

## BAB V

### HASIL ANALISIS DATA DENGAN WEKA

#### 5.1 PRAPROSES DATA

Pada tahapan awal ini penulis melakukan beberapa persiapan proses data yang meliputi data awal, seleksi data dan konversi hasil data.

##### 5.1.1 Data Awal

Data ini merupakan semua data penjualan pada Toko SRINGATIN, sebelumnya penulis telah menjelaskan data awal pada bab sebelumnya yaitu pada 4.2.1 Reprerentasi data, pada tabel 4.1 penulis telah melampirkan data penjualan pada toko SRINGATIN. Contoh tabel sebagai berikut:

**Tabel 5.1 Data Awal Penjualan Pada Toko Sringatin**

NO	NAMA BARANG	JUMLAH TRANSAKSI	TOTAL PENJUALAN	RATA-RATA PENJUALAN	SATUAN
1	Kecap botol bango	7	10	1	BOTOL
2	Kecap saset sedap	6	6	1	PAK
3	Kecap botol sedap	6	6	1	BOTOL
4	Terasi	3	5	2	PAK
5	Sarden	8	10	1	KALENG
...	...	...	...	...	...
117	wigo botol	10	20	2	DUS
118	Aqua 1500ml	9	9	1	DUS
119	Energen	9	10	1	RENTENG
120	Jahe merah	5	6	1	RENTENG
...	...	...	...	...	...
230	Pena E1000	7	8	1	KOTAK

231	Pensil grible	6	6	1	KOTAK
232	Lasegar kaleng	5	9	2	DUS
233	Lasegar botol	6	7	1	DUS

### 5.1.2 Seleksi Data Awal

Seleksi data awal ini merupakan proses yang dilakukan oleh penulis untuk menyeleksi data dengan cara memfokuskan pada atribut – atribut yang akan digunakan untuk melakukan perhitungan nama barang, jumlah transaksi, total penjualan , dan rata-rata penjualan. Adapun data tersebut dapat dilihat pada tabel 5.2.

**Tabel 5.2 Seleksi Data Awal Toko Sringatin**

NO	NAMA BARANG	JUMLAH TRANSAKSI	TOTAL PENJUALAN	RATA-RATA PENJUALAN
1	Kecap botol bango	7	10	1
2	Kecap saset sedap	6	6	1
3	Kecap botol sedap	6	6	1
4	Terasi	3	5	2
5	Sarden	8	10	1
...	...	...	...	...
117	wigo botol	10	20	2
118	Aqua 1500ml	9	9	1
119	Energen	9	10	1
120	Jahe merah	5	6	1
...	...	...	...	...
230	Pena E1000	7	8	1
231	Pensil grible	6	6	1
232	Lasegar kaleng	5	9	2
233	Lasegar botol	6	7	1

### 5.1.3 Hasil Konversi Data

Data yang terdapat di excel kemudian akan dipindahkan ke notepad dan disimpan dengan format.csv. Hal ini perlu dilakukan untuk tujuan agar dapat digunakan pada aplikasi weka, adapun bentuk data tersebut dapat dilihat pada tabel dan contoh gambar berikut :

**Tabel 5.3 Data Exel Format CSV**

JUMLAH TRANSAKSI	TOTAL PENJUALAN	RATA-RATA PENJUALAN
7	10	1
6	6	1
6	6	1
3	5	2
8	10	1
5	10	2
6	8	1
4	4	1
8	8	1
6	6	1
5	8	2
20	20	1
6	8	1
5	10	2
4	6	2
5	5	1
4	4	1
5	6	1
7	7	1
5	5	1
5	5	1
1	2	2
10	12	1
18	18	1
11	12	1
2	2	1
2	4	2

3	5	2
2	5	3
4	5	1
12	12	1
2	5	3
4	6	2
8	10	1
3	6	2
12	12	1
11	12	1
6	8	1
8	10	1
6	6	1
9	9	1
10	17	2
6	12	2
4	8	2
5	7	1
9	9	1
5	10	2
4	9	2
6	6	1
5	8	2
4	4	1
9	10	1
7	12	2
4	4	1
3	6	2
5	5	1
10	10	1
5	6	1
5	8	2
5	10	2
15	18	1
9	9	1
5	6	1
5	5	1
4	9	2
5	5	1
18	18	1

