
PENERAPAN DATA WAREHOUSE DAN DATA MINING UNTUK PENGAMBILAN KEPUTUSAN PEMASARAN TERHADAP DATA PENJUALAN PERUSAHAAN RETAIL MENGUNAKAN ALGORITMA APRIORI : STUDI KASUS KEDAUNG GROUP PALEMBANG

Rizaldi Hariyanto¹, M. Rizal Sirajuddin², Mardiani³

^{1,2} STMIK Global Informatika MDP Palembang : Jl. Rajawali No.14 , Telp : 0711 - 376400

³ Jurusan Teknik Informatika, STMIK GI MDP PALEMBANG

e-mail: bernardaugust@mhs.mdp.ac.id, frost.rizal@gmail.com, mardiani@mdp.ac.id

Abstrak

Perusahaan retail perlu melakukan pemasaran yang efektif untuk menarik pelanggan. Dalam pemasaran, pengembangan produk, pencarian, komunikasi, distribusi, penetapan harga dan pelayanan sangatlah penting. Untuk mengambil keputusan pemasaran, perlu diketahui karakteristik pelanggan dan keterkaitan produk-produk yang diminati pelanggan. Dengan mengetahui informasi tersebut, manajemen dapat mengetahui keputusan pemasaran yang efektif. Untuk mempermudah mendapatkan informasi tersebut di Kedaung Group Palembang, penulis akan menerapkan *data warehouse* dan *data mining* menggunakan aturan asosiasi algoritma apriori dan didapatkan hasil dimana dapat disimpulkan pelanggan sering membeli produk sebagai hadiah atau kado yang ditunjukkan dari hasil penerapan *data mining* yang menunjukkan mayoritas pasangan produk adalah bungkus kado.

Kata kunci: *data warehouse, data mining, algoritma apriori, pemasaran.*

Abstract

Retail companies needs to have an effective sales strategy to attract customers. In a sales strategy, product development, research, communication, distribution, pricing and services are very crucial. To make a decision on sales strategy, information on customer characteristics and the connection of the products bought are needed. By having these information, the sales manager can make an effective decision. To get these information for Kedaung Group Palembang, the writer will implement data warehouse and datamining using apriori algorithm association rule. By implementing data warehouse and datamining, we got some results where customers are tend to buy products for presents as shown in the results of the implementation of data mining that shows the majority of product assocition are with present wrapping.

Keywords: *data warehouse, data mining, apriori algorithm, sales.*

1. PENDAHULUAN

Di zaman yang semakin tergantung pada teknologi, hampir semua aspek kehidupan sudah dibantu oleh teknologi-teknologi, salah satu teknologi tersebut adalah teknologi informasi. Teknologi informasi dapat membantu dalam mengolah informasi yang sudah ada menjadi informasi baru yang lebih berguna untuk hal-hal penting misalnya pengambilan keputusan untuk

pemasaran yang lebih efektif pada suatu perusahaan dari informasi atau data transaksi pelanggan-pelanggan sebelumnya.

Perusahaan retail rata-rata telah menggunakan Teknologi Informasi dalam proses transaksi bisnisnya. Salah satu contoh Teknologi Informasi di perusahaan retail adalah penggunaan aplikasi untuk mempermudah proses transaksi penjualan dan pengambilan keputusan untuk manajemen. Secara perlahan data transaksi yang dilakukan oleh perusahaan retail makin bertambah setiap hari, dan membutuhkan ruang penyimpanan data yang besar. Dalam hal ini perlu dikembangkan aplikasi yang memaksimalkan penyimpanan data dalam jumlah besar tersebut.

Dari uraian di atas dapat dilihat bahwa dalam sebuah perusahaan retail, kebutuhan akan informasi dari data penjualan sangatlah besar. Tetapi, beberapa perusahaan masih mengalami kesulitan dalam memproses data itu sendiri, khususnya data transaksi penjualan yang dapat menjadi sebuah informasi yang akan mendukung proses pengambilan keputusan perusahaan.

Pemasaran adalah kegiatan manusia yang diarahkan untuk memuaskan kebutuhan dan keinginan melalui proses pertukaran. Penjual harus mencari pembeli, mengenali kebutuhan mereka, merancang produk yang tepat, mempromosikan produk tersebut, menyimpan dan mengangkutnya, menegosiasikan dan sebagainya. Kegiatan seperti pengembangan produk, pencarian, komunikasi, distribusi, penetapan harga dan pelayanan merupakan inti kegiatan pemasaran.

Dalam mengambil keputusan untuk pemasaran, setiap perusahaan pasti memiliki caranya masing-masing. Data transaksi yang besar dan terus bertambah akan digunakan untuk mengambil keputusan tersebut. Perusahaan retail Kedaung, khususnya di cabang Palembang, juga mengambil keputusan dengan melihat data transaksi ini juga, tetapi hanya mengandalkan kesimpulan dari bagian manajemennya secara implisit.

Penelitian ini berkaitan dengan desain dan implementasi aplikasi pendukung pengambilan keputusan yang mengolah data transaksi sehingga menghasilkan informasi yang dapat digunakan untuk membantu manajemen Kedaung Group dalam memecahkan permasalahan yang berhubungan dengan manajemen pemasaran Kedaung serta menentukan strategi pemasaran yang bisa digunakan. Untuk itu dalam penelitian ini akan dibuat sebuah aplikasi data warehouse yang dilanjutkan dengan analisis menggunakan data mining.

