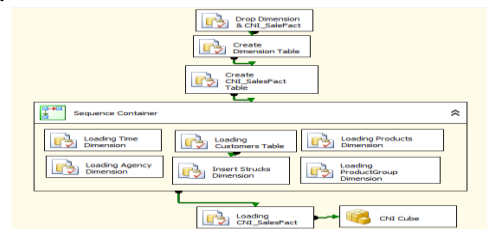
Setelah data-data berhasil dimasukkan ke dalam tabel CNI_sale pada *database* CNI, isi dari tabel CNI_sale akan diisi ke beberapa tabel baru yang dibuat di dalam *database* CNI yaitu : tabel Agency, Customer, Products, Strucks, ProductGroup dan TimeTable.



Gambar 6: Paket Transformasi pada SSIS

3.4 Loading Tabel Dimensi dan Tabel Fakta

Proses ini merupakan tahapan pembuatan tabel-tabel dimensi pada *database* DW_CNI yang bertujuan untuk mengisi tiap tabel-tabel dimensi, setiap isi tabel dimensi berasal dari tabel *database* CNI yang sudah terpisah. Berikut gambar proses *loading* tabel dimensi dan tabel fakta :



Gambar 7: Proses Loading Dimensi dan Fakta

4 ANALISIS DATA WAREHOUSE DAN HASIL DATA MINING

4.1 Presentasi Data Warehouse

Adapun hasil informasi yang didapatkan pada *data warehouse* PT. CNI Palembang yang berisikan data-data penjualan produk selama 3 tahun terakhir dari tahun 2011-2013 untuk setiap periode waktu, antara lain :

a. Total Pembayaran dan Jumlah Barang Berdasarkan Jenis Kelamin, Umur dan Pekerjaan Pelanggan

Total pembayaran dan jumlah barang berdasarkan jenis kelamin, umur dan pekerjaan pelanggan perbulan dapat dilihat pada tabel.

Tabel 3: Total Pembayaran dan Jumlah Barang Berdasarkan Jenis Kelamin, Umur dan Pekerjaan Pelanggan Perbulan

		Year		Month Name								
		2011		2012								
		January		February		March		April		May		
Gender	Age	Occupation	Quantity	Total Cost	Quantity	Total Cost	Quantity	Total Cost	Quantity	Total Cost	Quantity	Total Cost
L	19	Regenerasi	54	9954900.6335	49	3946936.4	32	5465972.1875	20	5972199.2		
		Programa Siswa										
		PHD	28	2834100.5	15	880850.3	11	830351.7499	15	667461.4		
		Universitas	73	17085394.6221	43	4934600	76	11370294.2375	97	1150914.0284		
		Total	101	19919495.1221	113	15770230.9335	125	15731230.6375	147	197168.9681	91	18150564.6384

Pada tabel 4.1 PT.CNI Palembang dapat melihat jumlah penjualan produk dari pelanggan pengendara dengan status mahasiswa yang berumur 19 tahun dan berjenis kelamin laki-laki sejumlah 54 item produk dengan total pembeliannya sebesar Rp. 9.954.900,6335 untuk bulan Februari 2011.

b. Informasi Data Warehouse CNI Berdasarkan Nama dan Kota Agen Penjualan

Tabel 4: Total Pembayaran dan Jumlah Barang Berdasarkan Nama dan Kota Agen Penjualan Tahun

		Year		Month Name	
		2011		2012	
City	Agency Name	Quantity	Total Cost	Quantity	Total Cost
Beltung		1530	2627731.21.2362	1272	20408391.1018
		1530	2627731.21.2362	1272	20408391.1018
Bengkulu		13364	2005987855.442	6610	999967237.1161
		13364	2005987855.442	6610	999967237.1161
Jambi	SEWAWAN SHOTANG	1468	21376985.9793	1349	21955848.5776
	CV. ANEKA CEMERLANG	2894	4758551.02.2863	3941	56325660.886
PT. SUKSES TUGU MANDIRI		4469	65302032.2771	3776	54403967.9699
		4469	65302032.2771	3776	54403967.9699
Total		8921	1244594139.5527	8566	1339722446.5335

Pihak CNI Palembang dapat melihat total jumlah penjualan produk dari 5 kota penjualan dan nama agen di wilayah Sumatera. Dapat dilihat pada agen 'CV. ANEKA CEMERLANG', pada kota 'Jambi' terjual 1382 item produk dengan total pembayaran sebesar Rp. 194.407.544,4229 yang jumlahnya menurun drastis dibandingkan hasil dari dua tahun sebelumnya.

