

BAB V

HASIL ANALISIS DAN VISUALISASI

1.1 PRAPROSES DATA

Pada tahap ini dilakukan beberapa persiapan proses data. Persiapan proses data tersebut yaitu :

1.1.1 Data Mentah

Data ini merupakan data pendaftar program bidikmisi sebelum dilakukan penyeleksian data. Data tersebut dapat dilihat pada gambar 5.1 berikut ini:

1	No.	Nama Siswa	NISN	NPSN	Asal Sekolah	Kota Sekolah	Provinsi Sekolah	Jurusan di sekolah	Tempat Lahir
2	1	RAFLI RACHMANDANI	9988809129	10504581	SMAN 5 KOTA JAMBI	Kota Jambi	Jambi		JAMBI
3	2	AAN ADIYANSYAH	9997392181	10501366	MERANGIN	Kab. Merangin	Jambi	IPA	Pengambang
4	3	AAN WAHYUNI	6055537	10504584	SMAN 8 KOTA JAMBI	Kota Jambi	Jambi	12 IPA 1	JAMBI
5	4	AAS NURMAYANG	5407143	10500236	SMAN 6 BATANG HARI	Kab. Batanghari	Jambi	XII IPA	Jambi
6	5	ABD MUTTAKIN	3871560	10500237	SMAN 7 BATANG HARI	Kab. Batanghari	Jambi	12 IPA	TEBING TINGGI
7	6	ABDUL HADI	9991727565	10303106	SMAN 1 TALAMAU	Barat	Sumatera Barat	12 IPS 1	BUKITTINGGI
8	7	ABIMAYU	9986417552	10505463	SMKN 1 KOTA JAMBI	Kota Jambi	Jambi		JAMBI
9	8	ADAM INSANU BAKTI	5573020	10502800	SMAN 9 MUARO JAMBI	Kab. Muara Jambi	Jambi		Bukit Subur
10	9	PRIANTI GAYO	13414137	10647871	SMAN 2 BANYUASIN II	Kab. Banyuasin	Sumatera Selatan		Betung
11	10	ADE INDRİYANI	9986294674	10507896	JAMBI	Kota Jambi	Jambi	12 IPA 2	SUNGAI LANDAI
12	11	NATANOEL	9994496426	10504587	SMAN 11 KOTA JAMBI	Kota Jambi	Jambi	-	Pekan Baru
13	12	ADE RAHMA FITRIA	9990684137	10302399	SMAN 1 SUNGAYANG	Kab. Tanah Datar	Sumatera Barat	IPS	batu SANGKAR
14	13	ADE SUKMA DEWI	20218207	10506165	FANIA SALSABILA	Kota Jambi	Jambi	ANALIS KESEHATAN	Jambi
15	14	ADE SYAFRIANI	7075230	10504585	SMAN 9 KOTA JAMBI	Kota Jambi	Jambi		JAMBI
16	15	ADE YOLANDA	4091718	10500236	SMAN 6 BATANG HARI	Kab. Batanghari	Jambi	XII IPA	Muara Bulian
17	16	NASUTION	19888738	10504587	SMAN 11 KOTA JAMBI	Kota Jambi	Jambi	IPS	Jambi
18	17	ADHE SUCI DESPIANTI	9996759807	10505059	SMAN 12 MERANGIN	Kab. Merangin	Jambi	-	PAMENANG
19	18	KUSUMA SARI	9797520	10504582	SMAN 6 KOTA JAMBI	Kota Jambi	Jambi	-	Jambi
20	19	ADIL FIKRI	6800383	10502388	SMAN 8 KERINCI	Kab. Kerinci	Jambi	IPS	SEBERANG MERANGIN
21	20	FACHRUNISSA	8523060	10504580	SMAN 4 KOTA JAMBI	Kota Jambi	Jambi	IPS	JAMBI
22	21	ADINDA PUTRI	7074811	10504553	SMAN 3 KOTA JAMBI	Kota Jambi	Jambi	12 IPA 4	jambi
23	22	DEWI	7713433	10504587	SMAN 11 KOTA JAMBI	Kota Jambi	Jambi	IPA	Palembang

Gambar 5. 1 Data Mentah Pendaftar Program Bidikmisi

1.1.2 Seleksi Data Mentah

Seleksi data mentah ini merupakan proses penyeleksian data terhadap data yang tidak lengkap dan data yang sama. Seleksi data ini memfokuskan pada atribut-atribut yang digunakan untuk perhitungan. Adapun data tersebut dapat dilihat pada gambar 5.2:

penghasilan	Jumlah	Kepemilikan Rumah	Luas Bangunan	Sumber Air	MCK	Prestasi	kelayaka
sangat rendah	sedikit	Sendiri	sangat kecil	sungai / mata air	Berbagi Pakai	tdk	layak
sangat rendah	sedikit	Sendiri	sangat kecil	sungai / mata air	Kepemilikan Sendiri Didalam	tdk	layak
sangat rendah	sedikit	Sendiri	sangat kecil	sumur	Berbagi Pakai	tdk	layak
sangat rendah	sedikit	Sendiri	sangat kecil	sumur	Kepemilikan Sendiri Diluar	tdk	layak
sangat rendah	sedikit	Sendiri	sangat kecil	PDAM	Berbagi Pakai	tdk	layak
sangat rendah	sedikit	Sendiri	kecil	sumur	Berbagi Pakai	tdk	layak
sangat rendah	sedikit	Sendiri	kecil	sumur	Kepemilikan Sendiri Diluar	ada	layak
sangat rendah	sedikit	Sendiri	kecil	sumur	Kepemilikan Sendiri Diluar	tdk	layak
sangat rendah	sedikit	Sendiri	kecil	sumur	Kepemilikan Sendiri Didalam	ada	layak
sangat rendah	sedikit	Sendiri	kecil	sumur	Kepemilikan Sendiri Didalam	tdk	layak
sangat rendah	sedikit	Sendiri	kecil	PDAM	Kepemilikan Sendiri Didalam	ada	layak
sangat rendah	sedikit	Sendiri	kecil	PDAM	Kepemilikan Sendiri Didalam	tdk	layak
sangat rendah	sedikit	Sendiri	sedang	sumur	Kepemilikan Sendiri Diluar	tdk	layak
sangat rendah	sedikit	Sendiri	sedang	PDAM	Kepemilikan Sendiri Diluar	tdk	layak
sangat rendah	sedikit	Sendiri	besar	sumur	Kepemilikan Sendiri Didalam	tdk	layak
sangat rendah	sedikit	Sendiri	besar	sumur	Kepemilikan Sendiri Diluar	tdk	layak
sangat rendah	sedikit	Sendiri	besar	sungai / mata air	Kepemilikan Sendiri Didalam	tdk	layak
sangat rendah	sedikit	Sendiri	sangat besar	sumur	Kepemilikan Sendiri Didalam	tdk	layak
sangat rendah	sedikit	Menumpang	sangat kecil	sumur	Berbagi Pakai	tdk	layak
sangat rendah	sedikit	Menumpang	sangat kecil	sumur	Kepemilikan Sendiri Diluar	tdk	layak
sangat rendah	sedikit	Menumpang	sangat kecil	PDAM	Kepemilikan Sendiri Didalam	tdk	layak
sangat rendah	sedikit	Menumpang	kecil	sumur	Berbagi Pakai	tdk	layak

Gambar 5. 2 Hasil Seleksi Data Mentah

1.1.3 Konversi Hasil Data

Data yang telah diseleksi kemudian disimpan dengan format .csv, ini dilakukan agar data dapat dimasukkan kedalam *tools* Weka. Adapun bentuk data tersebut dapat dilihat pada gambar 5.3 :

penghasilan, Jumlah Tanggungan, Kepemilikan, Rumah, Luas Bangunan, Sumber Air, MCK, Prestasi Ekstrakurikuler, kelayakan
 sangat rendah, sedikit, Sendiri, sangat kecil, sungai / mata air, Berbagi Pakai, tdk, layak
 sangat rendah, sedikit, Sendiri, sangat kecil, sungai / mata air, Kepemilikan Sendiri Didalam, tdk, layak
 sangat rendah, sedikit, Sendiri, sangat kecil, sumur, Berbagi Pakai, tdk, layak
 sangat rendah, sedikit, Sendiri, sangat kecil, sumur, Kepemilikan Sendiri Diluar, tdk, layak
 sangat rendah, sedikit, Sendiri, sangat kecil, PDAM, Berbagi Pakai, tdk, layak
 sangat rendah, sedikit, Sendiri, kecil, sumur, Berbagi Pakai, tdk, layak
 sangat rendah, sedikit, Sendiri, kecil, sumur, Kepemilikan Sendiri Diluar, ada, layak
 sangat rendah, sedikit, Sendiri, kecil, sumur, Kepemilikan Sendiri Diluar, tdk, layak
 sangat rendah, sedikit, Sendiri, kecil, sumur, Kepemilikan Sendiri Didalam, ada, layak
 sangat rendah, sedikit, Sendiri, kecil, sumur, Kepemilikan Sendiri Didalam, tdk, layak
 sangat rendah, sedikit, Sendiri, kecil, PDAM, Kepemilikan Sendiri Didalam, ada, layak
 sangat rendah, sedikit, Sendiri, kecil, PDAM, Kepemilikan Sendiri Didalam, tdk, layak
 sangat rendah, sedikit, Sendiri, sedang, sumur, Kepemilikan Sendiri Diluar, tdk, layak
 sangat rendah, sedikit, Sendiri, sedang, PDAM, Kepemilikan Sendiri Diluar, tdk, layak
 sangat rendah, sedikit, Sendiri, besar, sumur, Kepemilikan Sendiri Didalam, tdk, layak
 sangat rendah, sedikit, Sendiri, besar, sumur, Kepemilikan Sendiri Diluar, tdk, layak
 sangat rendah, sedikit, Sendiri, besar, sungai / mata air, Kepemilikan Sendiri Didalam, tdk, layak
 sangat rendah, sedikit, Sendiri, sangat besar, sumur, Kepemilikan Sendiri Didalam, tdk, layak
 sangat rendah, sedikit, Menumpang, sangat kecil, sumur, Berbagi Pakai, tdk, layak
 sangat rendah, sedikit, Menumpang, sangat kecil, sumur, Kepemilikan Sendiri Diluar, tdk, layak
 sangat rendah, sedikit, Menumpang, sangat kecil, PDAM, Kepemilikan Sendiri Didalam, tdk, layak
 sangat rendah, sedikit, Menumpang, kecil, sumur, Berbagi Pakai, tdk, layak