1.1 Tujuan

Tujuan dari penelitian ini adalah membantu manajemen Kedaung Group dalam menemukan strategi penjualan, dengan cara menyediakan informasi berupa pola pembelian dari pelanggan dan pola keterkaitan antar produk.

1.2 Manfaat

Adapun manfaat yang diperoleh, yaitu:

1. Dengan mengetahui informasi dari pola pembelian pelanggan, manajemen dapat mengetahui produk yang akan dipesan dan dijual.
2. Dengan mengetahui pola keterkaitan antar produk, manajemen bisa mengambil keputusan untuk menentukan pemasaran yang efektif.

1.3 Landasan Teori

1.3.1 *Data Mining*

Data mining adalah proses mengekstraksi pola-pola yang menarik (implisit, belum diketahui sebelumnya dan berpotensi untuk bermanfaat) dari data yang berukuran besar. Ada beberapa istilah yang mempunyai kemiripan dengan *data mining*, yaitu ekstraksi pengetahuan, analisis pola, pengerukan data dan lain-lain. Ada yang berpendapat *data mining* merupakan sinonim dari istilah *knowledge discovery in database* (KDD). *Data mining* yang mempunyai arti yang luas sehingga ada banyak jenis fungsi yang dapat digolongkan dalam *data mining*. Salah satu contohnya adalah *Association Rule*. [4]

1.3.2 Algoritma Apriori

Algoritma *Apriori* adalah suatu algoritma dasar yang diusulkan oleh Agrawal & Srikant pada tahun 1994 untuk penentuan *frequent itemsets* untuk aturan asosiasi *boolean*. [2]

Mencari kombinasi *item* yang memenuhi syarat minimum dari nilai *support* dalam basis data dengan Nilai *support item* dapat diperoleh dengan menggunakan rumus berikut: [6]

$$Support(A) = \frac{Jumlah\ transaksi\ mengandung\ A}{Total\ Transaksi} \dots\dots\dots(1)$$

Nilai *support* dari 2 *item* diperoleh dengan menggunakan rumus :

$$Support(A,B) = P(A \cap B)$$

$$Support(A,B) = \frac{\sum Transaksi\ mengandung\ A\ dan\ B}{\sum Transaksi} \dots\dots\dots(2)$$

Frequent itemset adalah pasangan *item* yang memiliki frekuensi kemunculan lebih dari nilai minimum yang ditentukan (Φ). Misalkan $\Phi = 2$, maka semua pasangan *item* yang frekuensi kemunculannya lebih dari atau sama dengan 2 kali disebut *frequent*. Himpunan dari *frequent k-itemset* dilambangkan dengan F_k .

Setelah semua pola dengan frekuensi tinggi ditemukan , cari aturan asosiasi dengan menghitung *confidence* dari aturan asosiasi $A \rightarrow B$. [2]

$$confidence = P(B|A) = \frac{\sum transaksi\ mengandung\ A\ dan\ B}{\sum transaksi\ mengandung\ A} \dots\dots\dots(3)$$

1.3.3 Database

Database adalah sekumpulan data yang memiliki hubungan secara logika dan diatur dengan susunan tertentu serta disimpan dalam media penyimpanan komputer. Data itu sendiri adalah representasi dari semua fakta yang ada pada dunia nyata. Database sering digunakan untuk melakukan proses terhadap data-data tersebut untuk menghasilkan informasi tertentu.[5]

1.3.4 PHP

PHP merupakan secara umum dikenal sebagai bahasa pemrograman script-script yang membuat dokumen HTML (Hypertext Markup Language) yang dihasilkan secara on-the-fly yang dieksekusi di server web, dokumen HTML yang dihasilkan dari suatu aplikasi bukan dokumen HTML yang dibuat menggunakan editor teks atau editor HTML dikenal juga sebagai bahasa pemrograman server side. [10]

1.3.5 Framework CodeIgniter

Framework codeigniter adalah framework PHP yang membantu mempercepat developer dalam mengembangkan aplikasi web berbasis PHP dibanding jika menulis semua kode program dari awal.[3]