c. Total Pembayaran dan Jumlah Barang Berdasarkan Jenis Transaksi dan Nama Produk

Tabel 5: Total Pembayaran dan Jumlah Barang Berdasarkan Jenis Transaksi dan Nama Produk

		Year		Month Name	
		2011		2012	
Payment	Product Name	Quantity	Total Cost	Quantity	Total Cost
Super SC 88 Laundry Compound - sachet		38	1669145.6876	40	1757457.1544
		38	1669145.6876	40	1757457.1544
		38	1669145.6876	40	1757457.1544
		38	1669145.6876	40	1757457.1544
		38	1669145.6876	40	1757457.1544

PT. CNI Palembang dapat melihat dari tabel 4.3 berdasarkan nama produk Super SC 88 Laundry Compound – Sachet yang jumlahnya ada 38 unit dan total pembayaran.

d. Total Pembayaran dan Jumlah Barang Berdasarkan Kelompok dan Tingkatan Kelompok Barang

Tabel 6: Total Pembayaran dan Jumlah Barang Berdasarkan Kelompok dan Tingkatan Kelompok Barang

		Year		Month Name	
		2011		2012	
Product Group Level	Product Group	Quantity	Total Cost	Quantity	Total Cost
3	SHHCDG88 - DETERGEN	159	8347070.1097	126	6462469.3006
	SHHCDG88 - DETERGEN	159	8347070.1097	126	6462469.3006
	SHHCDG88 - DETERGEN	159	8347070.1097	126	6462469.3006
	SHHCDG88 - DETERGEN	159	8347070.1097	126	6462469.3006
	SHHCDG88 - DETERGEN	159	8347070.1097	126	6462469.3006

PT. CNI Palembang dapat melihat tingkatan kelompok barang yang ke-3 pada tahun 2011 dan kelompok produk HSHCDG88-DETERGEN dari tahun 2011 sebanyak 159 item dengan total pembayaran sebesar Rp. 8.347.070,1097.

e. Analisis Pertumbuhan Jumlah Penjualan Produk dan Total Pembayaran Penjualan Produk

Tampilan analisis pertumbuhan jumlah penjualan produk dan total pembayaran transaksi perbulan yang tercatat dalam database PT. CNI Palembang.

Tabel 7: Pertumbuhan Data Perbulan

		Year		Month Name		Time Dim		Calculations	
		2011		2012					
		January		February		March			
		Quantity		Total Cost		Quantity		Total Cost	
		Quantity		Total Cost		Quantity		Total Cost	
City	Beltung	651	79741624.6	1193	178705836.3203	83.26%	124.11%	743	101640700.7607
	Bengkulu	317	54824954.5438	445	64498180.0263	40.38%	17.64%	723	117261303.2093
	Muara Bungo	1360	166371160.0287	628	104857437.7627	-53.82%	-36.97%	904	146274907.0315
	Palembang	2328	300937739.1725	2266	348061454.1093	-2.66%	15.66%	2431	382340949.1015
	Grand Total	2328	300937739.1725	2266	348061454.1093	-2.66%	15.66%	2431	382340949.1015

Dapat dilihat dari tabel 7 terjadi peningkatan jumlah penjualan produk pada kota penjualan Bengkulu sebesar 83.26 % dari bulan Januari ke bulan Februari 2013 dan total pembayaran transaksi sebanyak 124,11 %.

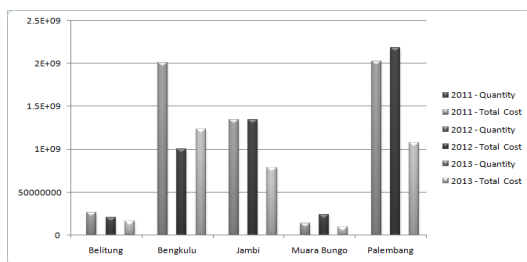
4.2 Laporan Hasil Analisis Data Warehouse

Berikut ini merupakan hasil laporan jumlah penjualan produk dan total Total Pembayaran untuk setiap nama agen per tahun dalam bentuk tabel pada *Microsoft Office Excel*.