Gambar 5. 3 Konversi Hasil Data

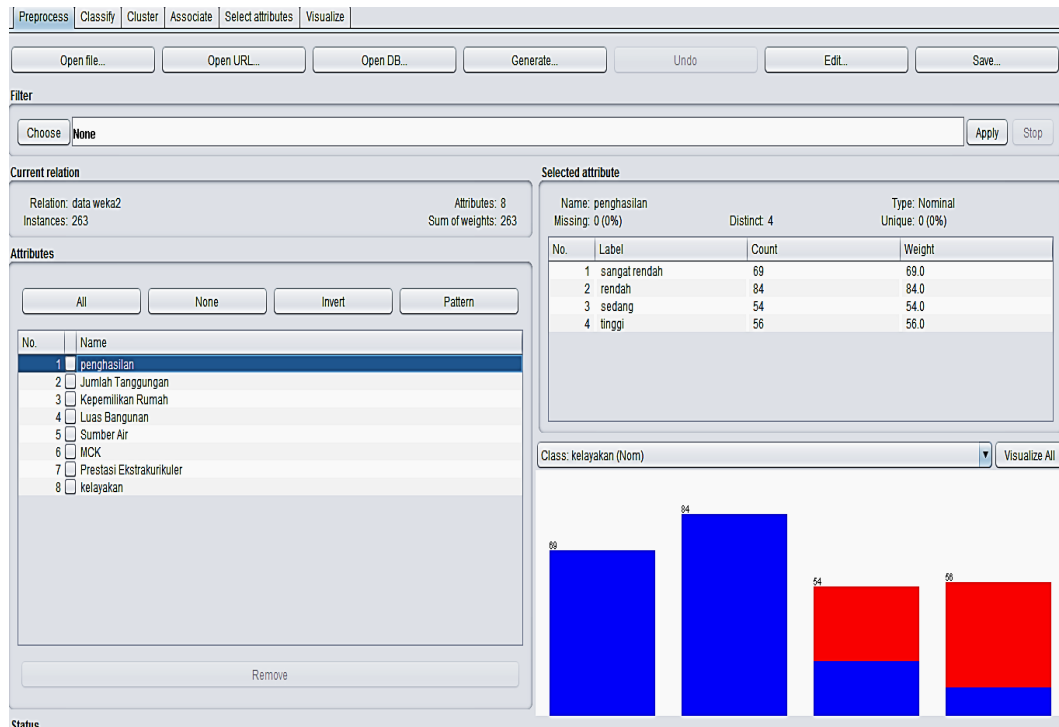
1.2 HASIL VISUALISASI DATA SETIAP ATRIBUT DENGAN MENGGUNAKAN WEKA

Tools Weka digunakan untuk melihat hasil visualisasi setiap atribut dan digunakan untuk mengetahui hasil perhitungan menggunakan algoritma *C4.5*.

Berikut visualisasi pada setiap atribut :

1. Visualisasi Atribut Penghasilan

Hasil visualisasi atribut tanggungan Jenis Kelamin dapat dilihat pada gambar 5.4 berikut ini :