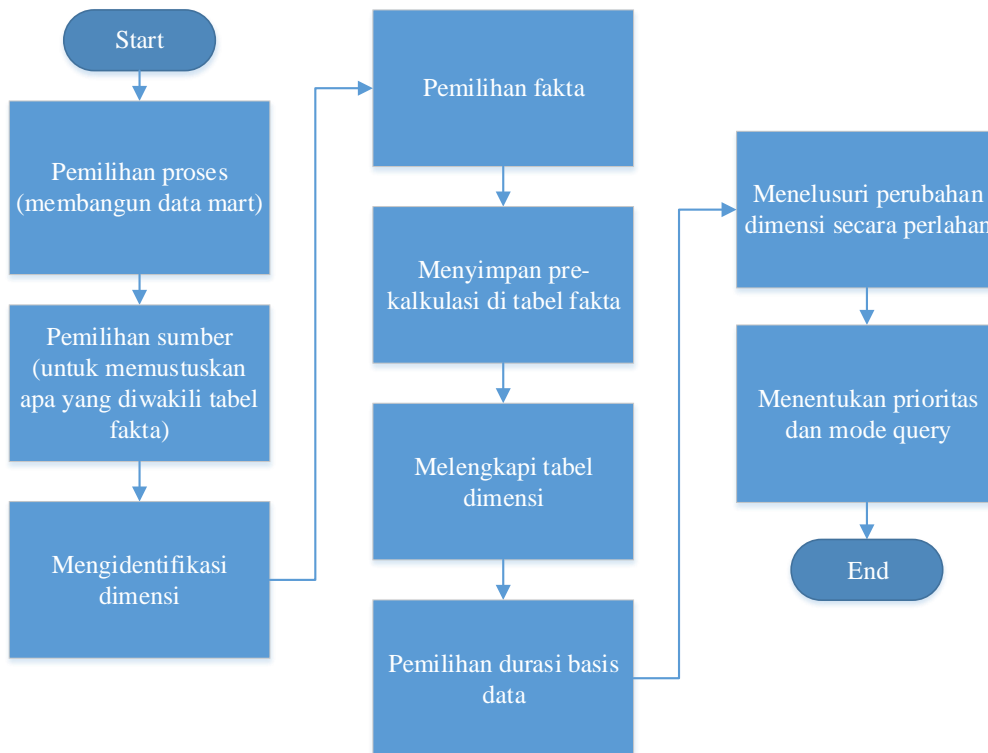
Codeigniter dibuat berdasarkan kaidah model-view-controller. Dengan MVC, memungkinkan pemisahan antara layer-application-logic dan presentation. Sehingga dalam pengembangan web, seorang programmer bisa berkonsentrasi pada core-system, sedangkan web designer bisa berkonsentrasi pada tampilan web. Skrip PHP, query MySQL, javascript dan CSS bisa saling terpisah , tidak dibuat dalam satu skrip besar yang membutuhkan resource yang besar untuk mengeksekusinya.[3]

1.3.6 MySQL

MySQL atau dibaca “My Sekuel” adalah suatu RDBMS (Relational Database Management System) yaitu aplikasi sistem yang menjalankan fungsi pengolahan basis data.[9]

2. METODE PENELITIAN

Metodologi yang digunakan pada penelitian ini ada dua yaitu metodologi untuk perancangan data warehouse dan metodologi untuk data mining. Pada perancangan data warehouse digunakan metodologi nine step(sembilan langkah) dari R.Kimball sedangkan untuk data mining digunakan metodologi CRISP-DM (Cross Industry Standard Process for Data Mining).



Gambar 2.1 Metodologi *nine-step*^[7]

2.1 Metodologi Perancangan Basis Data untuk *Data Warehouse*

Metodologi yang digunakan perancangan basis data untuk data warehouse adalah metodologi nine step(sembilan langkah) yang dikemukakan oleh R.Kimball (1996).[1]

Kesembilan langkah tersebut meliputi:[1]

Langkah 1: Pemilihan Proses

Pada tahap ini yang dilakukan adalah data mart yang pertama kali dibangun haruslah data mart yang dapat dikirim tepat waktu dan dapat menjawab semua pertanyaan bisnis yang penting.

Langkah 2: Pemilihan Sumber

Pemilihan sumber data untuk memutuskan secara pasti yang direpresentasikan oleh sebuah tabel fakta. Jika sumber data dari sebuah tabel fakta properti sale adalah properti sale

individual maka sumber dari sebuah dimensi pelanggan berisi rincian pelanggan yang membeli produk utama.

Langkah 3: Mengidentifikasi dimensi

Pada tahap identifikasi dimensi yang dilakukan adalah:

1. Set dimensi yang dibangun dengan baik, memberikan kemudahan untuk memahami dan menggunakan data mart.
2. Dimensi ini penting untuk menggambarkan fakta yang terdapat pada tabel fakta.
3. Jika ada dimensi yang muncul pada dua data mart, kedua data mart tersebut harus berdimensi sama atau salah satunya berupa subset matematis dari yang lainnya.
4. Jika sebuah dimensi digunakan pada dua data mart atau lebih, dan dimensi ini tidak disinkronisasi, maka keseluruhan data warehouse akan gagal, karena dua data mart tidak bisa digunakan secara bersamaan.

Langkah 4: Pemilihan fakta

Pada tahap pemilihan fakta yang dilakukan adalah:

- a) Sumber data dari sebuah tabel menentukan fakta yang bisa digunakan dalam data mart.
- b) Semua fakta diekspresikan pada tingkat yang telah ditentukan sumber.

Langkah 5: Menyimpan pre-kalkulasi di tabel fakta

Tabel fakta merupakan tabel utama dalam data warehouse, semua informasi yang ingin dicapai lewat data warehouse melalui tabel fakta. Banyak proses kalkulasi yang dilakukan terhadap tabel fakta, dan untuk memudahkan dalam implementasi ke data warehouse perlu menyimpan hasil pre-kalkulasi tersebut.

Langkah 6: Melengkapi tabel dimensi

Pada tahap ini yang kita lakukan adalah:

1. Menambahkan keterangan selengkapnya pada tabel dimensi.
2. Keterangan harus bersifat intuitif dan mudah dipahami oleh pengguna.

Langkah 7: Pemilihan durasi basis data

Berdasarkan kegunaan dari basis data yang dibuat maka pada tahap ini ditentukan berapa lama data tersebut tersimpan.

Langkah 8: Menelusuri perubahan dimensi secara perlahan

Ada tiga tipe perubahan dimensi, yaitu :

- a) Tipe 1. Atribut dimensi yang berubah ditulis ulang.
- b) Tipe 2. Atribut dimensi yang berubah menimbulkan dimensi baru.
- c) Tipe 3. Atribut dimensi yang berubah menimbulkan alternatif sehingga nilai atribut lama dan baru dapat diakses secara bersamaan pada dimensi yang sama.

