Metode Deskripsi Untuk Mengetahui Pola Belanja Konsumen Pada Data Penjualan

Wahju Tjahjo Saputro*
Teknologi Informasi, Universitas Muhammadiyah Purworejo, 54111, Indonesia

Abstrak

Teknologi internet merupakan salah satu bagian dari teknologi informasi. Internet mampu memberikan kemudahan bagi siapa pun dalam memperoleh data dalam jumlah besar. Melimpahnya data memunculkan pertanyaan Pengetahuan apakah yang dapat dihasilkan dari data atau informasi tersebut?. Untuk menggali pengetahuan dari bongkahan data yang melimpah dapat diselesaikan dengan metode deskripsi pada bidang data mining. Metode deskripsi digunakan untuk mendeskripsikan dan menggambarkan pola belanja konsumen.Data penjualan selalu memiliki transaksi dengan jumlah frekuensi yang tinggi setiap hari. Data transaksi berjumlah banyak ini dapat ditemukan pengetahuan pola barang yang dijual. Data mining memiliki banyak metode atau algoritma yang mampu menggali berbagai pengetahuan. Penelitian ini berusaha menemukan pola pada data transaksi konsumen menggunakan metode Deskripsi menggunakan Borland Delphi 7.0 dan database MySQL. Hasil penelitian ini mampu menampilkan informasi berbasis pengetahuan berupa informasi nilai rata-rata, nilai tengah, nilai modus, tingkat keberagaman data dan grafik.

Kata kunci: data mining, metode deskripsi, penjualan

Abstract

Internet technology is one part of information technology. The internet can make it easy for anyone to obtain large amounts of data. The abundance of data raises the question of what knowledge can be generated from this data or information?. To gain knowledge from the abundant chunks of data, it can be solved by the method described in the data mining field. The description method is used to describe and describe the patterns of consumer spending. Sales data always has a high number of transactions every day. This large amount of transaction data can be found in the knowledge of the patterns of goods being sold. Data mining has many methods or algorithms that can explore various knowledge. This study seeks to find patterns in consumer transaction data using descriptive methods using Borland Delphi 7.0 and MySQL database. The results of this study can display knowledge-based information in the form of information on average values, mean values, mode values, levels of data diversity, and graphs.

Keywords: *data mining, description methods, sales*

1. PENDAHULUAN

Melimpahnya data memunculkan pertanyaan "Pengetahuan apakah yang dapat dihasilkan dari data atau informasi tersebut?" (Fayyad, Piatetsky-shapiro, & Smyth, 1996), (Yusuf, 2006). Untuk menggali pengetahuan dari bongkahan data yang melimpah digunakan metode data mining, salah satunya yaitu metode deskripsi (Susanto & Suryadi, 2010). Metode deskripsi digunakan untuk mendeskripsikan dan menggambarkan pola belanja konsumen. Data yang digunakan dalam penelitian ini adalah data transaksi penjualan, seperti kategori barang apakah yang paling sering laku dalam satu bulan atau periode tertentu.

Supermarket X berada di lokasi yang strategis dan dekat pasar tradisional. Lokasi ini menjadi faktor pendukung atas besarnya tingkat kunjungan konsumen. Rata-rata mencapai 800 transaksi setiap hari. Saat ini database yang ada di kasir diolah menggunakan Foxpro For DOS dan laporan yang dihasilkan berbentuk tabular dan konvensional. Laporan-laporan tersebut belum dapat digunakan sebagai bahan pertimbangan dalam mengambil keputusan bagi pimpinan.

Berangkat dari banyaknya jumlah transaksi dan kondisi laporan yang dihasilkan, maka penulis tertarik untuk menganalisis database transaksi penjualan dengan menggunakan data mining

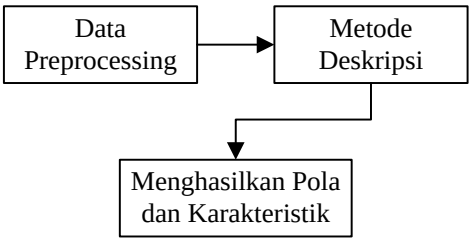
metode Deskripsi guna menemukan informasi berbasis pengetahuan yang bermanfaat dan belum diketahui dari data transaksi penjualan.

