

# IMPLEMENTASI DATA MINING PENJUALAN MOTOR PADA PT.THAMRIN BROTHER DI KABUPATEN LAHAT METODE INTERPOLASI

Evantri Septiawan<sup>1</sup>, A Yani Ranius<sup>2</sup>, Hutrianto<sup>3</sup>,  
Dosen Universitas Bina Darma<sup>2,3</sup>, Mahasiswa Universitas Bina Darma<sup>2</sup>  
Jalan Jenderal Ahmad Yani No.12 Palembang  
Pos-el : [3van3aslat@gmail.com](mailto:3van3aslat@gmail.com)

---

**Abstract :** PT . Thamrin Brother is a company engaged in the sale of the motor . In the face of competition and increase the company's revenue , the company should be able to take the right decision in determining the marketing strategy that will sell their products . Availability of abundant data , the need for information to support the decision to create business solutions and infrastructure support in the areas of information technology is the forerunner of the birth of data mining technology . Data mining is intended to provide a real solution for decision makers in the business world , to develop their business . One method of data mining that is used in this study is the grouping ( clustering ) , it is identify objects that have certain characteristics in common , and then used as a " characteristic vector " or " centroid " . The results of this research is the application that will make analyze of data that help provide information as a basic feature of corporate decisions .

**Keywords :** business competition , business solutions , information technology , data mining , clustering .

**Abstrak :** PT. Thamrin Brother merupakan perusahaan yang bergerak dibidang penjualan motor. Dalam menghadapi persaingan bisnis dan meningkatkan pendapatan perusahaan, maka perusahaan harus dapat mengambil keputusan yang tepat dalam menentukan strategi pemasaran produknya yang akan dijualnya. Ketersediaan data yang melimpah, kebutuhan akan informasi sebagai pendukung pengambilan keputusan untuk membuat solusi bisnis, dan dukungan infrastruktur di bidang teknologi informasi merupakan cikal-bakal dari lahirnya teknologi data mining. Data mining dimaksudkan untuk memberikan solusi nyata bagi para pengambil keputusan di dunia bisnis, untuk mengembangkan bisnis mereka. Salah satu metode yang terdapat dalam data mining yang digunakan dalam penelitian ini adalah pengelompokan (Clustering) dimana metode tersebut mengidentifikasi objek yang memiliki kesamaan karakteristik tertentu, dan kemudian digunakan sebagai "vektor karakteristik" atau "centroid". Hasil dari penelitian ini adalah aplikasi yang dapat mempermudah menganalisis sejumlah data yang membantu memberikan informasi berharga sebagai gambaran dasar pengambilan keputusan perusahaan.

**Kata Kunci :** persaingan bisnis, solusi bisnis, teknologi informasi, data mining, clustering.

---

## 1 PENDAHULUAN

Perkembangan dunia teknologi informasi saat ini semakin cepat memasuki berbagai bidang, kebutuhan akan informasi yang akurat sangat dibutuhkan dalam kehidupan sehari-hari, sehingga informasi akan menjadi suatu hal penting dalam perkembangan

masyarakat saat ini dan waktu mendatang. Namun kebutuhan informasi yang tinggi kadang tidak diimbangi dengan penyajian informasi yang memadai, seringkali informasi tersebut masih harus digali ulang dari data yang jumlahnya sangat besar. Kemampuan teknologi informasi untuk

mengumpulkan dan menyimpan berbagai tipe data jauh meninggalkan kemampuan untuk menganalisis, meringkas dan mengekstrak pengetahuan dari data. Pemanfaatan data yang ada di dalam sistem informasi untuk menunjang kegiatan pengambilan keputusan, tidak cukup hanya mengandalkan data operasional saja, diperlukan suatu analisis data untuk menggali potensi-potensi informasi yang ada. Para pengambil keputusan berusaha untuk memanfaatkan gudang data yang sudah dimiliki untuk menggali informasi yang berguna membantu mengambil keputusan, hal ini mendorong munculnya cabang ilmu baru untuk mengatasi masalah penggalian informasi atau pola yang penting atau menarik dari data dalam jumlah besar.

Persaingan disuatu perusahaan yang penuh persaingan membuat para pelakunya harus selalu memikirkan strategi-strategi terobosan yang dapat menjamin kelangsungan bisnis mereka. Pada suatu perusahaan memiliki ketersediaan data yang melimpah. Ini melahirkan kebutuhan akan adanya teknologi yang dapat memanfaatkannya untuk membangkitkan pengetahuan-pengetahuan baru, yang dapat membantu dalam pengaturan strategi dalam menjalankan bisnis. Prediksi minat konsumen akan sangat penting bagi

perusahaan, dimana dengan adanya prediksi minat konsumen perusahaan dapat mengambil suatu keputusan atau strategi yang benar dan tepat bagi konsumennya.

Begitu juga dengan PT.Thamrin Brothers, yang dahulunya merupakan CV. Thamrin Bersaudara. Pada mulanya CV. Thamrin Bersaudara merupakan sebuah toko yang bergerak dibidang penjualan sepeda dan sepeda motor berbagai merek, antara lain : merek Mitsubishi, Binter, Honda, Verspa, Yamaha dan sebagainya. Toko tersebut bernama "Toko Singasari", yang berdiri pada tahun 1965 dan beralokasi di Jalan Lembayung Lahat . Toko tersebut selain menjual sepeda dan berbagai jenis sepeda motor juga menjual berbagai merk ban kendaraan bermotor empat, diantaranya : merk Goodyear, Intirub, dan lain-lain. pada saat ini toko Singasai tersebut hanyalah merupakan toko pengecer biasa, belum menjadi agen atau distributor penyalur resmi dari salah satu produk diatas.