Langkah 9: Menentukan prioritas dan mode query

Setelah langkah 1 sampai 8 dilalui, maka pada tahap ini akan menggunakan perancangan fisik. Tahap perancangan fisik ini menghasilkan data warehouse yang siap diimplementasi. Untuk itu perlu sebuah sistem yang didalamnya berisi query yang digunakan untuk menampilkan data yang diinginkan oleh pengguna.

2.2 Metodologi Data Mining

Untuk memudahkan penerapan data mining, diperlukan sebuah metodologi standar yang dikenal dengan CRISP-DM (Cross Industry Standard Process for Data Mining). CRISP-DM

terdiri dari: business understanding, data understanding, modeling, evaluation, dan deployment. Keuntungan menggunakan metodologi ini adalah tidak tergantung pada industri tertentu, peranti lunak tertentu, dan fokus pada proses data mining itu sendiri..



Gambar 2.2 Metodologi CRISP-DM^[7]

1. Business Understanding

Pada tahap business understanding, hal-hal yang dilakukan disini adalah kita harus menentukan tujuan bisnis dari perusahaan atau organisasi yang akan diterapkan data mining. Pemahaman terhadap tujuan bisnis merupakan hal yang penting sebelum melakukan langkah selanjutnya. Setelah melakukan pemahaman terhadap tujuan bisnis, disini juga perlu dilihat situasi atau kondisi dari perusahaan yang meliputi komponen data dan aplikasi yang ada saat ini. Berdasarkan fakta-fakta yang diperoleh pada tahap pemahaman bisnis, maka dapat ditentukan tujuan dari data mining yang akan diterapkan serta dapat menghasilkan rencana proyek yang akan dibuat. [8]

2. Data Understanding

Pada tahap data understanding, hal-hal yang perlu dilakukan adalah pengumpulan data awal yang akan digunakan pada penerapan data mining. Kualitas data awal yang diperoleh dapat menentukan dalam mengidentifikasi permasalahan. Tahap pemahaman data awal ini merupakan pandangan awal dalam mendapatkan data atau melakukan hipotesa terhadap informasi yang tersembunyi dari data tersebut. [8]

3. Data Preparation

Tahap persiapan data meliputi semua kegiatan untuk membangun data yang akan digunakan atau dimasukkan dalam pembuatan data modeling yang berasal dari data awal. Tugas persiapan data kemungkinan besar akan dilakukan beberapa kali dan tidak dalam urutan tertentu. Hal yang dilakukan dalam tahap persiapan data adalah membuat tabel, merekam data dan menyeleksi atribut serta melakukan transformasi dan pembersihan data untuk keperluan data modeling. [8]

4. Modeling

Pada tahap ini, berbagai teknik pemodelan data dipilih dan diterapkan beserta parameternya yang dikalibrasi untuk mendapatkan nilai-nilai yang optimal. Biasanya, ada beberapa teknik yang digunakan untuk jenis permasalahan data mining yang sama. Beberapa teknik mempunyai persyaratan khusus terhadap bentuk data yang digunakan. Oleh karena itu, sangat dimungkinkan pada tahap ini akan kembali ke tahap persiapan data (Data Preparation). [8]

5. Evaluation

Pada tahap ini sebuah model yang memiliki kualitas yang cukup tinggi dari perspektif analisis data telah dibangun. Sebelum melanjutkan ke tahap pengembangan (deployment) dari

model yang dibuat, sangat penting untuk mengevaluasi lebih dalam model dan meninjau langkah-langkah eksekusi untuk membangun model yang akan memastikan apa yang dilakukan sudah benar dan sesuai dengan tujuan bisnis. Satu tujuan utama adalah menentukan apakah ada beberapa isu bisnis penting yang sudah cukup dipertimbangkan. Pada akhir tahap ini, keputusan mengenai tujuan penggunaan data mining hasilnya harus tercapai. [8]

6. Deployment

Tahap deployment dilakukan setelah tahap evaluasi diperoleh hasil bahwa tujuan bisnis bisa dicapai dengan penggunaan data mining. Tahap deployment menentukan rencana pengembangan, melakukan pemantauan, dan menghasilkan final report melakukan review proyek yang dibuat. [8]

3. ANALISIS DAN PEMBAHASAN

3.1 Perancangan *Data Warehouse*

3.1.1 Identifikasi dan Analisis Kebutuhan Manajemen

Dalam pembuatan data warehouse, perlu dilakukan identifikasi kebutuhan dari pihak manajemen untuk membuat data warehouse yang sesuai dengan tujuan perancangan data warehouse. Berdasarkan hasil wawancara dengan manajer pemasaran Kedaung Group Palembang diperoleh laporan-laporan yang diinginkan meliputi:

1. Barang-barang yang pelanggan minati.
2. Pola pembelian barang yang diminati dengan barang lainnya.

3.1.2 Perancangan Arsitektur *Data Warehouse*

Setelah dilakukan identifikasi dan analisa kebutuhan manajemen, akan dilakukan pendeskripsian data warehouse yang akan dikembangkan. Berdasarkan tahapan sebelumnya, perancangan data warehouse manajemen pemasaran Kedaung Group Palembang tujuannya untuk memberikan gambaran terhadap pola barang-barang yang diminati pelanggan. Hasil pengolahan data akan memberikan masukan pada bagian manajemen pemasaran Kedaung Group Palembang dalam menentukan keputusan terhadap strategi pemasaran Kedaung Group Palembang.