Data transaksi yang berskala besar dapat digunakan untuk mengetahui pola dan pengetahuan apa yang tersembunyi dalam bongkahan data tersebut (Huda, 2010), (Syaifullah, 2010), (Yusuf, 2006). Laporan yang bersifat tabular dapat diatasi dengan menggunakan metode deskripsi dari data mining. Dengan metode deskripsi, informasi yang dihasilkan akan berupa ringkasan, perbandingan sekelompok data dengan data yang lain atau gabungan keduanya dan ditampilkan berupa diagram (Susanto & Suryadi, 2010), (Santosa, 2007). Dengan informasi yang berupa ringkasan dan diagram tersebut akan membantu pihak menejemen dalam mengambil keputusan yang cepat dan efektif guna menghasilkan program promo yang sesuai dengan perilaku konsumen. Saat ini frekuensi kenaikan jumlah konsumen yang membeli barang sebatas bila ada promo untuk barang tertentu.

2. METODE

Dengan menggunakan metode Deskripsi pada fungsi minor data mining (Susanto & Suryadi, 2010), (Santosa, 2007). Permasalahan dalam penelitian ini membatasi pada data transaksi penjualan yang digunakan tahun 2007 sampai 2008. Kemudian data yang digunakan hanya data transaksi penjualan yang diolah oleh *cash register* pada bagian kasir. Karena ada beberapa transaksi pada supermarket X yang dilakukan manual. Seperti jasa layanan jam tangan, perbaikan alat elektronik dan stand yang berjualan dilingkungan supermarket. Penelitian ini menggunakan langkah penyelesaian seperti ditunjukkan Gambar 1.

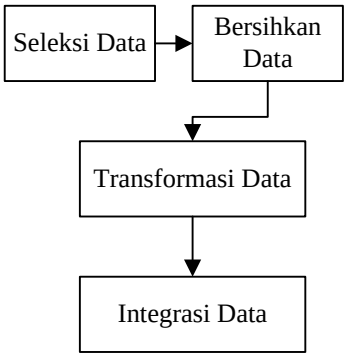
a. Melakukan *data preprocessing* pada database transaksi penjualan. *Data preprocessing* termasuk didalamnya meliputi *cleaning data*, filter data dan transformasi data.
b. Melakukan identifikasi metode deskripsi dengan query ke dalam Borland Delphi. Data yang telah melewati proses preprocessing ditranformasi ke MySQL, selanjutnya diolah ke aplikasi yang dibangun menggunakan Borland Delphi.
c. Menghasilkan pola dan karakteristik dengan melakukan uji coba aplikasi. Setelah diproses menggunakan aplikasi maka akan diperoleh sebuah pola atau karakteristik dari konsumen dalam berbelanja.



Gambar 1. Tahap metode Deskripsi pada data mining

2.1. Data Preprocessing

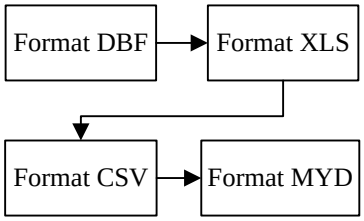
Tahap data *preprocessing* ditunjukkan pada Gambar 2. Selama tahap *preprocessing* dilakukan, data yang diperoleh mengalami berbagai perubahan format data, dimana setiap perubahan juga dilakukan proses data *cleaning* untuk semua data yaitu data barang, data penjualan maupun data detail penjualan.



Gambar 2. Tahap Data Preprocessing

Seleksi data adalah melakukan seleksi terhadap atribut yang akan dilibatkan dalam proses data mining. Beberapa atribut penting yang bersifat rahasia telah dilakukan pemotongan, seperti harga pokok barang, diskon, dan nilai transaksi. Dimana atribut tersebut tidak dilibatkan dalam proses data mining.

Membersihkan data yang dimaksud yaitu data yang diperoleh kondisinya masih belum siap digunakan karena ada *record* yang tidak memiliki nilai, ditemukan duplikasi *record*, format tanggal yang tidak konsisten, *incomplete* data dan data kosong. Pembersihan dilakukan agar data yang akan dipakai siap digunakan dalam proses data mining (Larose, 2005), (Fayyad, Piatetsky-shapiro, & Smyth, 1996), (Santosa, 2007). Langkah awal sebelum melakukan *cleaning* data dengan mengubah format data. Tahap perubahan format data tampak pada Gambar 3.



Gambar 3. Tahap melakukan perubahan format data

Transformasi data berguna untuk menstrukturkan data transaksi menjadi bentuk yang mudah di proses oleh data mining. Beberapa data dalam database transaksi masih perlu dilakukan transformasi. Seperti data kategori dimana potongan data kategori yang telah melalui proses transformasi ditunjukkan pada Tabel 1.