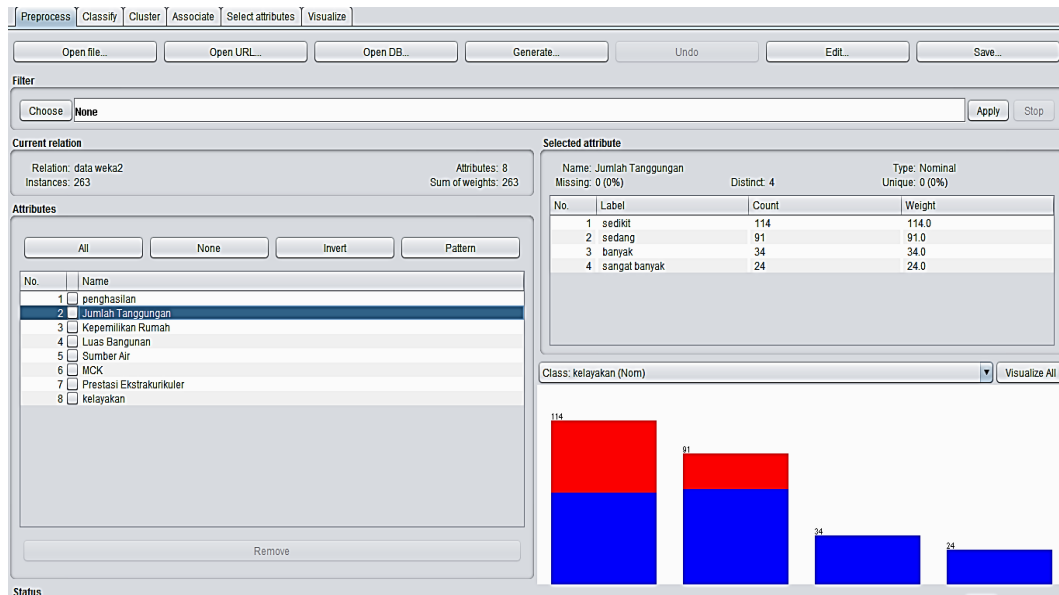


Gambar 5. 4 Visualisasi Atribut Penghasilan

Gambar 5.4 adalah visualisasi dari atribut penghasilan. Sebanyak 263 data pada column *selected atribut* tidak terdapat *missing*. Distinct 4 dan unique 0 dengan jumlah penghasilan sangat rendah sebanyak 69 orang, jumlah penghasilan rendah sebanyak 84 orang, jumlah penghasilan sedang sebanyak 54 orang dan jumlah penghasilan tinggi sebanyak 56 orang.

2. Visualisasi Atribut Jumlah Tanggungan

Hasil visualisasi atribut tanggungan Jenis Kelamin dapat dilihat pada gambar 5.5 berikut ini :

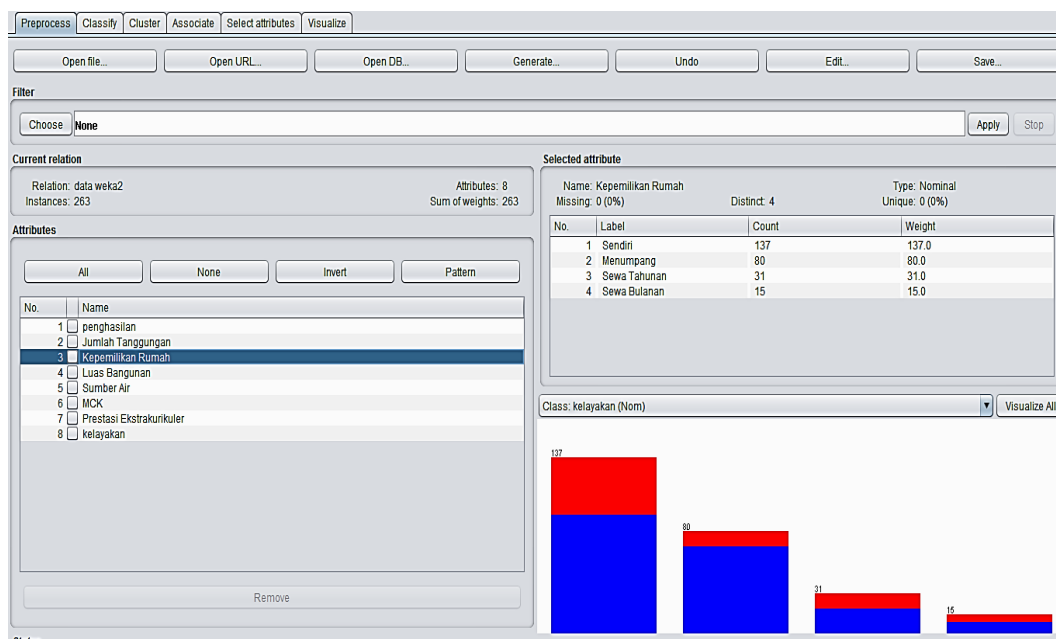


Gambar 5.5 Visualisasi Atribut Jumlah Tanggungan

Gambar 5.5 adalah visualisasi dari atribut jumlah tanggungan. Sebanyak 263 data pada column *selected atribut* tidak terdapat *missing*. Distinct 4 dan unique 0 dengan jumlah tanggungan sedikit sebanyak 114 orang, jumlah tanggungan sedang sebanyak 91 orang, jumlah tanggungan banyak sebanyak 34 orang dan jumlah tanggungan sangat banyak sebanyak 24 orang.

3. Visualisasi Atribut Kepemilikan Rumah

Hasil visualisasi atribut tanggungan Jenis Kelamin dapat dilihat pada gambar 5.6 berikut ini :

