

## BAB V

### HASIL DAN VISUALISASI

#### 5.1 PRAPROSES DATA

Tahapan ini melakukan persiapan proses data. Persiapan proses data tersebut yaitu :

##### 5.1.1 Seleksi Data Mentah

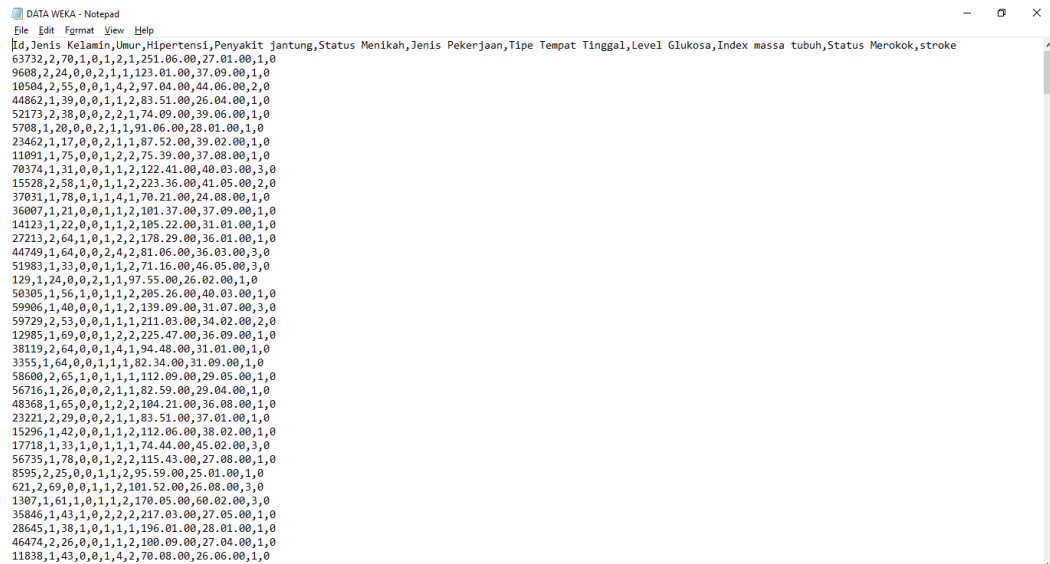
Seleksi data mentah ini merupakan proses penyeleksian data dengan memfokuskan pada atribut-atribut yang digunakan untuk perhitungan yaitu kondisi rumah, status kepemilikan rumah, penghasilan/bln, jumlah keluarga, dan status kesejahteraan. Adapun hasil seleksi data tersebut yaitu :

Id	Jenis Kelamin	Umur	Hipertensi	Penyakit jantung	Status Menikah	Jenis Pekerjaan	Tipe Tempat Tinggal	Level Glukosa	Index massa tubuh	Status Merokok	stroke
63732	2	70	1	0	1	2	1	251.0600	27.0100	1	0
9408	2	24	0	0	2	1	1	123.0100	37.0900	1	0
10504	2	55	0	0	1	4	2	97.0400	44.0600	2	0
44862	1	39	0	0	1	1	2	83.5100	26.0400	1	0
52173	2	38	0	0	2	2	1	74.0900	39.0600	1	0
3708	1	20	0	0	2	1	1	91.0600	28.0100	1	0
23462	1	17	0	0	2	1	1	87.3200	39.0200	1	0
11091	1	75	0	0	1	2	2	75.3900	37.0800	1	0
70374	1	31	0	0	1	1	2	122.4100	40.0300	3	0
13528	2	38	1	0	1	1	2	223.3600	41.0500	2	0
37031	1	78	0	1	1	4	1	70.2100	24.0800	1	0
36007	1	21	0	0	1	1	2	101.3700	37.0900	1	0
14123	1	22	0	0	1	1	2	105.2200	31.0100	1	0
27213	2	64	1	0	1	2	2	178.2900	36.0100	1	0
44749	1	64	0	0	2	4	2	81.0600	36.0300	3	0
51983	1	33	0	0	1	1	2	71.1600	46.0500	3	0
129	1	24	0	0	2	1	1	97.5500	26.0200	1	0
50305	1	56	1	0	1	1	2	205.2600	40.0300	1	0
59906	1	40	0	0	1	1	2	139.0900	31.0700	3	0
59729	2	53	0	0	1	1	1	211.0300	34.0200	2	0
12985	1	69	0	0	1	2	2	225.4700	36.0900	1	0
38119	2	64	0	0	1	4	1	94.4800	31.0100	1	0
3355	1	64	0	0	1	1	1	82.3400	31.0900	1	0
38600	2	65	1	0	1	1	1	112.0900	29.0500	1	0
56715	1	26	0	0	2	1	1	82.5900	29.0400	1	0
48368	1	65	0	0	1	2	2	104.2100	36.0800	1	0
23221	2	29	0	0	2	1	1	83.3100	37.0100	1	0
15296	1	42	0	0	1	1	2	112.0600	38.0200	1	0
17718	1	33	1	0	1	1	1	74.4400	45.0200	3	0
56735	1	78	0	0	1	2	2	115.4300	27.0800	1	0
8595	2	25	0	0	1	1	2	95.5900	25.0100	1	0
621	2	69	0	0	1	1	2	101.5200	26.0800	3	0
1307	1	61	1	0	1	2	2	170.0500	60.0200	3	0
33946	1	43	1	0	2	2	2	217.0300	27.0500	1	0
28645	1	38	1	0	1	1	1	195.0100	28.0100	1	0
46474	2	26	0	0	1	1	2	100.0900	27.0400	1	0
11838	1	43	0	0	1	4	2	70.0800	26.0600	1	0