Perencanaan arsitektur yang dibangun dalam pengembangan data warehouse ini terdiri dari penentuan sumber data dan arsitektur data warehouse. Perancangan arsitektur meliputi perancangan terhadap alur data dari sumber data yang menjadi acuan.

3.1.2.1 Sumber Data

Sumber data yang digunakan dalam pembuatan data warehouse berasal dari basis data penjualan di Kedaung Group Palembang yang terdiri dari tabel transaksi harian dan tabel barang. Data yang digunakan adalah data transaksi dari Januari 2016 sampai dengan Januari 2018.

Tabel transaksi harian memiliki atribut No_Transaksi, ID_Operator, ID_Article, Customer, Serial_No., Amount, Price, Disc, Disc_amount, PPN, Net, dan Value.

Tabel barang memiliki atribut ID_Article, Article, Description, Product, Price, Unit, dan Stock.

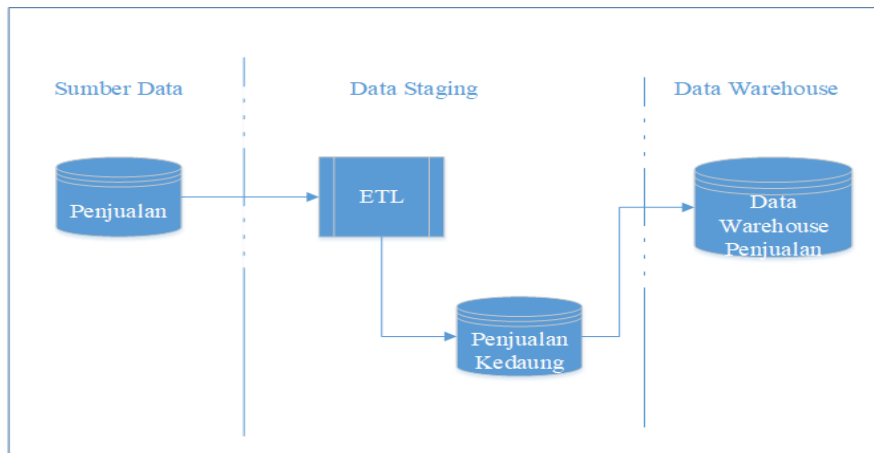
Tabel 3.1 Jumlah Keseluruhan Data

Tabel	Jumlah Data
Transaksi	928509
Artikel dan Barang	2453

3.1.2.2 Arsitektur *Data Warehouse*

Berdasarkan sumber data, dilakukan pemilihan data untuk menentukan data yang akan dimuat dalam data warehouse. Data yang telah dipilih akan diproses dan dilakukan pembersihan data. Pembersihan data ini meliputi membersihkan data duplikat, format data

yang tidak valid atau data yang bernilai null. Selanjutnya dilakukan transformasi data dari sumber data ke data warehouse. Pada tahap ini proses pembuatan data warehouse sudah berhasil dan dapat digunakan.



Gambar 3.1 Arsitektur Data Warehouse

3.1.3 ETL (Extraction, Transformation, Loading)

Data yang didapat dari Kedaung Group Palembang masih berupa data mentah yang harus dibersihkan dan dirapikan terlebih dahulu. Cleansing dari data warehouse ini memakan waktu yang lama dan meliputi beberapa proses yaitu:

a. Penghapusan kolom atribut yang tidak diperlukan

Pada tabel transaksi harian terdapat atribut No_Transaksi, ID_Operator, ID_Article, Customer, Serial_No., Amount, Price, Disc, Disc_amount, PPN, Net, dan Value. Dari atribut-atribut tersebut hanya diambil 6 atribut yaitu No_Transaksi, ID_Operator, ID_article, Customer, Amount dan Price.

Trans No	Status	Operator ID	Customer		Total Price	Total Disc.	Total PPN	Total Nett	Total Value	
Serial No	Type	Article ID	KIG Article ID	Amount/UOM	Price	Disc. %	Disc. Amount	PPN	Nett	Value
09843001	OK	YUNI	NILA		2,200.00	0.00	1,200.00		12,000.00	72,000.00
1 Sales		HGP12319	RB-5P	6 PCS	2,200.00	0.00	0.00	1,200.00	12,000.00	72,000.00
09843002	OK	YUNI	NILA		13,000.00		0.00	1,181.82	11,818.18	11,818.18
1 Sales		HGP12494	SQP-500SC	1PCS	13,000.00	0.00	0.00	1,181.82	11,818.18	11,818.18
09843003	OK	YUNI	BU SANTI		58,000.00		0.00	5,909.09	59,090.91	71,818.18
1 Sales		HGP14307	GIB-208(TP)6-1WB	1PCS	17,000.00	0.00	0.00	1,545.45	15,454.55	15,454.55
2 Sales		HGP14302	HMV-6.2B(TP)2-1WB	2 SET	7,000.00	0.00	0.00	1,272.73	12,727.27	25,454.55
3 Sales		HGP10321	608(TP)6-1WB	1SET	14,000.00	0.00	0.00	1,272.73	12,727.27	12,727.27
4 Sales		HNT10081	M-111	1PCS	20,000.00	0.00	0.00	1,818.18	18,181.82	18,181.82