Teknologi data mining hadir sebagai solusi. Aplikasi data mining ini memanfaatkan data masukan berupa data penjualan. Dari data tersebut, akan diolah dengan metode clustering dan untuk menganalisis data menggunakan teknik interpolasi. Setelah melewati proses tersebut, maka akan didapatkan pola-pola dalam pengambilan keputusan. Pada PT.Thamrin Brother yang merupakan perusahaan bergerak dibidang penjualan motor. Kendaraan bermotor yang dijual di khususkan untuk kendaraan Yamaha, kendaraan yang di jual oleh PT.Thamrin Brother terdapat banyak sekali tipe-tipe jenis motor. Oleh karena itu, pihak perusahaan sering mengalami kendala dalam penentuan jumlah penyediaan tipe barang karena konsumen memiliki selera berbeda dalam tipe pembelian motor. Oleh karena itu, dibutuhkan teknologi data mining untuk memaksimalkan kinerja perusahaan dalam memperkirakan jumlah barang yang akan laku sesuai dengan tipe motor pada periode yang akan datang sehingga perusahaan dapat menyediakan barang sesuai dengan perkiraan yang tepat.

Aplikasi data mining dengan metode interpolasi untuk memprediksi minat konsumen dalam pembelian tipe motor yang diinginkan konsumen. Berdasarkan permasalahan diatas maka

penulis tertarik untuk melakukan penelitian lebih lanjut yaitu **"Implementasi Data Mining Penjualan PT.Thamrin Brother Di Kabupaten Lahat Dengan Menggunakan Metode Interpolasi"**.

## 2 METODOLOGI PENELITIAN

### 2.1 Metode Data Mining pada Proses Knowledge Discovery in Databases (KDD)

Terdapat banyak algoritma yang digunakan untuk *clustering*. Salah satunya algoritma *K-Means*. Algoritma ini merupakan algoritma *clustering* data yang setiap datanya menjadi anggota dari suatu kluster dengan derajat didefinisikan dengan level keanggotaan. Proses *mining* untuk informasi mengenai data transaksi berdasarkan data yang sudah ada yang telah dipilih. Untuk proses yang dilakukan *clustering* mempunyai 6 langkah proses dalam menentukan nilai yang telah dipilih. Berikut langkah-langkah dari *clustering* :

$$I_m = \sum_{i=1}^N \sum_{j=1}^C (u_{ij})^m \|z_i - c_j\|^2$$

Dengan :

$u_{ij}$  merupakan level keanggotaan dari  $z_i$  dalam cluster  $j$

$z_i$  merupakan nilai data ke- $i$  dari  $d$ -dimensi data

$c_j$  merupakan nilai- $j$  dari  $d$ -dimensi cluster center

$m$  merupakan sembarang bilangan real lebih besar dari 1

Selanjutnya algoritma FCM disusun dengan langkah sebagai berikut :

Langkah1 : Tentukan himpunan data  $Z$ . Tentukan jumlah *cluster* yang diharapkan  $1 < c < N$ , nilai pembobot  $m > 1$ , toleransi penghentian  $\epsilon > 0$ .

Langkah2 : Inisialisasi matriks partisi secara acak,  $U \in M_{fc}$ . Ulangi untuk  $k = 3, 4, 5$ .

Langkah3 : Hitung cluster center (means).

$$v_i^{(l)} = \frac{\sum_{k=1}^N (u_{ik}^{(l-1)})^m z_k}{\sum_{k=1}^N (u_{ik}^{(l-1)})^m}, 1 \leq i \leq c$$

Langkah 4 : Hitung jarak

$$v_{ikA}^{(2)} = (z_k - v_i^{(1)})^T A (z_k - v_i^{(1)}), 1 \leq i \leq c, 1 \leq k \leq N$$

Langkah 5 : Perbaharui matriks partisi.

Untuk  $1 \leq k \leq N$

Jika  $D_{ikA} > 0$  untuk semua  $i = 1, 2, \dots, c$

$$u_{ik}^{(l)} = \frac{1}{\sum_{j=1}^c (D_{ikA} / D_{ikA})^{2/(m-1)}}$$

Atau dengan kata lain :

$$u_{ik}^{(l)} = 0 \text{ jika } D_{ikA} > 0 \text{ dan } u_{ik}^{(l)} \in [0, 1]$$

$$\text{dengan } \sum_{i=1}^c u_{ik}^{(l)} = 1$$

Ulangi sampai  $\|z_i - c_j\| < \epsilon$ .

## 2.2 Interpolasi Newton

Interpolasi Newton dinyatakan (Kurniawan Hidayat, 2007 : 116) :

$$P_n(\mu) = a_0 + a_1(\mu - \mu_0) + a_2(\mu - \mu_0)(\mu - \mu_1) + \dots + a_n(\mu - \mu_0)\dots(\mu - \mu_{n-1})$$

Dengan nilai konstanta  $a_0, a_1, \dots, a_n$  yang bersesuaian. Maka fungsi  $P_n(\mu)$  saat  $\mu = \mu_0$  adalah titik yang dikenai proses *interpolasi* dan mempunyai suatu konstanta yaitu  $a_0$

$$a_0 = P_n(\mu) = f(x_0)$$

dengan cara yang sama pada  $\mu = \mu_1$  Fungsi  $P_n(\mu_1)$  adalah

$$f(x_0) + a_1(\mu - \mu_0) = P_n(\mu) = f(x_1)$$

$$a_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

Metode *Divided difference* untuk ke- $k$

$$f[\mu_i, \mu_{i+1}, \dots, \mu_i + k-1, \mu_i + k] = \frac{f[\mu_i+1, \mu_i+k] - f[\mu_i, \mu_i+k-1]}{\mu_i+k - \mu_i}$$

Maka bentuk *Polinomial Newton*, menjadi

$$P_n(\mu) = f[\mu_0] + f[\mu_0, \mu_1](\mu - \mu_0) + a_1(\mu - \mu_0)(\mu - \mu_1) + \dots + a_n(\mu - \mu_0)(\mu - \mu_1) \dots (\mu - \mu_{n-1})$$

$$P_n(\mu) = f[\mu_0] + \sum_{k=1}^n f[\mu_0, \mu_1, \dots, \mu_k](\mu - \mu_0) \dots (\mu - \mu_{k-1})$$

Persamaan diatas dikenal Newton Interpolatory divided difference formula.