Tabel 1. Hasil potongan transformasi data kategori

Idketegori	Huruf
1	Agar-agar
2	Aneka perabot rumah tanggah
3	Aksesoris
4	Bahan roti
5	Beras
6	Bumbu-bumbu
7	Busana pria
8	Busana wanita

Data yang telah bersih dari *missing value*, nilai data tidak konsisten dan data yang tidak lengkap selanjutnya dilakukan integrasi data di MySQL (Larose, 2005). Proses *integrated* dilakukan dari berbagai data yang digabungkan untuk menghasilkan data yang disimpan pada tabel baru. Proses penggabungan ini dilakukan dengan perintah SQL. Hasil integrasi data tampak pada Gambar 4. Gambar 4 menunjukkan data yang siap untuk diproses dengan data mining. Data tersebut disimpan pada MySQL. Terdiri dari lima kolom yaitu atribut kode barang, nama barang, id kategori, tanggal dan nomor faktur.

2.2. Metode Deskripsi

Metode Deskripsi merupakan salah satu fungsi dalam data mining untukmenemukan atau menggali sekumpulan data yang banyak. Deskripsi juga merupakan fungsi minor pertama dalam data mining (Susanto & Suryadi, 2010). Tiga cara dalam mendeskripsikan data yaitu dengan menggunakan deskripsi grafis, deskripsi lokasi dan deskripsi keberagaman (Susanto & Suryadi, 2010), (Santosa, 2007).

Cara deskripsi grafis merupakan cara untuk mendeskripsikan data dalam bentuk gambar yang sesungguhnya. Gambar yang umum digunakan dalam deskripsi grafis adalah diagram titik (*dot diagram*) dan histogram (Susanto & Suryadi, 2010). Dengan menggunakan histogram dapat dilihat untuk dapat diperoleh pengetahuan dimana lokasi kecenderungan berkumpulnya data dari batang tertinggi.

Meskipun deskripsi grafis sudah menggambarkan karakteristik data namun sifatnya masih terlalu kasar dan kurang praktis untuk dilakukan, oleh karena itu masih diperlukan sebuah angka yang cukup dapat mewakili data yang ada serta dapat diperoleh dengan cara yang lebih praktis dari pada menggambar. Ada beberapa ukuran yang umum digunakan untuk menyatakan lokasi tersebut (Susanto & Suryadi, 2010). Rata-rata berarti ‘membuat menjadi rata’, dan nilai perataan tersebut dianggap sebagai lokasi pusat, titik berat, atau kesetimbangan data. Rumus menghitung rata-rata menggunakan persamaan (1).

kodebrg	namabrg	idkategori	tanggal	nofak
500017	DANISA MESIS WARNA 90 G	35	2007-08-20	5001
500017	DANISA MESIS WARNA 90 G	35	2007-08-20	7001
8300025	KAOS KETAT GEA PANJANG BLONG	8	2007-08-21	8001
8300222	KAOS KETAT JOLIE	8	2007-08-21	8002
3100097	BUNGA TANGKAI 1000	2	2007-08-21	8002
3100097	BUNGA TANGKAI 1000	2	2007-08-21	8002
3100098	NAMPAN OVAL A	2	2007-08-21	8002
3100098	NAMPAN OVAL A	2	2007-08-21	8002
5000188	HERS PROTEX S C MAXI WING 5'S	25	2007-08-21	8003
9400687	PRODANO C 324-325 HT	33	2007-08-21	9001

Gambar 4. Hasil integrasi data

$$\overline{X} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \sum_{i=1}^{n} \frac{x_i}{n} \quad \text{..... (1)}$$

Nilai Tengah (median), dengan ukuran ini, dicari nilai observasi yang secara harafiah 'bertempat di tengah-tengah. Langkah pertama yang tidak boleh dilupakan adalah mengurutkan semua data dari yang terkecil hingga yang terbesar. Apabila memiliki *n* buah data dan *n* adalah bilangan ganjil, maka letak nilai median dihitung menggunakan rumus persamaan (2).

$$\text{Median} = [(n+1)/2] \quad \text{..... (2)}$$

Sedangkan apabila jumlah *n* adalah bilangan genap maka mediannya dihitung dengan menggunakan rumus persamaan (3).

$$\frac{\text{data ke} - [n/2] \; + \; \text{data ke} - \; [(n/2) + 1]}{2} \quad \text{..... (3)}$$

Modus, dengan ukuran modus, dapat dicari nilai observasi yang paling sering muncul. Apabila terdapat dua atau lebih nilai yang kekerapan munculnya sama, semua nilai-nilai tersebut juga disebut modus.