5	10	2
1	2	2
4	5	1
6	6	1
8	8	1
4	10	3
7	9	1
5	10	2
10	10	1
5	6	1
8	8	1
9	10	1
5	6	1
6	9	2
4	4	1
10	10	1
5	7	1
6	6	1
6	9	2
4	5	1
5	5	1
4	6	2
2	5	3
2	8	4
2	4	2
4	5	1
4	9	2
8	12	2
7	7	1
12	14	1
9	9	1
8	10	1
4	6	2
5	5	1
10	15	2
8	8	1
3	3	1
7	10	1
10	12	1
10	10	1

5	8	2
7	9	1
6	6	1
5	9	2
10	10	1
7	7	1
15	19	1
8	8	1
6	6	1
10	20	2
9	9	1
9	10	1
5	6	1
9	9	1
2	2	1
7	7	1
8	10	1
12	12	1
8	9	1
10	12	1
10	10	1
9	12	1
17	17	1
15	15	1
9	10	1
9	9	1
2	3	2
17	17	1
10	10	1
5	8	2
5	10	2
6	12	2
8	9	1
6	6	1
9	10	1
8	8	1
4	4	1
6	6	1
9	10	1
1	2	2

6	7	1
5	9	2
17	19	1
12	12	1
5	7	1
6	6	1
9	9	1
4	4	1
4	4	1
4	6	2
2	4	2
7	9	1
5	5	1
16	16	1
7	7	1
10	15	2
16	18	1
14	17	1
15	15	1
5	12	2
9	9	1
6	12	2
10	11	1
7	7	1
5	10	2
8	8	1
1	2	2
5	5	1
8	9	1
6	6	1
10	10	1
9	10	1
5	9	2
6	6	1
5	8	2
20	20	1
20	20	1
20	20	1
18	18	1
9	10	1

6	6	1
9	9	1
6	6	1
5	9	2
5	10	2
7	7	1
5	5	1
2	4	2
28	28	1
19	19	1
50	50	1
2	4	2
9	9	1
5	6	1
2	4	2
1	2	2
10	15	2
15	20	1
20	20	1
42	42	1
30	50	2
5	6	1
4	5	1
9	9	1
10	10	1
5	6	1
3	4	1
12	14	1
4	4	1
1	1	1
7	7	1
5	5	1
8	8	1
6	7	1
8	10	1
9	10	1
2	2	1
17	17	1
12	14	1
20	20	1



12	12	1
10	10	1
7	8	1
6	6	1
5	9	2
6	7	1

```

File Edit Format View Help
JUMLAH TRANSAKSI,TOTAL PENJUALAN,RATA-RATA PENJUALAN
7,10,1
6,6,1
6,6,1
3,5,2
8,10,1
5,10,2
6,8,1
4,4,1
8,8,1
6,6,1
5,8,2
20,20,1
6,8,1
5,10,2
4,6,2
5,5,1
4,4,1
5,6,1
7,7,1
5,5,1
5,5,1
1 2 2

```

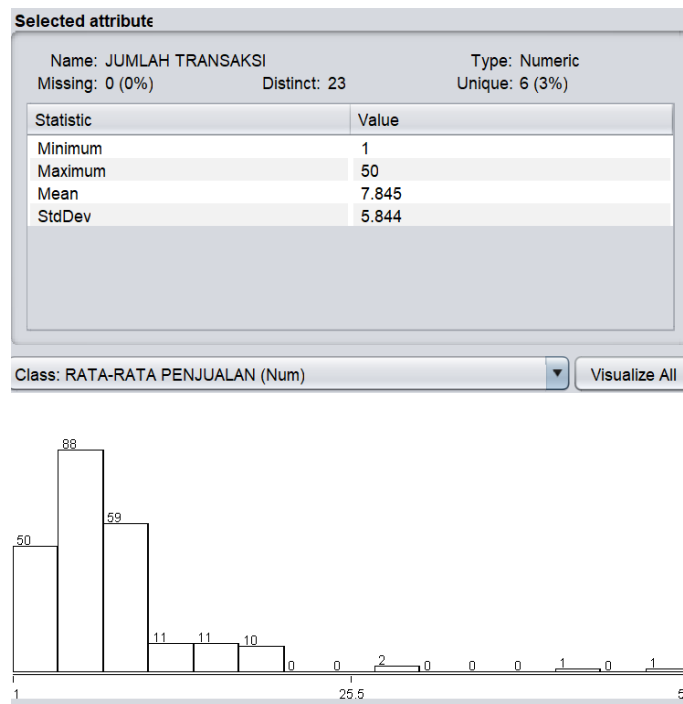
**Gambar 5.1 Hasil Konversi Data Ke Notpad**