### 2.3 Algoritma Interpolasi Newton

Algoritma interpolasi newton sebagai berikut (Kurniawan Hidayat, 2007 : 116) :

Input Number  $\mu_0, \mu_1, \dots, \mu_n$  ;  
value  $f(x_0), f(x_1), \dots, f(x_n)$  Output The  
Number  $F_{0,0}, F_{1,1}, \dots, F_{n,n}$  where

$$P(\mu) = \sum_{j=0}^{i-1} F_{i,j} [(\mu - \mu_i)]$$

Step 1 For  $i = 1, 2, \dots, n$

For  $j = 1, 2, \dots, i$

$$\text{Set } F_{i,j} = \frac{F_{i,j-1} - F_{i-1,j-1}}{\mu_i - \mu_{i-j}}$$

Stop

### 2.4 XLMiner

XLMiner adalah hanya *comprehensive data mining add-in* pada *excel*. Dengan *neural nets, classification and regression trees, bayes classifier, k-nearest neighbors, discriminant*

*analysis, association rules, clustering, principal components*, dan lainnya.. XLMiner menyediakan segala yang dibutuhkan untuk sample data dari beberapa sumber *powerpivot, Microsoft/IBM/Oracle database*, atau *spreadsheet*, mengeksplorasi dan memvisualisasikan data dengan grafik berhubungan dengan *preprocess* dan *cleaning data, fit model data mining*, dan mengevaluasi daya prediksi.

### 2.5 Analisis

#### a. Data Selection

Data yang digunakan dalam penelitian ini berasal dari perusahaan PT. Thamrin Brother Lahat yaitu data transaksi penjualan tahun 2012, terdiri dari tabel transaksi penjualan. Tabel transaksi penjualan berisi tentang informasi data penjualan. Dari data tabel transaksi penjualan ada beberapa atribut yaitu atribut nama barang yang menjelaskan produk barang dan atribut perbulan mulai dari bulan januari sampai bulan desember dengan menggunakan data transaksi tahun 2012 pada PT. Thamrin Brother Lahat.

Dari transaksi penjualan pada tahun 2012 seperti tabel 3.1 menjelaskan beberapa produk motor PT. Thamrin Brother Lahat.

**Tabel 3.1** Data Penjualan Sales



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Nama Sales	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec		
2	MARTINA RATNA NINGRUM	52	65	65	40	50	46	53	40	33	27	18	19		
3	IRKA DWILIAROZA	45	35	2	17	25	27	28	20	23	22	13	15		
4	PANSEN (24 JUNI 2006)	10	11	9	2	17	5	3	10	0	8	8	0		
5	SEFTA MIKEYANSYAH (02 MEI 2011)	10	19	11	7	6	12	15	16	5	9	18	6		
6	KUSNANDAR (01 MARET 2011)	7	4	7	9	7	6	9	16	9	7	9	4		
7	SIGIT DIDIK WIBOWO (02 AGUSTUS 2010)	11	8	3	2	7	10	9	9	7	3	3	1		
8	SUHERIYADI (02 AGUSTUS 2011)	8	11	3	5	1	2	4	9	9	4	9	3		
9	MEKA SILVIA (01 MARET 2011)	7	4	11	6	6	6	7	1	3	1	1	1		
10	BERLIANSYAH (02 MEI 2011)	7	3	11	4	8	8	9	8	7	2	2	6		
11	AHMAD IKBAL (02 MEI 2011)	2	1	6	8	9	6	2	11	8	8	2	2		
12	FEBRI KURNAWAN (05 SEPTEMBER 2011)	5	0	6	7	8	8	12	4	3	0	3	1		
13	ARI WAHYUDI (01 OKTOBER 2011)	5	4	6	8	7	5	7	9	3	9	0	4		
14	ZAINUL ABIDIN (01 AGUSTUS 2011)	6	7	7	6	10	7	11	7	4	3	4	7		
15	JEFRI SAPUTRA (N 15/11/2011)	4	6	6	6	7	7	8	11	7	9	5	3		
16	CHARLES KRISTIAN (N 19/11/2011)	5	6	1	6	7	7	7	8	4	3	7	0		
17	BERLAN W.KUSUMA (N 15/11/2011)	3	6	0	0	12	16	17	9	7	6	1	0		
18	YUDA PRAYOGI (N 02/01/2012)	3	2	7	0	15	17	21	16	14	13	6	6		
19	HENKY PUTRAWAN	14	18	10	7	13	9	5	18	19	15	13	10		

Pada tahapan penelitian ini penulis menambahkan atribut baru yaitu atribut jumlah untuk menentukan banyaknya transaksi penjualan barang yang terjual, yang di dapat melalui rumus berikut =  $\text{SUM (Data Transaksi Penjualan dari bulan Jan-Des)}$  maka munculah jumlah perhitungan transaksi penjualan dari bulan januari sampai bulan desember.

Pada tahap *transformasi* ini, data akan diubah menjadi bentuk yang sesuai untuk di mining, karena sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data. Untuk tahap ini maka penulis menambahkan atribut jumlah untuk menentukan banyaknya masing-masing transaksi penjualan. Berikut tabel jumlah dari masing-masing penjualan sales.