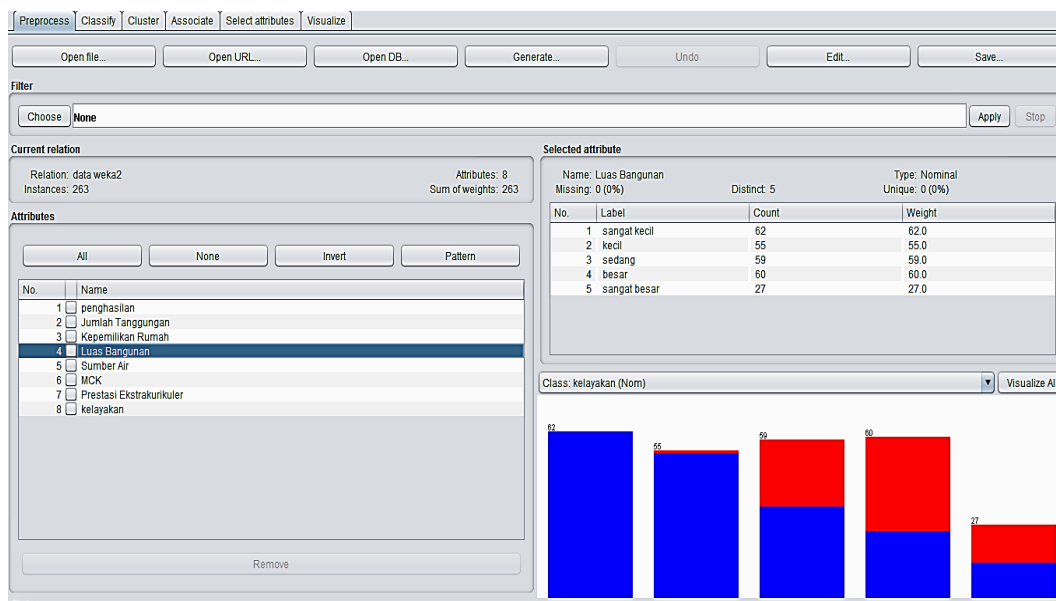


Gambar 5. 6 Visualisasi Atribut Kepemilikan Rumah

Gambar 5.6 adalah visualisasi dari atribut kepemilikan rumah. Sebanyak 263 data pada column *selected atribut* tidak terdapat *missing*. Distinct 4 dan unique 0 dengan jumlah kepemilikan rumah sendiri sebanyak 137 orang, jumlah kepemilikan rumah menumpang sebanyak 80 orang, jumlah kepemilikan rumah sewa bulanan sebanyak 15 orang dan jumlah kepemilikan rumah sewa tahunan sebanyak 31 orang.

4. Visualisasi Atribut Luas Bangunan

Hasil visualisasi atribut tanggungan Jenis Kelamin dapat dilihat pada gambar 5.7 berikut ini :

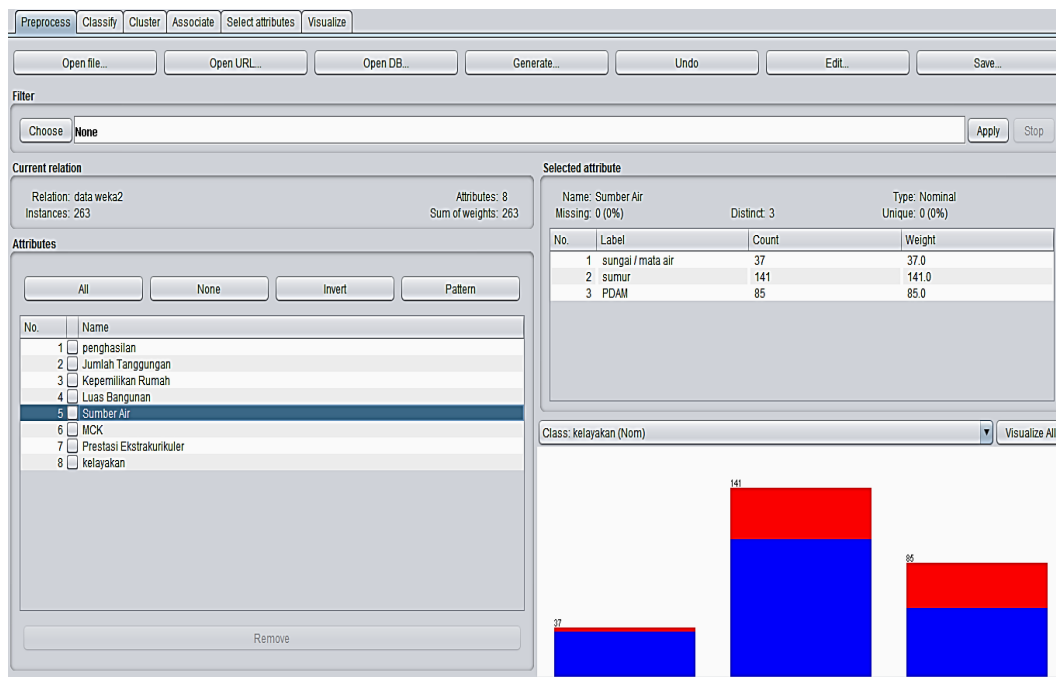


Gambar 5. 7 Visualisasi Atribut Luas Bangunan

Gambar 5.7 adalah visualisasi dari atribut luas bangunan. Sebanyak 263 data pada column *selected atribut* tidak terdapat *missing*. Distinct 5 dan unique 0 dengan jumlah luas bangunan sangat kecil sebanyak 62 orang, jumlah luas bangunan kecil sebanyak 55 orang, jumlah luas bangunan sedang sebanyak 59 orang, jumlah luas bangunan besar sebanyak 60 orang dan jumlah luas bangunan sangat besar sebanyak 27 orang.