## **5.2 HASIL VISUALISASI DATA SETIAP ATRIBUT DENGAN MENGGUNAKAN WEKA**

Berikut merupakan bentuk visualisasi atribut penjualan pada Toko SRINGATIN dengan menggunakan tools Weka :

### 5.2.1 Visualisasi Atribut Jumlah Transaksi

Hasil visualisasi atribut jumlah transaksi dapat dilihat pada gambar 5.2.

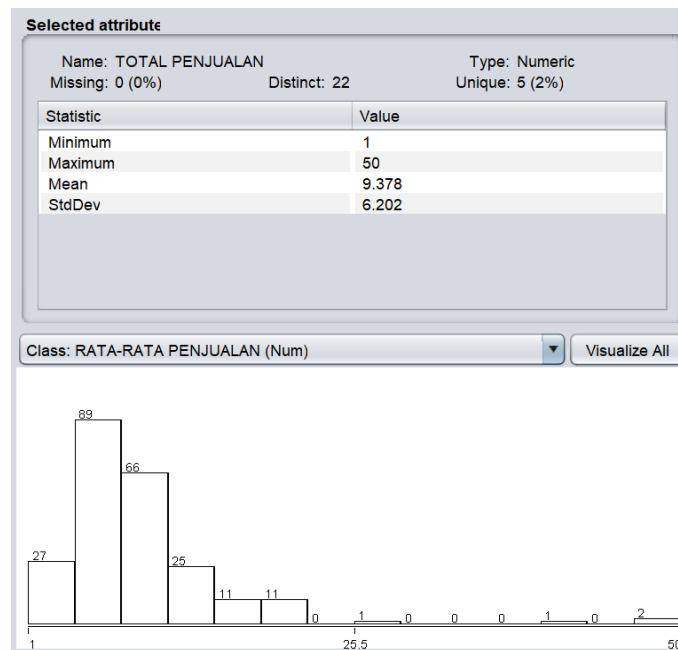


**Gambar 5.2 Visualisasi Atribut Jumlah Transaksi**

Pada gambar 5.4 adalah visualisasi dari atribut jumlah transaksi. Sebanyak 233 data pada *Column Selected Atribut* tidak terdapat missing. *Distinct* 23, *Attribute Unique* 6, *Statistic Minimum* terdapat nilai 1, *Statistic Maximum* terdapat nilai 50, *Statistic Mean* terdapat nilai 7,845, *Standar Deviasi* terdapat nilai 5,844.

## 5.2.2 Visualisasi Atribut Total Penjualan

Hasil visualisasi atribut total penjualan dapat dilihat pada gambar 5.3

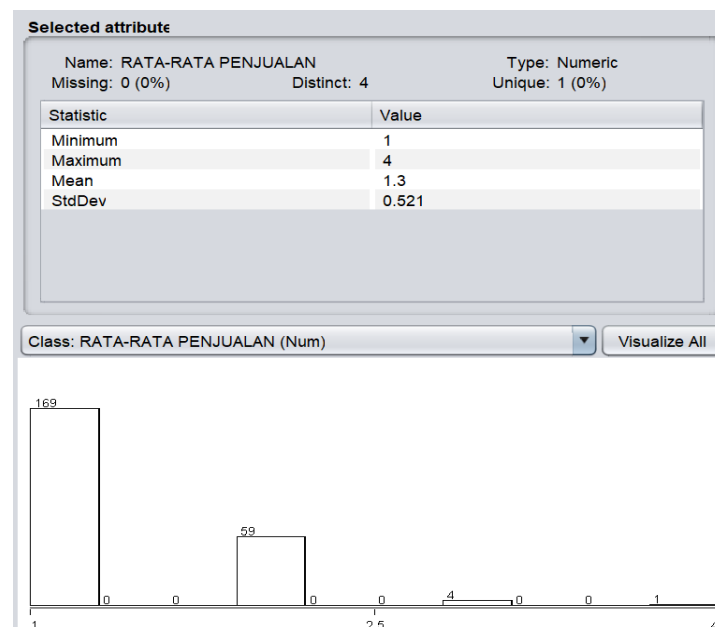


**Gambar 5.3 Visualisasi Atribut Total Penjualan**

Pada gambar 5.3 adalah visualisasi dari atribut total penjualan. Sebanyak 233 data pada Column Selected Attribute tidak terdapat missing. *Distinct* 22, *Attribute Unique* 5, *Statistic Minimum* terdapat nilai 1, *Statistic Maximum* terdapat nilai 50, *Statistic Mean* 9,378 dan *Standar Deviasi* terdapat nilai 6,202.