Gambar 3.2 Tabel transaksi harian sebelum cleansing (1)

no_transaksi	id_operator	id_customer	tanggal	kode_artikel	jumlah	Price
9843001	YUNI	NILA	1/8/2017	HGP12919	6	2200
9843002	YUNI	NILA	1/8/2017	HGP12494	1	13000
9843003	YUNI	BU SANTI	1/8/2017	HGP14307	1	17000
9843003	YUNI	BU SANTI	1/8/2017	HGP14302	2	7000
9843003	YUNI	BU SANTI	1/8/2017	HGP10321	1	14000
9843003	YUNI	BU SANTI	1/8/2017	HNT10081	1	20000
9843004	YUNI	RINI	1/8/2017	HGP12378	24	3000
9843005	YUNI	BU AZIFAH	1/8/2017	HGP12494	2	13000
9843005	YUNI	BU AZIFAH	1/8/2017	HGP14090	2	39000
9843005	YUNI	BU AZIFAH	1/8/2017	HN110001	1	2000
9843005	YUNI	BU AZIFAH	1/8/2017	HGP10319	6	6000
9843005	YUNI	BU AZIFAH	1/8/2017	HGP14300	12	10000

Gambar 3.3 Tabel transaksi harian setelah cleansing (1)

b. Penghapusan record yang kosong

Dikarenakan beberapa sebab seperti kesalahan data entry dan proses update, terdapat beberapa record dengan atribut yang kosong, sehingga perlu dilakukan penghapusan dari record-record tersebut.

ID Article	Article	Description	Harga Jual SR (include PPN)	Unit Jual
HNE10138	py-6	Pan	200,000	Set
HNE10038	py-6	Pan	155,000	Set
HNE10261	KI-424	Pan	247,000	Set
HNE10367	KI-102/24 DMH	Pan	92,000	Set
HNE10368	KI-102/20 DMH	Pan	62,000	Set
HGP11129	TP-370	Ice Cream Cup	12,000	Pcs
HGP10134	D-125	Ice Cream Cup	11,000	Pcs
	TPH-10ST	Steamware	6,000	Pcs
	TPF-10ST	Steamware	6,000	Pcs
	IIP-350	Steamware	7,000	Pcs
	PDJ-100	Tumbler	3,000	Pcs
	WDJ-500	Pitcher	12,000	Pcs
	DJP-500	Pitcher	12,000	Pcs
HGP10385	STH-11M/6-1WBP	Steamware Set	39,000	Set
	STH-7TUM/D5-1GB	Steamware Set	45,000	Set

Gambar 3.4 Tabel kode artikel sebelum cleansing (1)

Kode Artikel	Nama Artikel	Deskripsi	Harga	Uni
HNE10138	py-6	Pan	200,000	Set
HNE10038	py-6	Pan	155,000	Set
HNE10261	KI-424	Pan	247,000	Set
HNE10367	KI-102/24 DMH	Pan	92,000	Set
HNE10368	KI-102/20 DMH	Pan	62,000	Set
HGP11129	TP-370	Ice Cream Cup	12,000	Pcs
HGP10134	D-125	Ice Cream Cup	11,000	Pcs
HGP10385	STH-11M/6-1WBP	Steamware Set	39,000	Set

Gambar 3.5 Tabel kode artikel setelah cleansing (1)

c. Pembersihan data yang tidak konsisten

Data yang berulang atau tidak konsisten ini adalah dari tabel barang dimana jenis/deskripsi barang diinput dengan ragam tulisan seperti plate dan piring, sendok dan spoon.

HGP12421	P-32/D	Tumbler	2,000	Pcs
HGP12914	RS-6D	Piring	1,600	Pcs
HGP12081	RS-8D	Plate	2,300	Pcs
HGP13771	TPH-2200	Bowl	7,500	Pcs
HGP12334	NPR-9D	Plate	2,200	Pcs
HGP12117	NBB-9D	Plate	2,200	Pcs
HGP10945	AT-107/SAM	Ashtray	5,000	Pcs
HGP10861	B-130/(TP)6-1SW	Bowl Set	16,000	Pcs
HGP11244	BB-7S	Plate	3,000	Set
HGP14074	BLC-8.5B	Bowl	8,000	Pcs
HGP10947	CCS-6.5B/TP	Bowl	6,500	Pcs

Gambar 3.6 Tabel kode artikel sebelum cleansing (2)

HGP12421	P-32/D	Tumbler	2,000	Pcs
HGP12914	RS-6D	Plate	1,600	Pcs
HGP12081	RS-8D	Plate	2,300	Pcs
HGP13771	TPH-2200	Bowl	7,500	Pcs
HGP12334	NPR-9D	Plate	2,200	Pcs
HGP12117	NBB-9D	Plate	2,200	Pcs
HGP10945	AT-107/SAM	Ashtray	5,000	Pcs
HGP10861	B-130/(TP)6-1SW	Bowl Set	16,000	Pcs
HGP11244	BB-7S	Plate	3,000	Set

Gambar 3.7 Tabel kode artikel setelah cleansing (2)

3.1.4 Transformasi Data

Setelah proses cleansing selesai, dilakukan transformasi data. Proses pemilihan dan penggabungan beberapa field ke field yang pada tabel baru.