**Tabel 3.3** Penambahan atribut jumlah

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Nama Sales	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jumlah	
2	MARTINA RATNA NINGRUM	52	65	65	40	50	46	53	40	33	27	18	19	508	
3	IRKA DWILIAROZA	45	35	2	17	25	27	28	20	23	22	13	15	272	
4	PANSEN (24 JUNI 2006)	10	11	9	2	17	5	3	10	0	8	8	0	83	
5	SEFTA MIKEYANSYAH (02 MEI 2011)	10	19	11	7	6	12	15	16	5	9	18	6	134	
6	KUSNANDAR (01 MARET 2011)	7	4	7	9	7	6	9	16	9	7	9	4	94	
7	SIGIT DIDIK WIBOWO (02 AGUSTUS 2010)	11	8	3	2	7	10	9	9	7	3	3	1	73	
8	SUHERIYADI (02 AGUSTUS 2011)	8	11	3	5	1	2	4	9	9	4	9	3	68	
9	MEKA SILVIA (01 MARET 2011)	7	4	11	6	6	6	7	1	3	1	1	1	54	
10	BERLIANSYAH (02 MEI 2011)	7	3	11	4	8	8	9	8	7	2	2	6	75	
11	AHMAD IKBAL (02 MEI 2011)	2	1	6	8	9	6	2	11	8	8	2	2	65	
12	FEBRI KURNAWAN (05 SEPTEMBER 2011)	5	0	6	7	8	8	12	4	3	0	3	1	57	
13	ARI WAHYUDI (01 OKTOBER 2011)	5	4	6	8	7	5	7	9	3	9	0	4	67	
14	ZAINUL ABIDIN (01 AGUSTUS 2011)	6	7	7	6	10	7	11	7	4	3	4	7	79	
15	JEFRI SAPUTRA (N 15/11/2011)	4	6	6	6	7	7	8	11	7	9	5	3	79	
16	CHARLES KRISTIAN (N 19/11/2011)	5	6	1	6	7	7	7	8	4	3	7	0	61	
17	BERLAN W.KUSUMA (N 15/11/2011)	3	6	0	0	12	16	17	9	7	6	1	0	77	
18	YUDA PRAYOGI (N 02/01/2012)	3	2	7	0	15	17	21	16	14	13	6	6	120	
19	HENKY PUTRAWAN	14	18	10	7	13	9	5	18	19	15	13	10	151	

Setelah melakukan tahap seleksi data atribut yang relevan yaitu atribut nama sales dan bulan penjualan.

## 2) Integrasi Data

Tahap *integrasi* data adalah tahap penggabungan data dari berbagai sumber. *Dataset* transaksi penjualan PT. Thamrin Brother Lahat ini berasal dari satu sumber, dan dari data yang sudah dapat digunakan untuk teknik *clustrering* yaitu tabel transaksi penjualan.

## 3) Transformation Data

Tahap *transformasi* data yaitu pada tahap ini data diubah menjadi bentuk yang sesuai untuk di *mining*. Karena dalam penelitian ini hanya dilakukan uji coba secara teoritis. Maka data yang telah di *cleaning* di *transformasi* menjadi data yang siap di *mining*.

Setelah dilakukan penambahan atribut jumlah dalam proses *transformasi* maka dihasilkan *dataset transformasi*.

**Tabel 3.4** Sebagian *dataset* hasil transformasi

1	Nama Sales	Jumlah
2	MARTINA RATNA NINGRUM	508
3	IRKA DWI LIAROZA	272
4	PANSEN ( 24 JUNI 2006)	83
5	SEFTA MIKEYANSYAH (02 MEI 2011)	134
6	KUSNANDAR (01 MARET 2011)	94
7	SIGIT DIDIK WIBOWO (02 AGUSTUS 2010)	73
8	SUHERYADI 02 AGUSTUS 2011)	68
9	MEKA SILVIA (01 MARET 2011)	54
10	BERLIANSYAH (02 MEI 2011)	75
11	AHMAD IKBAL (02 MEI 2011)	65
12	FEBRI KURNIAWAN (05 SEPTEMBER 2011)	57
13	ARI WAHYUDI (01 OKTOBER 2011)	67
14	ZAINUL ABIDIN ( 01 AGUSTUS 2011)	79
15	JEFRI SAPUTRA (IN 15/11/2011)	79
16	CHARLES KRISTIAN (IN 19/11/2011)	61
17	BERLAN W.KUSUMA (IN 15/11/2011)	77

Dataset pada tabel 3.4 di atas telah memiliki bentuk yang sesuai untuk tahap data *mining* menggunakan teknik *clustering*, data transaksi penjualan yang telah di lakukan p roses penjumlahan maka data di atas siap untuk dilakukan proses selanjutnya.

### 3. HASIL

#### 3.1. Hasil Data Mining

*Data mining* merupakan tahapan untuk menemukan pola atau informasi dalam sekumpulan data dengan menggunakan teknik dan algoritma tertentu. Pemilihan

teknik dan algoritma yang tepat sangat tergantung pada proses KDD secara keseluruhan.

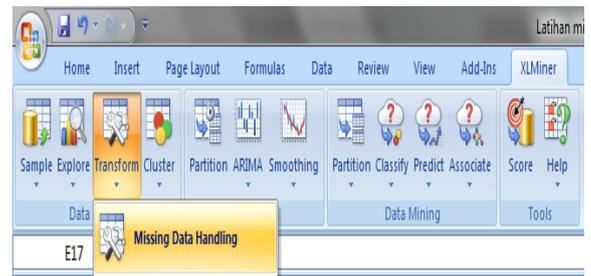
#### 3.2. Penerapan Metode Clustering dengan Algoritma K-means

Penerapan metode *clustering* dengan algoritma *k-means* pada PT. Thamrin Brother Lahat dalam mengolah data penjualan saat ini sudah menggunakan sistem komputer. Untuk yang sudah menggunakan sistem komputerisasi dimana setiap ada penjualan motor, maka pegawai akan langsung memasukkan datanya di komputer.