**Gambar 5.1 Seleksi Data Mentah**

### 5.1.2 Konversi Hasil Data

Hasil dari data tersebut, kemudian dipindahkan ke notepad dengan format .csv. Adapun bentuk data tersebut yaitu :



```

DATA WEKA - Notepad
File Edit Format View Help
Id,Jenis,Kelamin,Umur,Hipertensi,Penyakit jantung,Status Menikah,Jenis Pekerjaan,Tipe Tempat Tinggal,Level Glukosa,Index massa tubuh,Status Merokok,stroke
63752,2,70,1,0,1,2,1,251.06.00,27.01.00,1,0
9608,2,24,0,0,2,1,1,123.01.00,37.09.00,1,0
10504,2,55,0,0,1,4,2,97.04.00,44.06.00,2,0
44862,1,39,0,0,1,1,2,83.51.00,26.04.00,1,0
52173,2,38,0,0,2,2,1,74.09.00,39.06.00,1,0
5708,1,20,0,0,2,1,1,91.06.00,28.01.00,1,0
23462,1,17,0,0,2,1,1,87.52.00,39.02.00,1,0
11091,1,75,0,0,1,2,2,75.39.00,37.08.00,1,0
70374,1,31,0,0,1,1,2,122.41.00,40.03.00,3,0
15528,2,58,1,0,1,1,2,223.36.00,41.05.00,2,0
37031,1,78,0,1,1,4,1,70.21.00,24.08.00,1,0
36007,1,21,0,0,1,1,2,101.37.00,37.09.00,1,0
14123,1,22,0,0,1,1,2,105.22.00,31.01.00,1,0
27213,2,64,1,0,1,2,2,178.29.00,36.01.00,1,0
44749,1,64,0,0,2,4,2,81.06.00,36.03.00,3,0
51983,1,33,0,0,1,1,2,71.16.00,46.05.00,3,0
129,1,24,0,0,2,1,1,97.55.00,26.02.00,1,0
50305,1,56,1,0,1,1,2,205.26.00,40.03.00,1,0
59906,1,40,0,0,1,1,2,139.09.00,31.07.00,3,0
59729,2,53,0,0,1,1,1,211.03.00,34.02.00,2,0
12985,1,69,0,0,1,2,2,225.47.00,36.09.00,1,0
38119,2,64,0,0,1,4,1,94.48.00,31.01.00,1,0
3355,1,64,0,0,1,1,1,82.34.00,31.09.00,1,0
58600,2,65,1,0,1,1,1,112.09.00,29.05.00,1,0
56716,1,26,0,0,2,1,1,82.59.00,29.04.00,1,0
48368,1,65,0,0,1,2,2,104.21.00,36.08.00,1,0
23221,2,29,0,0,2,1,1,83.51.00,37.01.00,1,0
15296,1,42,0,0,1,1,2,112.06.00,38.02.00,1,0
17718,1,33,1,0,1,1,1,74.44.00,45.02.00,3,0
56755,1,78,0,0,1,2,2,115.43.00,27.08.00,1,0
8595,2,25,0,0,1,1,2,95.59.00,25.01.00,1,0
621,2,69,0,0,1,1,2,101.52.00,26.08.00,3,0
1307,1,61,1,0,1,1,2,170.05.00,60.02.00,3,0
35846,1,43,1,0,2,2,2,217.03.00,27.05.00,1,0
28645,1,38,1,0,1,1,1,196.01.00,28.01.00,1,0
46474,2,26,0,0,1,1,2,100.09.00,27.04.00,1,0
11838,1,43,0,0,1,4,2,70.08.00,26.06.00,1,0

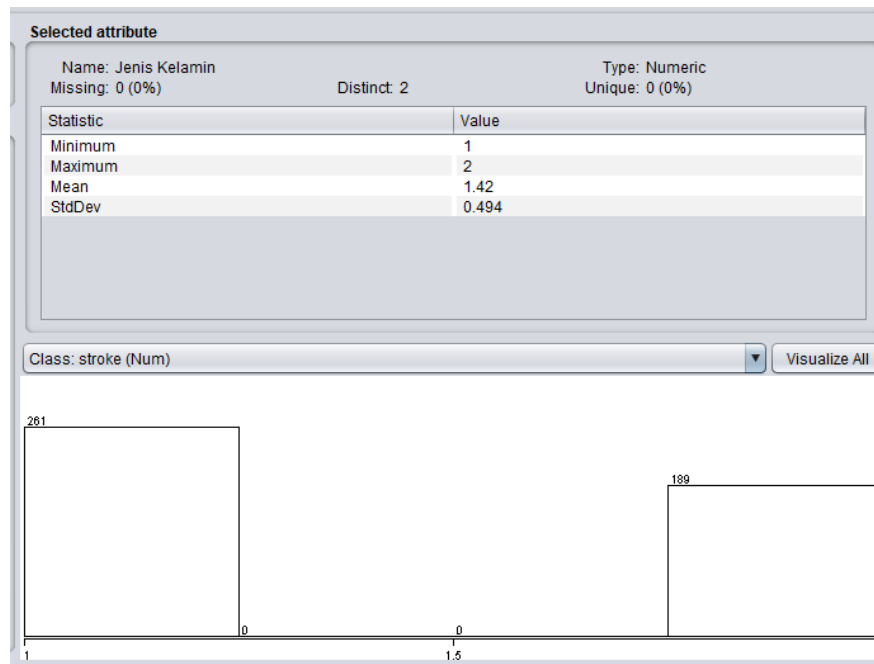
```