5. Visualisasi Atribut Sumber Air

Hasil visualisasi atribut tanggungan Jenis Kelamin dapat dilihat pada gambar 5.8 berikut ini :

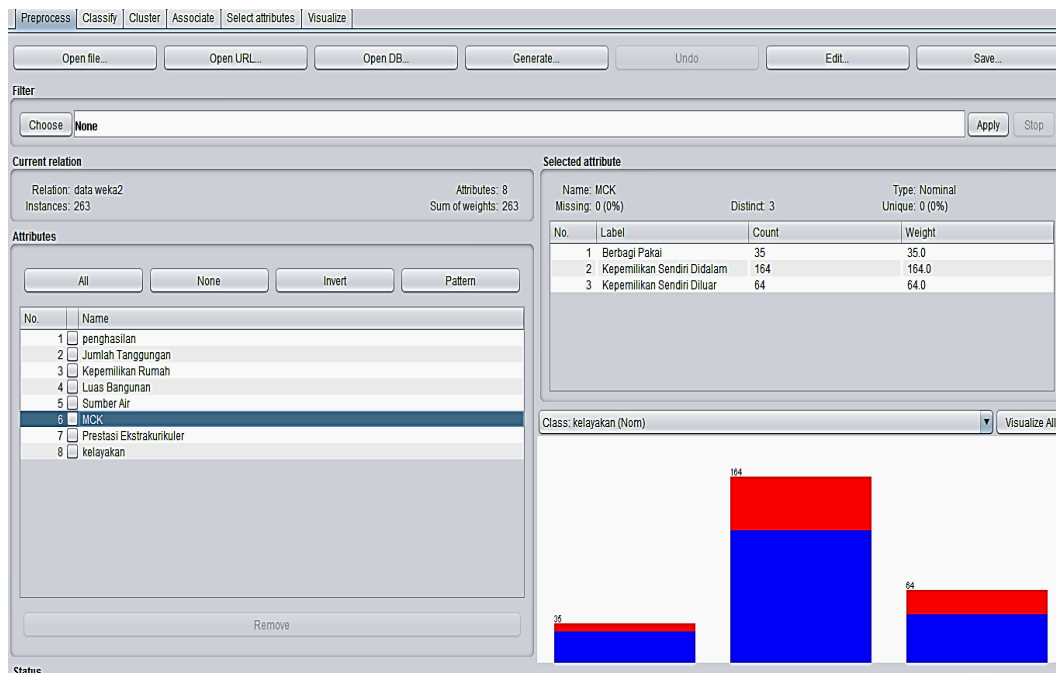


Gambar 5. 8 Visualisasi Atribut Sumber Air

Gambar 5.8 adalah visualisasi dari atribut sumber air. Sebanyak 263 data pada column *selected atribut* tidak terdapat *missing*. Distinct 5 dan unique 0 dengan jumlah sumber air PDAM sebanyak 85 orang, jumlah sumber air sumur sebanyak 141 orang dan jumlah sumber air sungai / matai air sebanyak 37 orang.

6. Visualisasi Atribut MCK

Hasil visualisasi atribut tanggungan Jenis Kelamin dapat dilihat pada gambar 5.9 berikut ini :

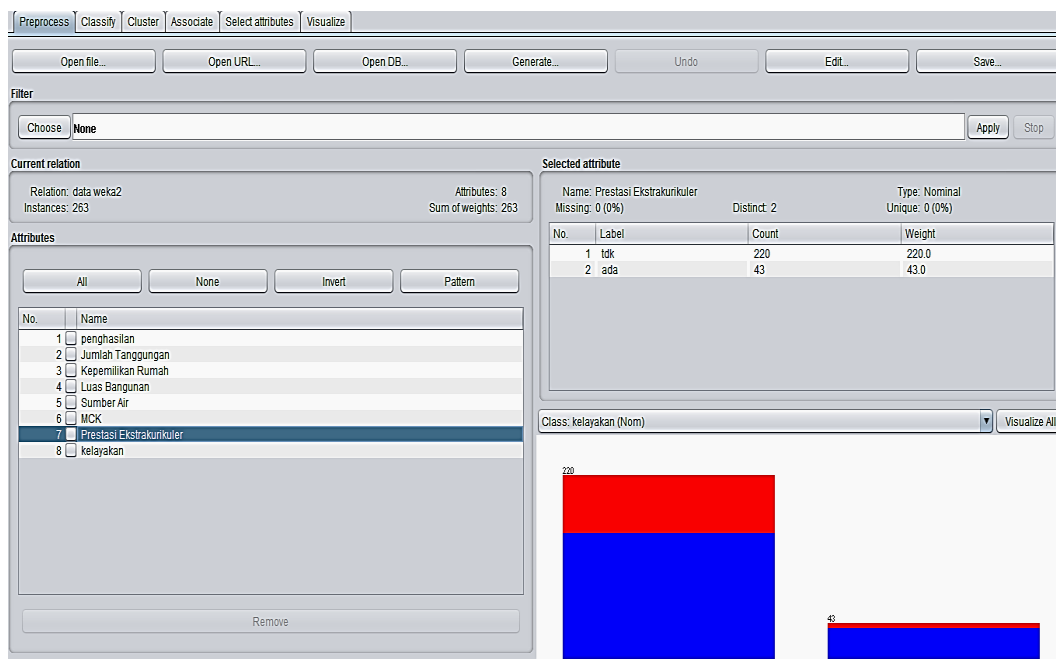


Gambar 5. 9 Visualisasi Atribut MCK

Gambar 5.9 adalah visualisasi dari atribut MCK. Sebanyak 263 data pada column *selected atribut* tidak terdapat *missing*. Distinct 3 dan unique 0 dengan jumlah MCK kepemilikan sendiri di dalam sebanyak 164 orang, jumlah MCK kepemilikan sendiri di luar sebanyak 64 orang dan jumlah MCK berbagi pakai sebanyak 35 orang.