Meskipun deskripsi lokasi sudah memberikan gambaran tentang lokasi pusat data (rata-rata, median, modus), masih belum memiliki gambaran atas keberagaman data. Untuk melengkapi gambaran terhadap data yang ada diperlukan ukuran keberagaman (Susanto & Suryadi, 2010). Varians, merupakan ukuran sebaran data yang mengunakan prinsip pencarian jarak antara setiap data dengan pusatnya (rata-rata). Untuk mencari nilai *varians* digunakan rumus seperti ditunjukkan pada persamaan (4).

$$S^2 = \frac{\sum_{i=1}^{n}(x_1 - \overline{X})}{n - 1} \quad \text{..... (4)}$$

Secara sederhana, rumus tersebut berarti setiap data observasi dikurangi dengan rata-rata seluruh data. Setiap hasil pengurangan tersebut dikuadratkan, kemudian semuanya dijumlahkan. Terakhir, hasil penjumlahan tersebut dibagi dengan (n-1), dengan *n* yang menyatakan banyaknya data.

2.3. Menghasilkan Pola dan Karatekristik

Merupakan proses menampilkan *output* dari hasil proses metode deskripsi data mining ketika aplikasi dijalankan. Pola dihasilkan dari proses mining menggunakan aplikasi yang dibangun dengan Delphi dan database tersimpan di MySQL.

3. HASIL DAN PEMBAHASAN

3.1. Mendapatkan Pengetahuan Nilai Rata-rata

Berdasarkan data yang diperoleh dari pihak Supermarket X, transaksi penjualan yang terjadi di Supermarket X selama periode bulan Agustus 2008 ditunjukkan Tabel 2.

Tabel 2. Frekuensi penjualan periode Agustus 2008

ID Kategori	Nama Kategori	Jumlah Transaksi
1	Agar-Agar	184
2	Aneka Perabot RT	8.093
3	Asesoris	3.480
4	Bahan Roti	324
5	Beras	0
6	Bumbu-Bumbu	955
7	Busana Pria	4.125
8	Busana Wanita	9.713
9	Daging Siap Saji	225
10	Elektronik	374
11	Gula	311
12	Ice Cream	404
13	Kopi	551
14	Kosmetik	2.336
15	Kue Basah	443
16	Mainan Anak	670
17	Makanan Kering	62
18	Mie Instan	2.287
19	Minuman	1.988
20	Minyak Goreng	396
21	Obat-Obatan	748
22	Pakaian Anak	2.015
23	Parfum Deodoran	866
24	Pasta Sikat Gigi	1.434
25	Pembalut Kapas	1.264
26	Pembasmi Serangga	346
27	Pembersih Kamar Mandi Lantai	276
28	Perlengkapan Alat Tulis	12.691
29	Perlengkapan Bayi	4.868
30	Perlengkapan Busana Muslim	3.255
31	Sabun Cuci	2.958
32	Sabun Mandi	2.465
33	Sepatu, Sandal Kaos Kaki	6.171
34	Shampo	1.465
35	Snack Biskuit	5.482

36	Susu Makanan Bayi	2.766
37	Teh	465
38	Telur	92
39	Rokok	194
	Jumlah	**86.742**

$$\overline{X} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \sum_{i=1}^{n} \frac{x_i}{n}$$

$$\overline{X} = \frac{86742}{39}$$

$$= 2.224{,}1538$$

Jadi rata-rata transaksi penjualan yang terjadi di Supermarket X periode bulan Agustus 2008 berdasarkan perhitungan di atas adalah 2.224,1538.

3.2. Mendapatkan Pengetahuan Nilai Median

Sebelum mencari nilai tengah terlebih dahulu harus diketahui apakah jumlah data yang akan diobservasi berjumlah genap atau ganjil. Selain itu data juga harus diurutkan dimulai dari data dengan jumlah frekuensi transaksi terkecil hingga terbesar.

Apabila jumlah data yang akan diobservasi genap, maka informasi letak nilai tengah diperoleh dari perhitungan rumus {[(n/2)+((n/2)+1)]/2}, dengan *n* adalah jumlah frekuensi transaksi penjualan yang terjadi pada periode yang diinputkan. Sedangkan apabila jumlah data yang akan diobservasi ganjil, maka informasi letak nilai tengah diperoleh dari perhitungan rumus [(n+1)/2], dengan *n* adalah jumlah frekuensi transaksi penjualan yang terjadi pada periode yang diinputkan. Data transaksi penjualan yang sudah diurutkan berdasarkan jumlah frekuensi transaksi dari yang terkecil hingga terbesar ditunjukkan Tabel 3.