Tabel 8: Laporan Jumlah Penjualan Produk dan Total Pembayaran CNI Palembang Selama Tiga Tahun dalam Bentuk Tabel

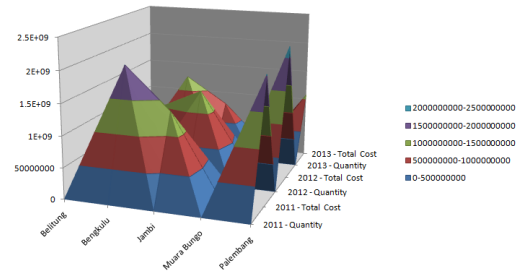
Row Labels	2011		2012		2013		Total	
	Quantity	Total Cost	Quantity	Total Cost	Quantity	Total Cost	Quantity	Total Cost
BERMAWAN SHOTANG	1468	211706986	1349	211933848.6	645	91418869.81	3462	517081704.4
CV. ANEKA CEMERLANG	2984	475853172.3	3441	563256604	1382	134407544.4	7807	1233519321
CV. MAMINTA NALU	3266	488306308.3	2673	431403306	1605	244612641.8	7544	1164324256
CV. MANDIRI BERSAMA	1430	207788033	2387	363937769.7	861	116215272.9	4678	687941075.7
CV. BANGUN JAYA	1393	232184302.1	1564	227429329.1	753	99002451.96	3710	558596113.2
CV. GLOBALINDO INSAN MANDIRI	3115	47454021.4	2589	456347408.8	2271	11694049.8	7955	1246750380
CV. JAYA SUKUNDI PERSADA	9879	1478034952	4482	685382129.9	6739	989832193.2	21100	3125150945
HADI OKTORA LUBIS	4052	619935886.8	4463	702861261.8	1544	287346633.2	10059	1629203090
IR. ANTONIUS MARBUN	3385	535932163.4	2128	314704077.2	1792	267987025.8	7305	1118645986
PT. SUKSES TUGU MANDIRI	4469	655020252.3	3776	540309988	3150	502078475.6	11395	1701616896
RACHMAT BASUKI	840	139679988.9	1393	234490543.6	637	99366322.13	2869	475567054.6
TIORU SIMBOLON	1520	262773121.2	1272	204308391.1	772	159759313.8	3564	626840644.1
Grand Total	37809	5714684608	31527	4960589718	22571	3359022030	91907	14894296356

Berikut ini tampilan beberapa laporan dalam bentuk grafik batang dan *surface*. Laporan dalam bentuk grafik batang dapat dilihat pada gambar yang menunjukkan jumlah Total Pembayaran untuk setiap kota penjualan selama 3 tahun.



Gambar 8: Laporan Jumlah Penjualan Produk dan Total Pembayaran Per Kota Penjualan dalam Bentuk Grafik Batang

Laporan jumlah Penjualan produk untuk setiap Kota Penjualan perbulan selama tiga tahun dalam bentuk diagram *surface*.



Gambar 9: Laporan Jumlah Penjualan Produk Per Kota Penjualan Per Bulan dalam Bentuk Diagram Surface

Pihak PT. CNI Palembang juga dapat melihat grafik informasi jumlah Penjualan produk dan Total Pembayaran dari berbagai bentuk grafik lainnya ataupun dapat juga mengkombinasikan antara dimensi dengan dimensi lainnya, dengan kategori yang dipilih sesuai kebutuhan.

4.3 Penerapan Data Mining

Pada Proses penerapan data mining pada CNI Palembang, metode yang digunakan adalah metodologi CRISP-DM (*Cross Industry Standard Process for Data Mining*).

a. Business Understanding (Pemahaman Bisnis)

Tahap pertama dari penerapan data mining ini ialah menentukan tujuan bisnis dan tujuan *Data Mining* serta rencana proyek lebih lanjut. Adapun tujuan bisnis yang ingin dicapai dari penerapan data mining pada PT. CNI Palembang ialah peningkatan keuntungan dan kinerja pada perusahaan.

b. Data Understanding (Pemahaman Data)

Adapun data yang digunakan untuk dapat memenuhi tujuan bisnis dan tujuan *data mining* yang telah ditetapkan adalah tabel transaksi pada CNI_DataWarehouse dalam bentuk format *CSV file* yang dapat dibentuk melalui *MS. excel*. Tabel transaksi penjualan yang ada pada *database CNI_Data Warehouse* akan diubah sesuai kebutuhan.