7. Visualisasi Atribut Prestasi Ekstrakurikuler

Hasil visualisasi atribut tanggungan Jenis Kelamin dapat dilihat pada gambar 5.10 berikut ini :



Gambar 5. 10 Visualisasi Atribut Prestasi Ekstrakurikuler

Gambar 5.10 adalah visualisasi dari atribut prestasi ekstrakurikuler. Sebanyak 263 data pada column *selected atribut* tidak terdapat *missing*. Distinct 2 dan unique 0 dengan jumlah prestasi ekstrakurikuler ada sebanyak 43 orang, dan jumlah prestasi ekstrakurikuler tidak sebanyak 220 orang.

1. Hasil Klasifikasi *C4.5* menggunakan *tool weka (use data training)*

Pada proses ini dilakukan pengetesan data menggunakan data training.

Adapun hasilnya dapat dilihat pada gambar 5.11 berikut ini :

The screenshot shows the WEKA Classifier output window. The 'Test options' panel on the left is configured with 'Use training set' selected, 'Cross-validation' set to 10 folds, and 'Percentage split' at 66%. The 'Classifier output' panel on the right displays the following results:

```

=== Evaluation on training set ===

Time taken to test model on training data: 0.03 seconds

=== Summary ===

Correctly Classified Instances      260      98.8593 %
Incorrectly Classified Instances    3         1.1407 %
Kappa statistic                    0.9724
Mean absolute error                 0.0203
Root mean squared error             0.1007
Relative absolute error              4.9655 %
Root relative squared error         22.3023 %
Total Number of Instances          263

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Class
      0.984    0.000    1.000     0.984   0.992     0.973  0.997    0.998    layak
      1.000    0.016    0.962     1.000   0.980     0.973  0.997    0.988    tdk layak
Weighted Avg.   0.989    0.005    0.989     0.989   0.989     0.973  0.997    0.995

=== Confusion Matrix ===

  a  b  <-- classified as
185  3  |  a = layak
  0 75 |  b = tdk layak

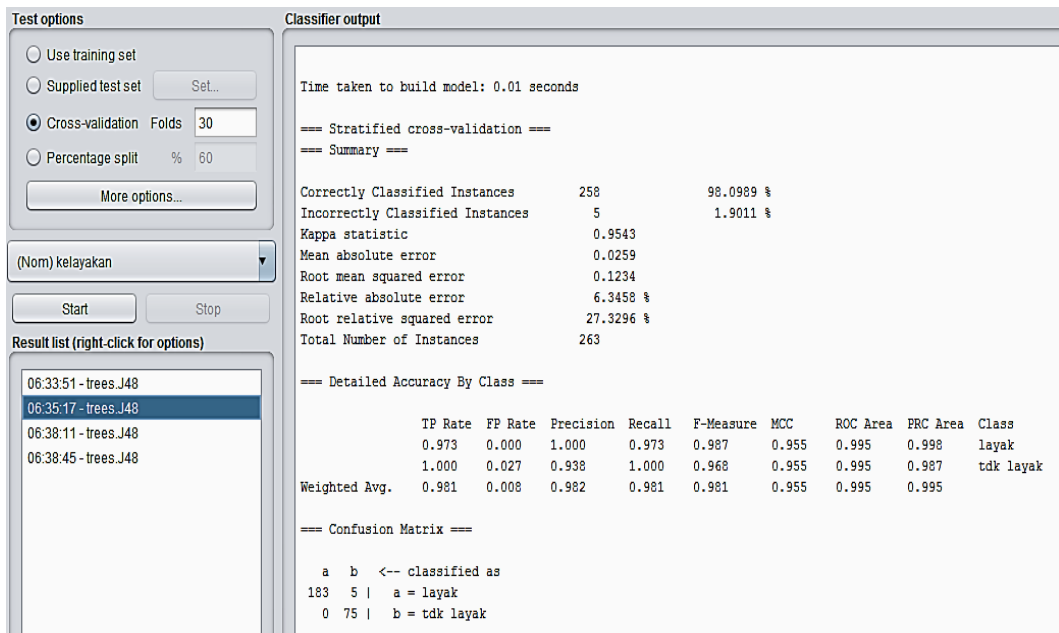
```

Gambar 5. 11 Hasil Klasifikasi Menggunakan Data Training

Gambar 5.11 merupakan hasil klasifikasi *C4.5* pada *tools* WEKA dengan menggunakan *use training set* yang menunjukkan hasil 260 prediksi benar dengan akurasi sebesar 98.8593% dan 3 prediksi salah dengan persentasi 1.1407% dengan waktu klasifikasi selama 0.03 detik.