### 5.2.3 Visualisasi Atribut Rata-rata Penjualan

Hasil visualisasi atribut rata-rata penjualan dapat dilihat pada gambar 5.4



**Gambar 5.4 Visualisasi Atribut Rata-rata Penjualan**

Pada gambar 5.4 adalah visualisasi dari atribut rata-rata penjualan. Sebanyak 233 data pada Column Selected Attribute tidak terdapat missing. *Distinct* 4, *Attribute Unique* 1, *Statistic Minimum* terdapat nilai 1, *Statistic Maximum* terdapat nilai 4, *Statistic Mean* 1,3 dan *Standar Deviasi* terdapat nilai 0,521.

### 5.2.4 Pengambilan Pusat Cluster

Log Listing Coding Pengambilan Pusat Cluster :

⇒ Started weka.clusterers.SimpleKMeans

⇒ Command: weka.clusterers.SimpleKMeans -init 0 -max-candidates

100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 3 -

A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S  
10

⇒ Finished weka.clusterers.SimpleKMeans.

Pengambilan pusat cluster pada *tools* weka dapat dilihat pada gambar 5.5

```

=== Run information ===

Scheme:      weka.clusterers.SimpleKMeans -init 0 -max-candidates 100 -periodic-prun
Relation:    dataweka
Instances:   233
Attributes:  3
              JUMLAHTRANSAKSI
              TOTALPENJUALAN
              RATA-RATAPENJUALAN
Test mode:   evaluate on training data

=== Clustering model (full training set) ===

kMeans
=====

Number of iterations: 6
Within cluster sum of squared errors: 3.948024588096435

Initial starting points (random):

Cluster 0: 12,12,1
Cluster 1: 11,12,1
Cluster 2: 8,8,1

```

### Gambar 5.5 Pengambilan Cluster

Pada gambar 5.5 Pada *tools* weka pengambilan pusat *cluster* diambil secara random dan disebutkan nilainya dari 3 atribut, pada *tools* weka juga dijelaskan berapa banyak jumlah iterasi yang dilakukan yaitu sebanyak 6 iterasi.

### 5.2.5 Hasil Centroid Akhir

Hasil implementasi dari *tools* weka yang menampilkan hasil *Centroid* Akhir, dapat dilihat pada gambar berikut :

```
Missing values globally replaced with mean/mode

Final cluster centroids:

Attribute          Full Data          Cluster#
                   (233.0)           0           1           2
=====
JUMLAHTRANSAKSI   7.8455            20.4615      7.0625      4.4286
TOTALPENJUALAN    9.3777            21.9615      7.6528      8.127
RATA-RATAPENJUALAN 1.3004            1.0385       1           2.0952
```

**Gambar 5.6 Hasil Centroid Akhir**

Pada gambar 5.6 yaitu hasil centroid akhir yang setiap atributnya mendapatkan nilai pada masing – masing cluster .

### 5.2.6 Jumlah Data Dari Setiap Cluster

Hasil implementasi dari *tools* weka yang menampilkan jumlah data dari sertiap *Cluster*, dapat dilihat pada gambar berikut :

```
Time taken to build model (full training data) : 0.03 seconds

=== Model and evaluation on training set ===

Clustered Instances

0          26 ( 11%)
1          144 ( 62%)
2           63 ( 27%)
```

**Gambar 5.7 Jumlah Data Dari Setiap Cluster**

Kemudian diperoleh hasil dengan jumlah setiap clusternya adalah *cluster* 1 sebanyak 26 data yang dikategorikan “Terjual Sedikit” dengan presentase sebesar 11% , *cluster* 2 sebanyak 144 data yang dikategorikan “Terjual Sedang” dengan presentase sebesar 62%, dan *cluster* 3 sebanyak 63 data yang dikategorikan “Terjual Banyak” dengan presentase sebesar 30%.