[illegible]

Setelah sumber data telah tersedia untuk diidentifikasi, pada tahap ini akan dilakukan persiapan. Pembentukan tabel yang dibutuhkan dalam proses *mining* menggunakan *apriori* pada Aplikasi WEKA.

Langkah keempat dari proses *data mining* adalah membangun sebuah model *data mining*. Adapun langkah-langkah pembentukan model mining menggunakan algoritma Apriori pada *Mining Tools WEKA* antara lain :

- ### *e. Evaluation*

[illegible]

```

Bear rules found:

1. SERATI KECAP MANIS= 1294 ==> HARISH SACHET= 1037 conf:(0.8)
2. SUPER MNC SACHET new= 1560 ==> HARISH SACHET= 1149 conf:(0.75)
3. STIG-FR 1,1 GR STIG= 1565 ==> STIG-FR 1165 conf:(0.74)
4. STIG-FR HAND + BODY LOTION (NEW BOTTLE)= 1173 ==> CMT VITAGIJI F RASA STRAWBERRE= 861 conf:(0.73)
5. CMT SANGING SELATAPAR= 1439 ==> MIE KINSENG SOTO AYAM - EKONOMI= 1042 conf:(0.72)
6. MIE KINSENG SOTO AYAM - EKONOMI= 1441 ==> CMT SANGING SELATAPAR= 1042 conf:(0.72)
7. CMT VITAGIJI F RASA STRAWBERRE= 1193 ==> STIG-FR HAND + BODY LOTION (NEW BOTTLE)= 861 conf:(0.72)
8. PROGRES (LISITIP)= 1717 ==> BIO PUL 1,1 GR STIG= 1165 conf:(0.68)
9. HARISH SACHET= 2635 ==> SUPER MNC SACHET new= 1169 conf:(0.44)
10. HARISH SACHET= 2635 ==> SERATI KECAP MANIS= 1037 conf:(0.39)

```

f. Deployment

Dengan persetujuan pihak PT. CNI Palembang, pembentukan *association rules* menggunakan algoritma *apriori* pada aplikasi WEKA dijalankan dengan *minimum support* sebesar 0,1 atau 84 kali kemunculan barang dan nilai *minimum confidence* sebesar 0,25 yang artinya aturan dengan tingkat kepercayaan rendah akan ditampilkan. Dari *mining* yang dilakukan dihasilkan 10 *assosiation rules*.

Setelah analisis selesai dilaksanakan maka dapat ditarik beberapa informasi yang telah didapatkan diantaranya:

- Hal - 8

confidence 0,75, “Jika membeli Super MPC Sachet new maka membeli Narish Sachet”.

confidence 0,74, “Jika membeli Bio Plus 1.2 gr isi 10 stick maka membeli Progene”.

confidence 0,73, “Jika membeli Sigi-F Hand & Body Lotion (new bottle) maka membeli CNI Vitasigi F Rasa Strawberry”.

confidence 0,72, “Jika membeli CNI Sarung Sejahtera maka membeli Mie Ginseng Soto Ayam – Ekonomis”, “ Jika membeli Mie Ginseng Soto Ayam – Ekonomis maka membeli CNI Sarung Sejahtera”, dan “Jika membeli CNI Vitasigi F Rasa Strawberry maka membeli Sigi-F Hand & Body Lotion (New Bottle)”.

confidence 0,68, “Jika membeli Progene (1strip) maka membeli Bio Plus 1.2 Gr Isi 10 Stick”.

Untuk melihat gambaran hasil prediksi yang telah dilakukan dengan menggunakan WEKA dapat dilihat pada gambar dibawah ini.

```
Best rules found:
1. SEHATI RECAP MANIS=t 1294 ==> NARISH SACHET=t 1037 conf:(0.8)
2. SUPER MPC SACHET new=t 1560 ==> NARISH SACHET=t 1169 conf:(0.75)
3. BIO PLUS 1.2 GR ISI 10 STICK=t 1565 ==> PROGENE (1STRIP)=t 1165 conf:(0.74)
4. SIGI-F HAND & BODY LOTION (NEW BOTTLE)=t 1173 ==> CNI VITASIGI F RASA STRAWBERRY=t 961 conf:(0.73)
5. CNI SARUNG SEJAHTERA=t 1439 ==> MIE GINSENG SOTO AYAM - EKONOMIS=t 1042 conf:(0.72)
6. MIE GINSENG SOTO AYAM - EKONOMIS=t 1441 ==> CNI SARUNG SEJAHTERA=t 1042 conf:(0.72)
7. CNI VITASIGI F RASA STRAWBERRY=t 1193 ==> SIGI-F HAND & BODY LOTION (NEW BOTTLE)=t 961 conf:(0.72)
8. PROGENE (1STRIP)=t 1717 ==> BIO PLUS 1.2 GR ISI 10 STICK=t 1165 conf:(0.68)
9. NARISH SACHET=t 2635 ==> SUPER MPC SACHET new=t 1169 conf:(0.44)
10. NARISH SACHET=t 2635 ==> SEHATI RECAP MANIS=t 1037 conf:(0.39)
```