2. Hasil Klasifikasi *C4.5* menggunakan *tool weka* (10 cross validation)

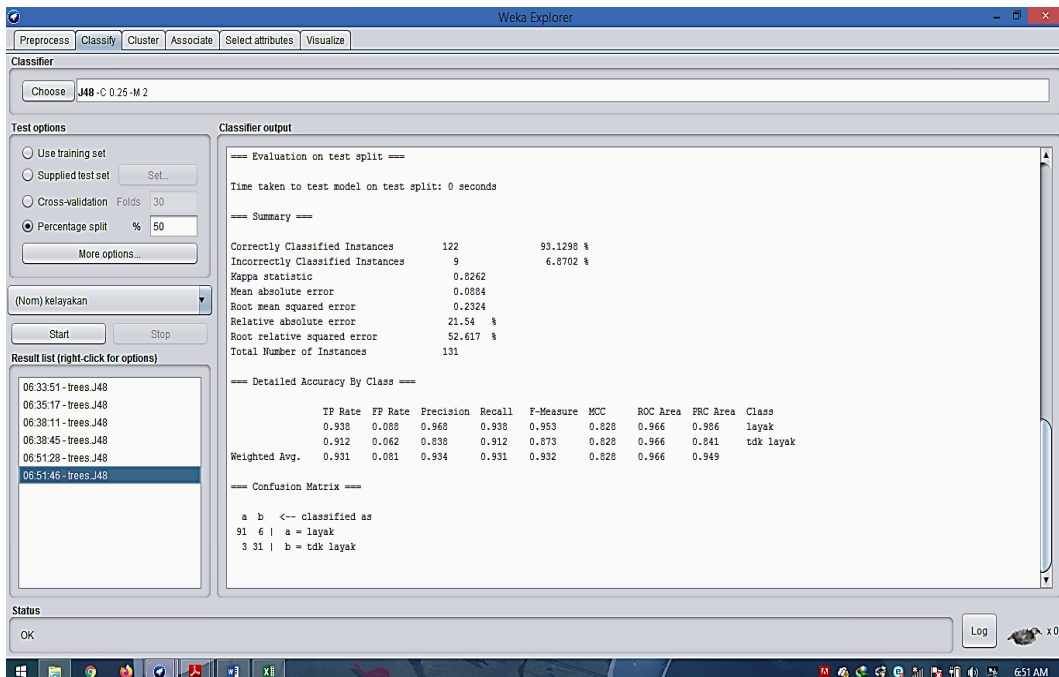
Test 10 *Cross-Validation* melakukan pengetesan data dimana data *training* dibagi menjadi k buah *subset* (subhimpunan). Dimana k adalah nilai dari *fold*. Pada pengetesan ini nilai *fold* adalah 10. Selanjutnya untuk tiap dari *subset*, akan dijadikan data tes dari hasil klasifikasi yang dihasilkan dari k-1 *subset* lainnya. Jadi, akan ada 10 kali tes. Dimana setiap data akan menjadi data tes sebanyak 1 kali dan menjadi data training sebanyak k-1 kali. Dapat dilihat pada gambar 5.12



Gambar 5. 12 Hasil Klasifikasi Dengan 10 Cross Validation

3. Hasil Klasifikasi C4.5 menggunakan tool WEKA (50% Percentage Split)

Tes Percentage Split hasil klasifikasi akan dites dengan menggunakan k% dari data tersebut. Pada tes ini akan digunakan 50% Percentage Split dari data.



Gambar 5. 13 Hasil Klasifikasi Dengan 50% Percentage Split

Gambar 5.12 merupakan hasil klasifikasi *C4.5* pada *tools* WEKA dengan menggunakan *50% Percentage Split* yang menunjukkan hasil 122 prediksi benar dengan akurasi sebesar 93.1298% dan 9 prediksi salah dengan persentasi 6.8702% dengan waktu klasifikasi selama 0 detik.

1.3 PERBANDINGAN HASIL KLASIFIKASI ALGORITMA C4.5

Perbandingan hasil klasifikasi algoritma *C4.5* menggunakan Excel dan Weka dapat dilihat pada tabel 5.1 berikut ini :

Tabel 5. 1 Perbandingan Hasil Klasifikasi Algoritma C4.5

Perbandingan	Excel	Weka
Jumlah kelas “Layak”	188	185
Jumlah kelas “Tidak Layak”	75	78
Akurasi	95.4372%	98.8593%

Pada tabel 5.1 diperoleh perbandingan hasil perhitungan klasifikasi prediksi kelayakan penerima program bidikmisi menggunakan algoritma *C4.5* dengan menggunakan keseluruhan data training menjadi data testing pada tool Excel dan Weka. Hasil akurasi pada tool Excel yaitu sebesar 95.4372%, sedangkan pada tool Weka *Correctly Classified Instance* sebesar 98.8593% dan *Incorrectly Classified Instance* sebesar 1.1407%.