### 5.3 PERBANDINGAN HASIL (Output) DARI 2 JENIS PROSES PERHITUNGAN K-MEANS CLUSTERING

Perbandingan Hasil (*Output*) dari 2 Jenis Proses Perhitungan *K-Means Clustering* dari beberapa pembahasan mengenai hasil (*output*) yang telah di peroleh baik itu untuk perhitungan *k-means clustering* secara manual maupun menggunakan *tools* data mining sebelumnya. Adapun beberapa perbandingan atau perbedaan hasil yang dapat dilihat secara langsung dari kedua hasil (*output*) yang di dapat, yaitu sebagai berikut :

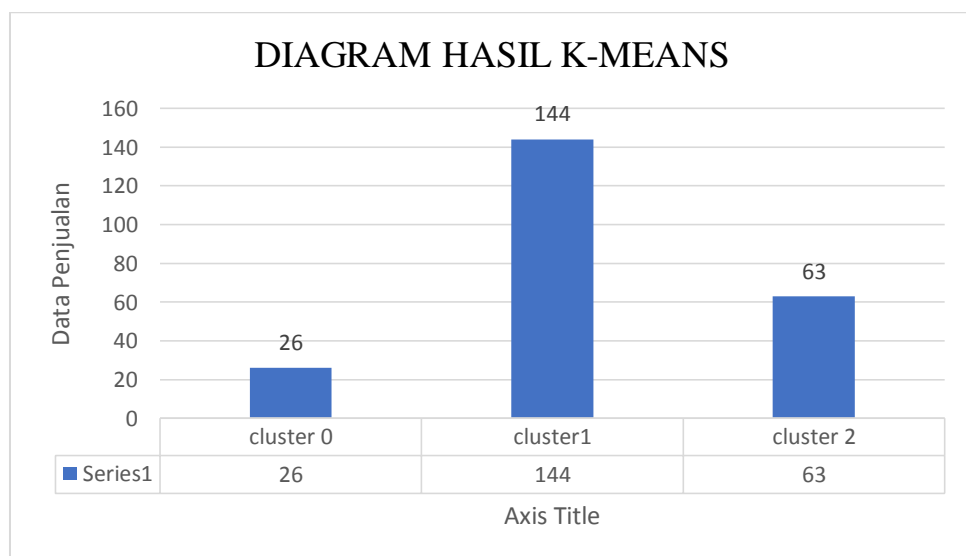
**Tabel 5.4 Perbandingan Hasil Cluster**

Keterangan	Proses	
	Manual	Weka
Jumlah iterasi	4	6
Jumlah Data		
Cluster 0	128	26 (11%)
Cluster 1	83	144 (62%)
Cluster 2	22	63 (27%)
Terjual Banyak	cluster 3	cluster 2
Terjual Sedang	cluster 2	cluster 1
Terjual Sedikit	cluster 1	cluster 0
Centroid Akhir		
Cluster 1		
Jumlah Transaksi	4,539	20,4615
Total Penjualan	6,0859	21,9615

Rata-rata Penjualan	1,4289	1,0385
Cluster 2		
Jumlah Transaksi	9,3373	7,0625
Total Penjualan	10,8192	7,6528
Rata-rata Penjualan	1,1847	1
Cluster 3		
Jumlah Transaksi	21,4545	4,4286
Total Penjualan	23,0909	8,127
Rata-rata Penjualan	1,0776	2,0952

Pada tabel 5.4 yaitu perbandingan antara perhitungan manual dan weka yang di ambil dari centroid akhir mulai dari Cluster 0, Cluster 1, dan Cluster 2, dapat disimpulkan pada diagram berikut :

### 5.3.1 Diagram Hasil Perhitungan Tools Weka



**Gambar 5.8 Diagram Hasil Perhitungan Tools Weka**



1. Dari hasil perhitungan diatas pada *cluster* 1 terdapat 26 barang yang Dikategorikan “Terjual Sedikit” maka, tidak direkomendasikan untuk menambah stok/persediaan barang.
2. Dari hasil perhitungan diatas pada *cluster* 2 terdapat 144 barang yang Dikategorikan “Terjual Sedang” maka, direkomendasikan untuk menambah stok/persediaan barang.
3. Dari hasil perhitungan diatas pada *cluster* 3 terdapat 63 barang yang Dikategorikan “Terjual Banyak” maka, sangat direkomendasikan untuk menambah stok/persediaan barang.