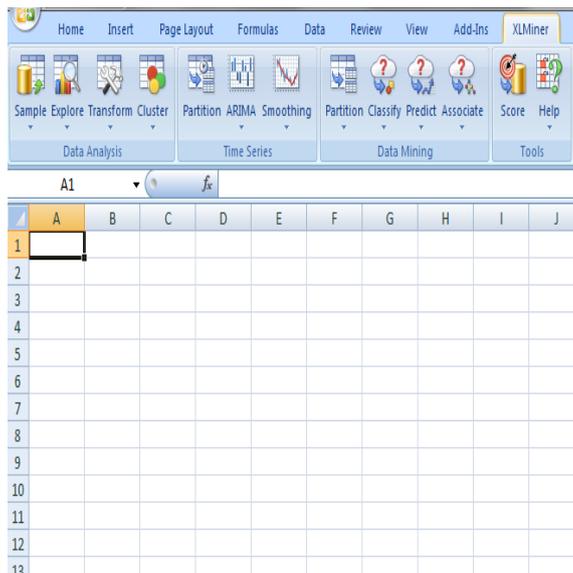
Tabel 4.1 Data Penjualan Sales

Setelah dilakukan penambahan atribut jumlah, maka proses selanjutnya adalah *transformasi*. Untuk tahap *transformasi* dan seterusnya nanti maka penelitian ini menggunakan *software data mining XLMiner* yaitu *SolverSetup*.

*XLMiner* adalah *comprehensive data mining add-in* pada *excel*. Dengan *neural nets*, *classification* and *regression trees*, *bayes classifier*, *k-nearest neighbors*, *discriminant analysis*, *association rules*, *clustering*, *principal components*, dan lainnya.



Gambar 4.2 Transformasi Missing Data

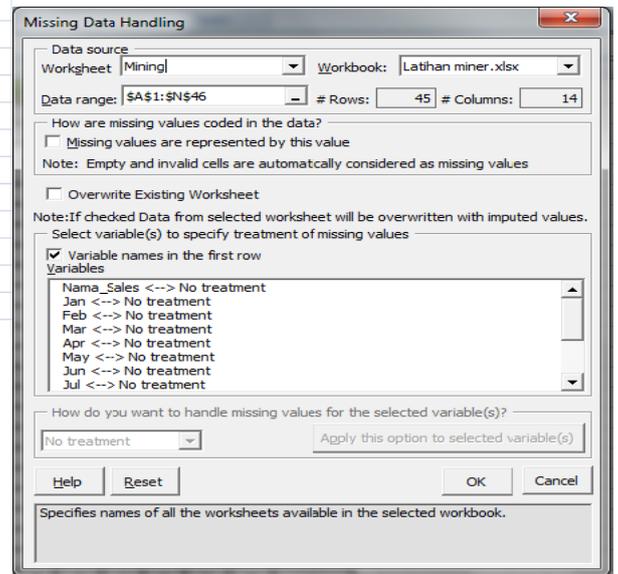


Gambar 4.1 Tampilan *XLMiner*

Tampilan *XLMiner* sama seperti *Microsoft excel* pada umumnya, karena *software SolverSetup* yang telah di install maka ditambahkan *add ini XLMiner* pada *Microsoft Excel 2007* ini.

Untuk proses selanjutnya menggunakan menu *Transformasi Missing data Handling*, seperti gambar dibawah ini.

Setelah memilih menu diatas, maka akan ke proses berikutnya yaitu mengatur proses dan hasil yang diinginkan, seperti gambar dibawah ini.



Gambar 4.3 Setting Missing Data Handling

Setelah gambar itu di proses maka klik tombol ok, maka hasil dari transformasi data seperti pada gambar dibawah ini.

The screenshot shows the 'Missing Data Handling' dialog box in XLMiner. The 'Data Source' is 'MissingData1\$B6:B6'. The 'Selected variables' are 'Nama\_Sales', 'Jan', 'Feb', 'Mar', 'Apr', 'May', 'Jun', 'Jul', 'Aug', 'Sep', 'Oct', 'Nov', 'Dec', and 'Jumlah'. The 'Missing values treatment' is set to 'None' for all variables.