Shift No	POS ID	Status	Operator ID	Customer	Total Price	Total Disc.	Total PPN	Total Nett	Total Value	
Serial No	Type	Article ID	KIG Article ID	Amount/UOM	Price	Disc. %	Disc. Amount	PPN	Nett	Value
9843001	1 Sales	YUNI	NLA		2,200.00	0.00	1,200.00	12,000.00	72,000.00	
		HGP12919	RB-5P	6 PCS	2,200.00	0.00	0.00	12,000.00	72,000.00	
9843002	1 Sales	YUNI	NLA		13,000.00		0.00	1,181.82	11,818.18	
		HGP12494	SGP-500SC	1PCS	13,000.00	0.00	0.00	1,181.82	11,818.18	
9843003	1 Sales	YUNI	BU SANTI		58,000.00		0.00	5,309.09	53,090.91	
		HGP14307	GB-208(TP)6-1vBP	1PCS	17,000.00	0.00	0.00	1,545.45	15,454.55	
		HGP14302	HNV-6.2B(TP)2-1vBP	2 SET	7,000.00	0.00	0.00	1,272.73	12,727.27	
		HGP10321	608(TP)6-1vBP	1SET	14,000.00	0.00	0.00	1,272.73	12,727.27	
		HNT10081	M-111	1PCS	20,000.00	0.00	0.00	1,818.18	18,181.82	

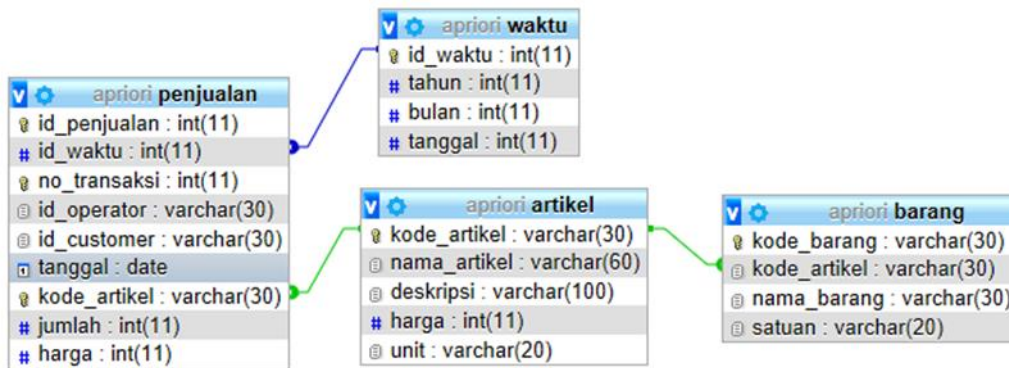
Gambar 3.8 Tabel transaksi harian sebelum cleansing (2)

no_transaksi	id_operator	id_customer	tanggal	kode_artikel	jumlah	Price
9843001	YUNI	NILA	1/8/2017	HGP12919	6	2200
9843002	YUNI	NILA	1/8/2017	HGP12494	1	13000
9843003	YUNI	BU SANTI	1/8/2017	HGP14307	1	17000
9843003	YUNI	BU SANTI	1/8/2017	HGP14302	2	7000
9843003	YUNI	BU SANTI	1/8/2017	HGP10321	1	14000
9843003	YUNI	BU SANTI	1/8/2017	HNT10081	1	20000
9843004	YUNI	RINI	1/8/2017	HGP12378	24	3000
9843005	YUNI	BU AZIFAH	1/8/2017	HGP12494	2	13000
9843005	YUNI	BU AZIFAH	1/8/2017	HGP14090	2	39000
9843005	YUNI	BU AZIFAH	1/8/2017	HN110001	1	2000
9843005	YUNI	BU AZIFAH	1/8/2017	HGP10319	6	6000
9843005	YUNI	BU AZIFAH	1/8/2017	HGP14300	12	10000

Gambar 3.9 Tabel transaksi harian setelah cleansing (2)

3.1.5 Pemodelan Data Multi Dimensi

Pemodelan data multi dimensi data warehouse manajemen pemasaran Kedaung Group Palembang menggunakan snowflake schema. Dengan tabel penjualan sebagai tabel fakta dan tabel waktu, artikel dan barang sebagai tabel dimensi. Pemodelan data multi dimensi dapat dilihat pada Gambar 3.10.

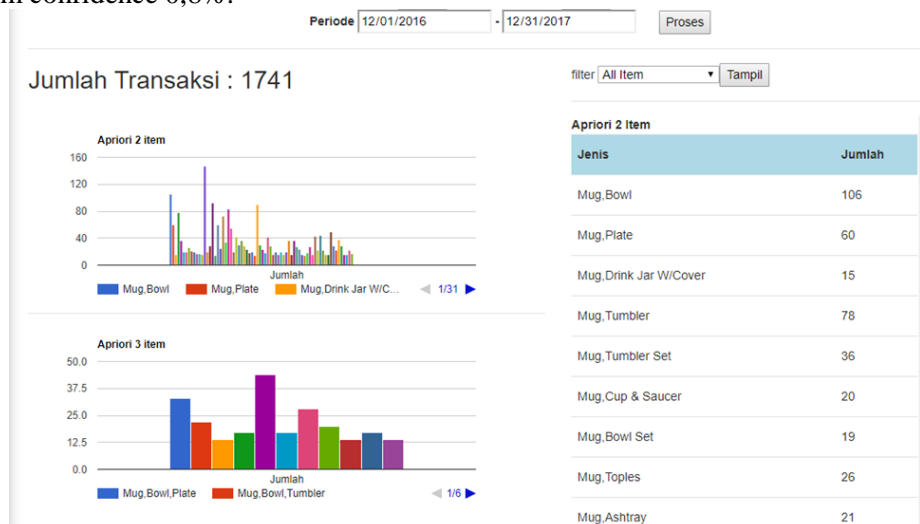


Gambar 3.10 Snowflake Schema Data Warehouse manajemen pemasaran Kedaung Group Palembang

3.2 Penerapan Data Mining

Proses penerapan data mining dibuat menggunakan bahasa pemrograman web PHP dan database menggunakan MySQL.