Tabel 3. Frekuensi penjualan periode Agustus 2008 yang telah diurutkan

ID Kategori	Nama Kategori	Jumlah Transaksi
5	Beras	0
17	Makanan Kering	62
38	Telur	92
1	Agar-Agar	184
39	Rokok	194
9	Daging Siap Saji	225
27	Pembersih Kamar Mandi Lantai	276
11	Gula	311
4	Bahan Roti	324
26	Pembasmi Serangga	346
10	Elektronik	374
20	Minyak Goreng	396
12	Ice Cream	404
15	Kue Basah	443
37	Teh	465
13	Kopi	551
16	Mainan Anak-Anak	670
21	Obat-Obatan	748
23	Parfum Dan Deodorant	866
6	Bumbu-Bumbu	955
24	Pasta Dan Sikat Gigi	1434
34	Shampo	1465
19	Minuman	1988
22	Pakaian Anak-Anak	2015
18	Mie Instan	2287
14	Kosmetik	2336
32	Sabun Mandi	2465
36	Susu Dan Makanan Bayi	2766
31	Sabun Cuci	2958
30	Perlengkapan Busana Muslim	3255
3	Asesoris	3480
7	Busana Pria	4125
29	Perlengkapan Bayi	4865
35	Snack Dan Biskuit	5482
33	Sepatu, Sandal Dan Kaos Kaki	6171
2	Aneka Perabot Rumah Tangga	8093
8	Busana Wanita	9713
28	Perlengkapan Alat Tulis	12691
	Jumlah	**86.742**

Jumlah transaksi yang tejadi pada periode bulan Agustus 2008 bernilai genap yaitu 86.742 maka perhitungan nilai *varians* menggunakan rumus berikut:

Nilai tengah terletak pada data ke-
= {[(n/2)+((n/2)+1)]/2}
= {[(86.742 / 2) + ((86.742/ 2) + 1)] / 2}
= {[43.371+ 43.372] / 2} = 43.371,5

Berdasarkan perhitungan di atas maka nilai median terletak pada data transaksi penjualan dengan urutan data ke- 43.371,5.

Pada dasarnya fungsi dari nilai rata-rata dan nilai median sama yaitu untuk mencari kategori barang yang kurang laku dan sering laku. Tetapi hasil yang diperoleh akan terjadi perbedaan yaitu terdapat kategori barang yang ketika diproses berdasarkan nilai rata-rata maka kategori barang tersebut masuk ke dalam kategori barang sering laku, sedangkan apabila diproses menggunakan nilai median maka kategori barang tersebut masuk ke dalam kategori barang yang kurang laku. Jika dilihat dari data transaksi yang ditunjukkan Tabel 3. kategori barang tersebut yaitu:

1 Mie instan

2 Kosmetik
3 Sabun mandi
4 Susu dan makanan bayi
5 Sabun cuci
6 Perlengkapan busana muslim
7 Asesoris
8 Busana pria

Hal ini dikarenakan adanya pengaruh nilai ekstrim atau kategori barang tertentu yang mempunyai nilai frekuensi yang besar dan jauh dari nilai frekuensi kategori barang yang lainnya pada perhitungan nilai rata-rata.

3.3. Mendapatkan Pengetahuan Nilai Modus

Nilai frekuensi penjualan tertinggi merupakan nilai observasi yang paling sering muncul (modus). Nilai frekuensi penjualan tertinggi diperoleh dari mencari kategori barang yang mempunyai frekuensi transaksi penjualan terbanyak. Berdasarkan data transaksi yang ditunjukkan pada Tabel 2. Nilai modusnya adalah 12.691 yaitu untuk kategori perlengkapan alat tulis.

3.4. Pengetahuan Keberagaman Data

Informasi tingkat keberagaman data, dalam hal ini adalah kategori barang yang terjual digunakan untuk mengetahui perbandingan tingkat keberagaman frekuensi transaksi penjualan dari dua periode yang berbeda. Transaksi penjualan yang beragam berarti bahwa dalam satu periode transaksi penjualan jumlah frekuensi transaksi penjualan antara kategori barang yang satu dengan kategori barang lainnya beragam atau tidak mengelompok pada jumlah tertentu. Sedangkan transaksi penjualan yang kurang beragam berarti bahwa dalam satu periode transaksi penjualan jumlah frekuensi transaksi antara kategori barang yang satu dengan yang lainnya banyak yang sama.

Data yang akan digunakan sebagai perbandingan adalah data periode bulan Agustus 2008 dan September 2008. Tingkat keberagaman ini dihitung menggunakan ukuran sebaran *varians.* V*arians* merupakan ukuran sebaran data yang menggunakan prinsip mencari jarak antara setiap data dengan pusatnya (rata-rata) yaitu apabila diaplikasikan ke dalam rumus, maka rumus *varians* adalah : $S^2 = \frac{\sum_{i=1}^{n}(x_1 - \overline{X})}{n-1}$. Secara sederhana rumus tersebut dapat dijelaskan bahwa nilai *varians* diperoleh dengan mengurangkan setiap data observasi dengan rata-rata seluruh data, kemudian setiap hasil pengurangan tersebut dikuadratkan untuk selanjutnya semuanya dijumlahkan. Terakhir, hasil penjumlahan tersebut dibagi dengan (n - 1), dengan *n* menyatakan banyaknya data. Data yang akan digunakan ditunjukkan pada Tabel 4.