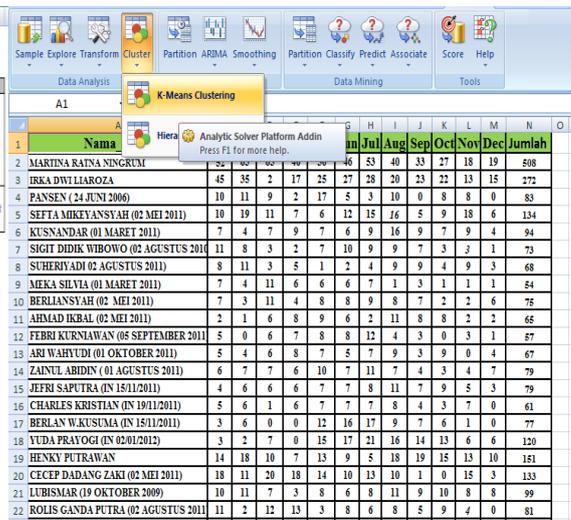
Row #	Nama_Sales	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jumlah
174	MARTINA RATNA NINGRUM	52	65	65	40	50	46	53	40	32	27	18	19	508
175	IRKA DWI LIARORA	45	35	2	17	25	23	28	20	23	22	13	15	272
176	PANSEN (24 JUNI 2006)	10	11	9	2	17	5	3	10	0	8	8	0	83
177	SEFTA MIKEYANSYAH (02 MEI 2011)	10	19	11	7	6	12	15	16	5	9	18	6	134
178	KUSNANDAR (01 MARET 2011)	7	4	7	9	7	6	9	16	9	7	9	4	94
179	SIGIT DIDIK WIBOWO (02 AGUSTUS 2011)	11	8	3	2	7	10	9	9	7	3	3	1	73
180	SUHERYADI (02 AGUSTUS 2011)	8	11	3	5	1	2	4	9	9	4	9	3	68
181	MEKA SILVIA (01 MARET 2011)	7	4	11	6	6	6	7	1	3	1	1	1	54
182	BERLIANSYAH (02 MEI 2011)	7	3	11	4	8	8	9	8	7	2	2	6	75
183	AHMAD IKBAL (02 MEI 2011)	2	1	6	8	9	6	2	11	8	8	2	2	65
184	FEBRI KURNIAWAN (05 SEPTEMBER 2011)	5	0	6	7	8	8	12	4	3	0	3	1	57
185	ARI WAHYUDI (01 OKTOBER 2011)	5	4	6	8	7	5	7	9	3	9	0	4	67
186	ZAINUL ABDIN (01 AGUSTUS 2011)	6	7	7	6	10	7	11	7	4	3	4	7	79
187	JEFRI SAPUTRA (IN 15/11/2011)	4	6	6	6	7	7	8	11	7	9	5	3	79
188	CHARLES KRISTIAN (IN 19/11/2011)	5	6	1	6	7	7	7	8	4	3	7	0	61
189	BERLAN WAKUSTAMA (IN 15/11/2011)	3	6	0	0	12	16	17	9	7	6	1	0	77
190	YUDA PRAYOGI (IN 02/01/2012)	3	2	7	0	15	17	21	16	14	13	6	6	120
191	HENKY PUTRAWAN	14	18	10	7	13	9	5	18	19	15	13	10	151
192	CECEP DADANG ZAKI (02 MEI 2011)	18	11	20	18	14	10	13	10	1	0	15	3	133
193	LUBISMAR (19 OKTOBER 2009)	10	11	7	3	8	6	8	11	9	10	8	8	99
194	ROLIS GANDA PUTRA (02 AGUSTUS 2011)	11	2	12	13	3	8	6	8	5	9	7	0	81

Gambar 4.4 Hasil Missing Data Handling

### 3.3. Data Mining

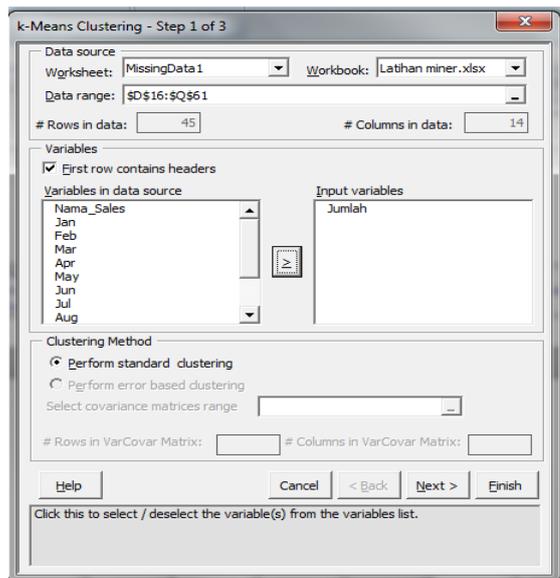
Pada proses data mining disini akan memproses dan mencari pola atau informasi dalam data terpilih dengan menggunakan teknik atau metode tertentu. Untuk penelitian ini penulis menggunakan metode *cluster*, yang berfungsi untuk mengetahui struktur dalam data yang dapat dipakai lebih lanjut dalam *klasifikasi*, pengolahan gambar dan pengolahan pola.

Untuk metode *clustering* dibagi menjadi dua, yaitu *K-Means Clustering* dan *Hierarchical Clustering*. Pada penelitian ini penulis menggunakan metode *K-Means cluster*. Selanjutnya pilih *cluster* yang ada pada menu *toolbar XLMiner*, seperti gambar dibawah ini.



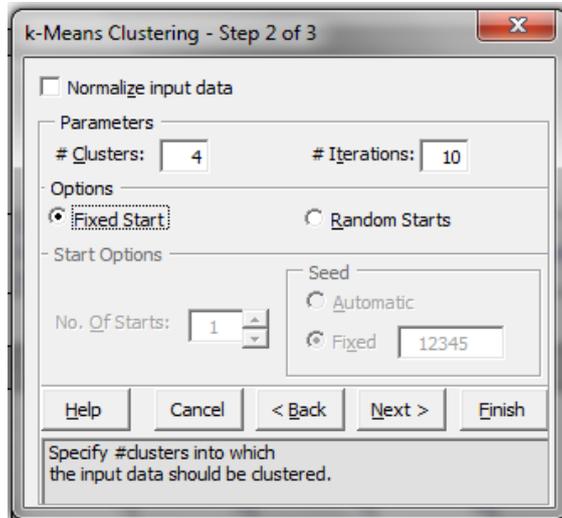
Gambar 4.5 Proses untuk menentukan K-Means

Kemudian gunakan *worksheet* yang telah di *transformasi* tadi yang diberi nama *missingdata1* dan pada *input variables* gunakan yang jumlah karena hasil yang diinginkan adalah keseluruhan penjualan dari nama sales. Seperti gambar dibawah ini.