Berikut ini tampilan untuk hasil penerapan data mining dengan menggunakan data 1 tahun (Desember 2016 – Januari 2018) dari keseluruhan data dan dengan minimum support 2% dan minimum confidence 0,8%:



Gambar 3.11 Hasil Penerapan Data Mining (1)

Apriori 3 Item	
Jenis	Jumlah
Mug,Bowl,Plate	33
Mug,Bowl,Tumbler	22
Mug,Bowl,Toples	14
Mug,Plate,Tumbler	17
Bowl,Plate,Tumbler	44
Bowl,Plate,Tumbler Set	17
Bowl,Plate,Bowl Set	28
Bowl,Plate,Toples	20
Bowl,Tumbler,Bowl Set	14
Bowl,Tumbler Set,Bowl Set	17

Gambar 3.11 Hasil Penerapan Data Mining (2)

Dari hasil penerapan data mining menggunakan algoritma apriori dari 1 tahun yang didapat (Desember 2016 – Januari 2018) dengan minimum support 2% dan minimum confidence 0.8% didapatkan hasil yaitu

1. Pada bulan November 2017 didapatkan pola pasangan produk yang paling sering muncul yaitu (bungkus kado dan bowl), (bowl dan plate), (bungkus kado dan tumbler set), (mug dan bowl), (mug, bowl dan plate).
2. Pada bulan Desember 2017 didapatkan pola pasangan produk yang paling sering muncul yaitu (bungkus kado dan tea set), (bungkus kado dan bowl set), (bungkus kado dan tumbler set), (bungkus kado dan bowl), (mug dan bowl), (bowl dan plate), (mug, bowl dan plate), (mug, bowl dan tumbler), (bowl, tumbler dan plate).
3. Pada Januari 2018 didapatkan pola pasangan produk yang paling sering muncul yaitu (tumbler dan mug), (mug dan bowl), (bowl dan plate), (bowl dan bungkus kado), (bungkus kado dan tea set), (bungkus kado dan tea set), (mug, bowl dan plate).

4. KESIMPULAN

Kesimpulan yang dapat diambil dari penerapan data warehouse dan data mining di Kedaung Group Palembang adalah:

1. Dari hasil penerapan *data mining* dapat dilihat bahwa pola produk-produk yang dibeli berpasangan adalah produk-produk yang berhubungan dengan kebutuhan rumah tangga biasanya, misalnya pasangan mangkuk, piring dan gelas.
2. Dari hasil penerapan *data mining* dapat disimpulkan bahwa mayoritas pembelian produk bukan untuk pemakaian sendiri tapi sebagai hadiah yang dapat dibuktikan dengan mayoritas pola pasangan yang sering muncul adalah dengan bungkus kado.
3. Dengan mengetahui pola pasangan produk yang dibeli pelanggan, manajemen dapat membuat keputusan untuk memperbanyak produk dan penempatan produk yang sering berpasangan misalnya mangkuk, teko dan piring.
4. Dari hasil penerapan *data mining* dapat diketahui pelanggan Kedaung Group Palembang jarang untuk membeli produk lebih dari 2 jenis, ini ditunjukkan dengan kemunculan pola pasangan produk untuk lebih dari 2 hanya muncul jika minimum support yang sangat kecil.

5. SARAN

Berdasarkan penelitian ini, saran yang dapat diusulkan untuk penelitian selanjutnya, yaitu mengingat database yang digunakan perlu diproses dari aplikasi yang berbeda, penelitian berikutnya dapat eneruskan untuk mengkoneksikan database pada aplikasi tersebut dengan aplikasi yang dibuat untuk memaksimalkan penggunaan data mining.

DAFTAR PUSTAKA

- [1] Connoly, T. & Begg, C. (2005). *Database Systems-A Practical Approach to Design, Implementation and Management*. Fourth Edition. Addison Wesley.
 - [2] Gunadi, G., & Sensuse, D. I. (2012). Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori dan Frequent Pattern Growth (FP-Growth) : Studi Kasus Percetakan PT. Gramedia.
 - [3] Hakim, L. (2010). *Langkah Menjadi Master Framework Codeigniter*. Yogyakarta: Lokomedia.
 - [4] Jiawei, H., Kamber, M., & Tung, A. (2006). *Spatial Clustering Methods in Data Mining: A Survey*. Canada: School of Computing Science SIMon Fraser University Barnaby.
 - [5] Komputer, W. (2010). *SQL Server 2008 Express*. Yogyakarta: Andi Offset.
 - [6] Kusrini. (n.d.). *Algoritma Data Mining*. Yogyakarta: Andi Offset.
 - [7] Poe, V. (1998). *Building Data Warehouse for Decision Support*. Second Edition. Prentice Hall.
 - [8] Shearer, C.(2000). *The CRISP-DM Model: The New Blueprint for Data mining*. Journal of Data Warehouse. Volume 5 Number 4.
 - [9] Sibero, A. F. (2013). *Web Programming Power Pack*. Jakarta: Media kom.
 - [10] Sidik, B. (2012). *Pemrograman Web dengan PHP*. Bandung: Informatika.
-