Gambar 12: Rules hasil Pengujian

5 PENUTUP

5.1 Kesimpulan

- Terjadi penurunan drastis dua tahun sebelumnya pada agen ‘CV.ANEKA CEMERLANG’ Kota ‘Jambi’ dengan 1382 item produk terjual.
- Pada tahun 2011 dapat dilihat berdasarkan dari nama produk Super SC 88 Laundry Compound – Sachet yang jumlahnya ada 38 unit dengan total pembayaran sebesar Rp. 1.669.145,6876.

- Dapat dilihat tingkatan kelompok barang yang ke 3 pada tahun 2011 dan kelompok produk dari tahun 2011 sebanyak 159 item dengan total pembayaran sebesar Rp. 8.347.070,1097.
- Terjadi peningkatan jumlah penjualan produk pada Kota ‘Bengkulu’ sebesar 83.26% dari bulan Januari ke bulan Februari 2013 dan total pembayaran transaksi sebanyak 124,11%.
- Transaksi penjualan produk selama tiga tahun pada PT.CNI Palembang terdiri dari 8443 baris transaksi dan 293 kolom nama produk.
- Pembentukan *Association Rules* menggunakan *Algoritma Apriori* pada aplikasi *WEKA* dijalankan dengan *support* 0,1 dan *confidence* 0,25.
- Dapat dilihat jumlah penjualan produk dari pelanggan PT.CNI Palembang dengan status pekerjaan, umur, jenis kelamin perbulannya.

5.2 Saran

- Diharapkan data yang digunakan dalam pembentukan *datawarehouse* ini bukan hanya pada wilayah Sumatera Selatan dan sekitarnya saja, melainkan seluruh cabang penjualan PT. CNI yang ada di Indonesia.
- PT. CNI Palembang agar menambah proses lain selain penjualan barang, misalnya pembelian barang ataupun proses lainnya.
- Digunakan *software data mining* yang lain selain Weka misalnya Ms. Visual Studio, Rapidminer ataupun software lainnya dengan algoritma yang sama ataupun berbeda sehingga pihak PT. CNI dapat membandingkan *association rules* yang dihasilkan.

DAFTAR PUSTAKA

- [1] Connolly, Thomas, C. B 2010, *Database System :A Practical Approach to Design, Implementation, and Management Fifth Edition*, Pearson Education Inc.
- [2] Cuzzocrea, Alfredo 2011, *Data Warehouseing and Knowledge Discovery*. London : Springer-Verlag Berlin Heidelberg.
- [3] Indrajani 2009, *Sistem Basis Data dalam Paket Five In One*, PT. Elex Media Komputindo, Jakarta
- [4] Kimball, R, Margy R, Warren T, Joy M and Bob B 2008. *The Data Warehouse Lifecycle Toolkit*, Wiley Publishing Inc, Canada
- [5] Komputer, Wahana 2010, *Panduan Belajar MySQL Database Server*, Media Kita, Jakarta Selatan
- [6] Kusrini 2009, *Algoritma Data Mining*, Andi Offset, Yogyakarta
- [7] Nuraida, Ida 2008, *Manajemen Administrasi Perkantoran*, Kanisius, Yogyakarta
- [8] Ponniah, Paulraj 2011, *Data Warehouseing*, John Wiley & Sons Inc, Canada
- [9] Rainaldi, Vincent 2008, *Building A Data Warehouse With Examples In SQL Server*, Apress
- [10] Sulianta, Feri dan Dominikus Juju 2010, *Data Mining Meramalkan Bisnis Perusahaan*, Elex Media Komputindo, Jakarta
- [11] Tantra, Rudi 2012, *Manajemen Proyek Sistem Informasi*, Andi Offset, Yogyakarta