Gambar 4.6 K-Means Clustering

Setelah itu, tekan tombol next maka akan tampil gambar seperti di bawah ini.



Gambar 4.7 Menentukan jumlah cluster

Untuk gambar diatas terdapat 4 cluster yang menentukan hasil dari nilai transaksi penjualan dan 10 iterations untuk menentukan perulangan cluster, kemudian klik finish.

Setelah semua tahapan clustering dilakukan maka akan tampil output navigator terdiri dari beberapa tabel, yaitu inputs, cluster centers, data summary, elapsed time, predicted clusters.

Tabel 4.4 XLMiner : K-Means Clustering

Output Navigator		
<a href="#">Inputs</a>	<a href="#">Cluster Centers</a>	<a href="#">Data Summ.</a>
<a href="#">Elapsed Time</a>	<a href="#">Predicted Clusters</a>	<a href="#">Random Starts Summary</a>

1. Tabel input maka dihasilkanlah seperti gambar dibawah ini.

Tabel 4.5 Tabel input

Inputs	
<b>Data</b>	
Input data	[Lathan miner.xlsx]*MissingData!*SD\$17-\$Q\$61
# Records in the input data	45
Input variables normalized	No
<b>Variables</b>	
# Selected Variables	1
Selected variables	Jumlah
<b>Parameters/Options</b>	
# Clusters	4
Start Option	Fixed Start
# Iterations	10
Show data summary	Yes
Show distance from each cluster	Yes

Tabel ini adalah sumber dari proses data awal yaitu data penjualan yang dimasukkan ke dalam XLMiner untuk di mining sampai dengan #cluster 4 dan iterations 10.

2. Cluster centers adalah hasil dari jumlah cluster yang telah ditentukan, seperti gambar dibawah ini.

Tabel 4.6 Tabel Cluster Centers

Cluster centers

Cluster	Jumlah
Cluster-1	40.6364
Cluster-2	272
Cluster-3	90.619
Cluster-4	508

Distance between cluster centers	Cluster-1	Cluster-2	Cluster-3	Cluster-4
Cluster-1	0	231.3636	49.9826	467.3636
Cluster-2	231.3636	0	181.381	236
Cluster-3	49.9826	181.381	0	417.381
Cluster-4	467.3636	236	417.381	0

3. Tabel data summary ini terdapat beberapa atribut, yaitu atribut cluster menjelaskan banyaknya cluster 1 sampai dengan cluster 4 sedangkan overall 45 adalah nilai dari jumlah keseluruhan cluster. Untuk atribut #Obs merupakan penjelasan banyak cluster dari setiap record, sedangkan atribut dari average distance in cluster adalah nilai dari rata-rata di dalam setiap cluster.

Tabel 4.7 Tabel Summary

Data summary

Cluster	#Obs	Average distance in cluster
Cluster-1	22	17.636
Cluster-2	1	0
Cluster-3	21	18.766
Cluster-4	1	0
Overall	45	17.38

4. *Elapsed Time* merupakan waktu dalam *K-Means Clustering*.

Tabel 4.8 Tabel *Elapsed Time*

Elapsed Time

Overall (secs)	27.0
----------------	------

Dari proses *XLMiner* yang dilakukan maka seluruh *algorithm K-Means Clustering* menampilkan informasi pengelompokan dari masing-masing *cluster* yang ditentukan

Tabel 4.9 Tabel hasil pengelompokan *cluster*

Row id.	Cluster id	Dist clust-1	Dist clust-2	Dist clust-3	Dist clust-4	Jumlah
1	4	467.36	236	417.38	0	508
2	2	231.36	0	181.38	236	272
3	3	42.364	189	7.619	425	83
4	3	93.364	138	43.381	374	134
5	3	53.364	178	3.381	414	94
6	3	32.364	199	17.619	435	73
7	3	27.364	204	22.619	440	68
8	1	13.364	218	36.619	454	54
9	3	34.364	197	15.619	433	75
10	1	24.364	207	25.619	443	65
11	1	16.364	215	33.619	451	57
12	3	26.364	205	23.619	441	67
13	3	38.364	193	11.619	429	79
14	3	38.364	193	11.619	429	79
15	1	20.364	211	29.619	447	61
16	3	36.364	195	13.619	431	77
17	3	79.364	152	29.381	388	120
18	3	110.36	121	60.381	357	151
19	3	92.364	139	42.381	375	133
20	3	58.364	173	8.381	409	99
21	3	40.364	191	9.619	427	81
22	3	52.364	179	2.381	415	93
23	3	57.364	174	7.381	410	98
24	1	23.364	208	26.619	444	64

3.4 *Interpretation/ Evaluation*

Untuk proses *Interpretation*, *XLMiner* akan menggunakan *toolbar Exponential Smoothing* yang akan

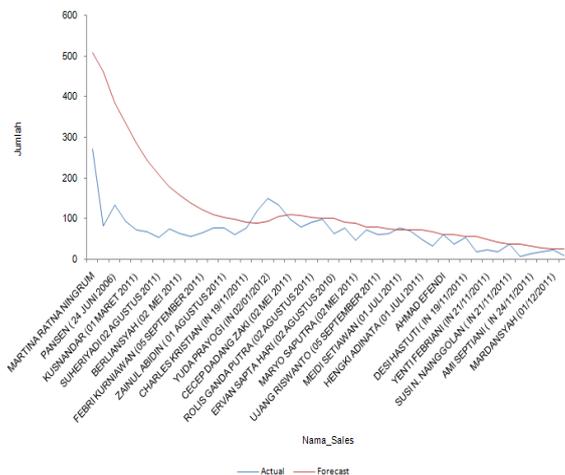
menampilkan atribut *Nama\_Sales*, *Actual*, *Forecast*, *Residuals*. Seperti pada gambar dibawah ini.