Tabel 4. Rincian perhitungan *varians* penjualan Agustus dan September 2008

ID Kategori	Frekuensi Agustus 2008	S1	Frekuensi September 2008	S2
1	184	4.162.227,52765444	280	5.106.093,59508889
2	8.093	34.443.355,71925440	5.422	8.307.845,25228889
3	3.480	1.577.149,67805444	3.042	252.338,74428889
4	314	3.610.584,46365444	275	5.128.715,26208889
5	0	4.946.860,12605444	0	6.449.906,94708889
6	955	1.610.751,36805444	738	3.246.002,89788889
7	4.125	3.613.216,27605444	8.993	41.645.510,68088890
8	9.713	56.082.817,40725440	18.292	248.136.004,39428900
9	225	3.996.615,91605444	257	5.210.567,26328889
10	374	3.423.069,08365444	298	5.025.069,59388889
11	311	3.660.157,46245444	359	4.755.307,25648889
12	404	3.312.959,85565444	214	5.408.725,59948889
13	551	2.799.443,63845444	397	4.591.020,58728889
14	2.336	12.509,57245444	1.855	468.768,49008889
15	443	3.172.508,85925444	615	3.704.341,90608889
16	670	2.415.394,03405444	388	4.629.669,58788889
17	62	4.674.909,05485444	36	6.268.346,94468889
18	2.287	3.949,64485444	1.553	973.511,17688889
19	1.988	55.768,61725444	2.464	5.725,44948889
20	396	3.342.146,31645444	274	5.133.245,59548889
21	748	2.179.030,04125444	509	4.123.607,24648889
22	2.015	43.745,31205444	6.266	13.885.559,86268890
23	866	1.844.581,74445444	731	3.271.275,23168889

24	1.434	624.343,02765444	978	2.438.802,88188889
25	1.264	921.895,31965444	847	2.865.120,55728889
26	346	3.527.461,69645444	245	5.265.495,26408889
27	276	3.795.303,22845444	230	5.334.560,26508889
28	12.691	109.554.869,37445400	7.736	27.001.879,76468890
29	4.868	6.989.922,72925444	6.653	16.919.510,83688890
30	3.255	1.062.643,88805444	8.828	39.543.135,69188890
31	2.958	538.530,24525444	2.167	138.880,46928889
32	2.465	58.006,89205444	1.664	766.792,16948889
33	6.171	15.577.594,92645440	8.222	32.288.911,73228890
34	1.465	576.314,49205444	1.059	2.192.373,87648889
35	5.482	10.613.561,86285440	4.264	2.973.325,32948889
36	2.766	293.597,30445444	2.349	36.353,79048889
37	465	3.094.622,09205444	357	4.764.033,92328889
38	92	4.546.079,82685444	65	6.123.975,27608889
39	194	4.121.524,45165444	125	5.830.615,27208889
Jumlah	**86.742**	**310.880.023,07692256**	**99.047**	**540.210.926,66666688**
Rata-rata	**2.224,1538**		**2.539,6667**	

Berdasarkan data yang ditunjukkan Tabel 4 total transaksi penjualan yang terjadi pada periode bulan Agustus 2008 adalah 86.742 dan jumlah kategori barang 39, sehingga rata-rata transaksi yaitu:

$$\overline{X} = \frac{x_1+x_2+\cdots+x_n}{n} = \frac{86.742}{39} = 2.224,1538$$

Sebagai contoh, dari Tabel 4 berikut rincian perhitungan *varians* untuk ID Kategori Barang 1 (Agar-agar), Barang 2 (Perabot rumahtangga), Barang 3 (asesoris) dan Barang 4 (Bahan roti) periode bulan Agustus 2008 yaitu, perhitungan *varians* untuk ID Kategori Barang 1 (Agar-agar) di dapat dengan perhitungan berikut:

$$S1 = (x_1 - \bar{X})^2 = (184 - 2.224,1538)^2$$
$$= 4.162.227,52765444$$

Perhitungan *varians* untuk ID Kategori Barang 2 (Perabot Rumah Tangga) di dapat dengan perhitungan berikut:

$$S1 = (x_2 - \bar{X})^2 = (8.093 - 2.224,1538)^2$$
$$= 34.443.355,71925440$$

Perhitungan *varians* untuk ID Kategori Barang 3 (Asesoris) di dapat dengan perhitungan berikut:

$$S1 = (x_3 - \bar{X})^2 = (3.480 - 2.224,1538)^2$$
$$= 1.577.149,67805444$$

Perhitungan *varians* untuk ID Kategori Barang 4 (Bahan Roti) di dapat dengan perhitungan berikut:

$$S1 = (x_4 - \bar{X})^2 = (324 - 2.224,1538)^2$$
$$= 3.610.584,46365444$$

Varians transaksi penjualan periode bulan Agustus 2008 yaitu:

$$Varians = \frac{\sum_{i=1}^{n}(x_1 - \bar{X})^2}{n-1}$$
$$= \frac{310.880.023,07692256}{39-1}$$
$$= \frac{310.880.023,07692256}{38}$$
$$= 8.181.053,23886638$$

Berdasarkan data yang ditunjukkan Tabel 4 total transaksi penjualan yang terjadi pada periode bulan September 2008 adalah 99.047 dan jumlah kategori barang 39, sehingga rata-rata transaksinya yaitu:

$$\overline{X} = \frac{x_1+x_2+\cdots+x_n}{n}$$
$$= \frac{99.047}{39} = 2.539,6667$$

Sebagai contoh, dari Tabel 4 rincian perhitungan *varians* untuk ID Kategori Barang 1 (Agar-agar), Barang 2 (Perabot rumahtangga), Barang 3 (asesoris) dan Barang 4 (Bahan roti) periode bulan September 2008 yaitu:Perhitungan *varians* untuk ID Kategori Barang 1 (Agar-agar) di dapat dengan perhitungan berikut:

$$S2 = (x_1 - \bar{X})^2 = (280 - 2.539,6667)^2$$
$$= 5.106.093,59508889$$

Perhitungan *varians* untuk ID Kategori Barang 2 (Perabot Rumah Tangga) di dapat:

$$S2 = (x_2 - \bar{X})^2 = (5.422 - 2.539,6667)^2$$
$$= 8.307.845,25228889$$

Perhitungan *varians* untuk ID Kategori Barang 3 (Asesoris) di dapat dengan perhitungan berikut:

$$S2 = (x_3 - \bar{X})^2 = (3.480 - 2.539,6667)^2$$

= 252.338,74428889

Perhitungan *varians* untuk ID Kategori Barang 4 (Bahan Roti) di dapat dengan:

$S2 = (x_4 - \bar{X})^2 = (275 - 2.539{,}6667)^2$
= 5.128.715,26208889

Varians transaksi penjualan periode bulan September 2008 yaitu:

$$Varians = \frac{\sum_{i=1}^{n}(x_1 - \bar{X})^2}{n-1}$$

$$= \frac{540.210.926{,}66666688}{39-1}$$

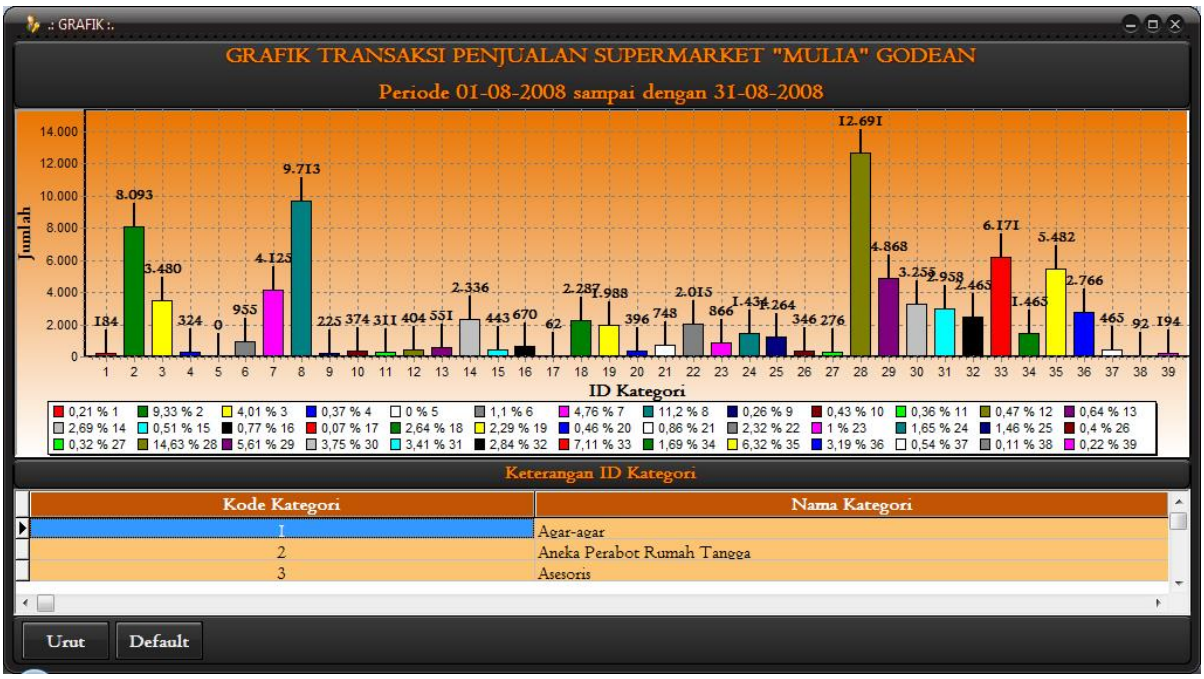
$$= \frac{540.210.926{,}66666688}{38}$$

= 14.216.077,01754390

Nilai *varians* transaksi penjualan periode bulan September 2008 lebih besar dibandingkan dengan nilai *varians* transaksi penjualan periode bulan Agustus 2008, ini berarti bahwa transaksi penjualan periode September 2008 lebih beragam dibandingkan periode Agustus 2008. Frekuensi transaksi penjualan untuk masing-masing kategori barang pada periode bulan September 2008 lebih beragam dan terdapat kategori barang yang lebih dominan atau diminati konsumen pada periode bulan tersebut.

3.5. Analisis Pengetahuan Grafik

Berdasarkan data yang ditunjukkan pada Tabel 2. Dalam implementasi data mining yang dibangun akan terbentuk *output* seperti ditunjukkan Gambar 5.



Gambar 5. Luaran grafik transaksi penjualan

Grafik merupakan visualisasi atau gambaran dari sekelompok data. Dengan melihat grafik dapat diketahui dimana letak kecenderungan berkumpulnya data dengan ditunjukkan oleh batang grafik yang tinggi dalam hal ini adalah kategori barang yang sering laku, sedangkan batang grafik yang rendah menunjukkan kategori barang yang jarang laku.

4. KESIMPULAN

Hasil penelitian ini dapat menampilkan pola transaksi penjualan pada Supermarket X per periode. Informasi yang ditampilkan berupa nilai rata-rata, nilai tengah, modus, varians dan grafik beserta informasi yang berkaitan dengan masing-masing nilai rata-rata, nilai tengah, modus dan varians.

Hasil dari proses data mining ini dapat digunakan sebagai pertimbangan dalam mengambil keputusan lebih lanjut untuk menghasilkan program

promo dan memperkirakan pola belanja konsumen di masa mendatang.

REKOMENDASI

Untuk pengembangan aplikasi data mining lebih lanjut, peneliti dapat menambahkan seleksi periode berdasarkan bulan dan tahun agar lebih *user friendly* karena dalam penelitian ini *input*-an periode masih berdasarkan tanggal. Selanjutnya peneliti juga dapat menambahkan *output* berupa saran untuk mendukung pengambilan keputusan oleh manager. Dapat pula peneliti melakukan penelitian menggunakan metode yang lain untuk dapat menampilkan informasi berbasis pengetahuan yang dapat digunakan untuk membantu menejer pemasaran dalam mengambil keputusan.

DAFTAR PUSTAKA

Fayyad, U., Piatetsky-shapiro, G., & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *AAAI and The MIT Press*, 37-53.

Huda, N. M. (2010). *Aplikasi Data Mining Untuk Menampilkan Informasi Tingkat Kelulusan Mahasiswa (Studi kasus di FMIPA Universitas Diponegoro)* (Skripsi S1 ed.). Semarang, Jawa Tengah, Indonesia: FMIPA Universitas Diponegoro.

Larose, D. T. (2005). *Discovery Knowledge in Data: An Introduction to Data Mining.* Hoboken: Wiley-Interscience, John Wiley & Sons Inc.

Santosa, B. (2007). *Data Mining: Teknik Pemanfaatan Data Untuk Keperluan Bisnis.* Yogyakarta, DIY, Indonesia: Graha Ilmu.

Susanto, S., & Suryadi, D. (2010). *Pengantar Data Mining.* Yogyakarta: Andi Publisher.

Syaifullah, M. (2010). *Implementasi Data Mining Algoritma Apriori Pada Sistem Penjualan* (Skripsi S1 ed.). Yogyakarta, DIY, Indonsesia: Teknik Informatika, AMIKOM.

Yusuf. (2006). Penerapan Data Mining Dalam Penentuan Aturan Asosiasi Antar Jenis Item. *SNATI (Seminar Nasional Teknologi Informasi)* (pp. E53-E56). Yogyakarta: Universitas Islam Indonesia.