Tabel 4.10 Tabel *Fitted Mode*

Nama_Sales	Actual	Forecast	Residuals
MARTINA RATNA NINGRUM	508	*	*
IRKA DWI LIAROZA	272	508	-236
PANSEN ( 24 JUNI 2006)	83	460.8	-377.8
SEFTA MIKEYANSYAH (02 MEI 2011)	134	385.24	-251.24
KUSNANDAR (01 MARET 2011)	94	334.992	-240.992
SIGIT DIDIK WIBOWO (02 AGUSTUS 2010)	73	286.7936	-213.7936
SUHERYADI (02 AGUSTUS 2011)	68	244.03488	-176.03488
MEKA SILVIA (01 MARET 2011)	54	208.8279	-154.8279
BERLIANSYAH (02 MEI 2011)	75	177.86232	-102.86232
AHMAD IKBAL (02 MEI 2011)	65	157.28986	-92.28986
FEBRI KURNIAWAN (05 SEPTEMBER 2011)	57	138.83189	-81.831887
ARI WAHYUDI (01 OKTOBER 2011)	67	122.46551	-55.465509
ZAINUL ABIDIN ( 01 AGUSTUS 2011)	79	111.37241	-32.372408
JEFRI SAPUTRA (IN 15/11/2011)	79	104.89793	-25.897926
CHARLES KRISTIAN (IN 19/11/2011)	61	99.718341	-38.718341
BERLAN W.KUSUMA (IN 15/11/2011)	77	91.974673	-14.974673
YUDA PRAYOGI (IN 02/01/2012)	120	88.979738	31.020262
HENKY PUTRAWAN	151	95.183791	55.816209
CECEP DADANG ZAKI (02 MEI 2011)	133	106.34703	26.652968
LUBISMAR (19 OKTOBER 2009)	99	111.67763	-12.677626
ROLIS GANDA PUTRA (02 AGUSTUS 2011)	81	109.1421	-28.142101
FEBRIANSYAH (01 JULI 2010)	93	103.51368	-10.513681
ERVAN SAPTA HARI (02 AGUSTUS 2010)	98	101.41094	-3.4109445
DIDIK PRAYOGO (01 AGUSTUS 2011)	64	100.72876	-36.728756
MARYO SAPUTRA (02 MEI 2011)	78	93.383004	-15.383004
AHMAD R ZAZALI (02 MEI 2011)	47	90.306404	-43.306404
UJANG RISWANTO (05 SEPTEMBER 2011)	73	81.645123	-8.6451229
NURHAYADI (01 OKTOBER 2011)	62	79.916098	-17.916098
MEIDI SETIAWAN (01 JULI 2011)	64	76.332879	-12.332879
DEPI ARDIANTO (01 AGUSTUS 2011)	78	73.866303	4.1336971
HENGI ADINATA (01 JULI 2011)	70	74.693042	-4.6930423
DAHRI LUBIS (IN 21/11/2011)	51	73.754434	-22.754434
AHMAD EFENDI	34	69.203547	-35.203547

Atribut nama sales didapat dari keseluruhan data nama sales penjualan motor sedangkan *actual* adalah jumlah penjualan motor dari masing-masing sales. Atribut *residuals* merupakan hasil sisa dari atribut *actual* dikurang *forecast*.

Untuk lebih dimengerti, maka hasil *Interpretation* saya buat berupa grafik dalam menghasilkan suatu informasi yang baru. Seperti pada gambar dibawah.



Gambar 4.8 Grafik dari data penjualan sales

Hasil dari grafik diatas adalah informasi banyaknya jumlah hasil dari penjualan sales yang terjual. Pada grafik terdiri dari 2 warna yaitu biru dan merah, untuk yang warna biru adalah actual yang berarti informasi dari transaksi penjualan sales yang banyak terjual. Untuk nilai tertinggi ada pada martina ratna ningrum dengan banyak penjualan  $\geq 500$ .

#### 4. KESIMPULAN

Kesimpulan yang saya dapatkan dari pembahasan ini antara lain :

1. Sistem yang dibangun dapat membantu PT. Thamrin Brother di Kabupaten Lahat sebagai gambaran untuk pengambilan keputusan perusahaan dalam rangka mendapatkan pola penjualan motor yang dihasilkan.

2. Sistem yang dibangun dapat mengurangi penumpukkan data yang kurang dimanfaatkan sebelumnya.

## DAFTAR RUJUKAN

- Anhar. (2010). *PHP dan MySQL Secara Otodidak*. Jakarta Selatan : Media Kita Ilustrasi.
- Hariyanto, Bambang. (2004). *"Rekayasa Sistem Berorientasi Objek"*. Bandung: INFORMATIKA.
- Iko, Pramudiono. (2012). *Pengantar Data Mining: Menambang Permata Pengetahuan di Gunung Data*, Diakses Pada 4 Maret 2012 dari ([http://ilmu.komputer.org/2008-11/25/-pengantar data mining](http://ilmu.komputer.org/2008-11/25/-pengantar-data-mining)).
- Krisnawati. (2007). *"Implementasi Interpolasi Lagrange Untuk Prediksi Nilai Data Berpasangan Dengan Menggunakan Matlab"*, STMIK AMIKOM .
- Kurniawan, Hidayat. (2007). *"Penerapan Data Mining Dengan Metode Interpolasi Untuk Memprediksi Minat Konsumen Asuransi"*, Universitas Islam Indonesia.
- Nugroho, Adi. (2005). *Analisis Dan Perancangan Sistem Informasi Dengan Metodologi Berorientasi Objek*. Bandung : Informatika.