Gambar 5.2 Konversi Hasil Data

## 5.2 HASIL VISUALISASI DATA SETIAP ATRIBUT DENGAN MENGGUNAKAN WEKA

Berikut merupakan bentuk visualisasi atribut dengan menggunakan tools WEKA, yaitu :

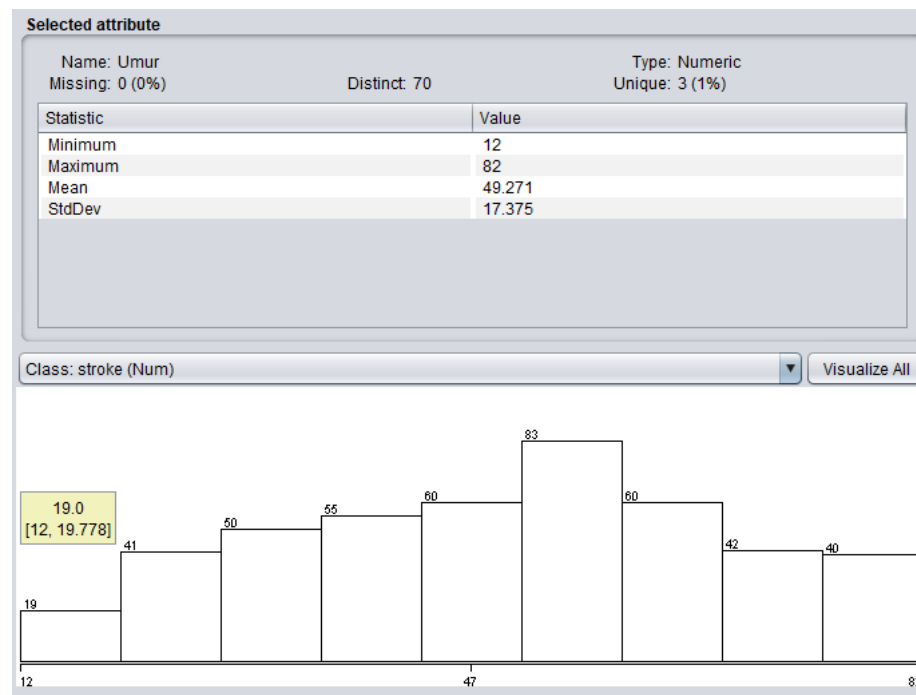
## 1. Visualisasi Atribut Jenis Kelamin



**Gambar 5.3 Visualisasi Atribut Jenis Kelamin**

Gambar 5.3 adalah visualisasi dari atribut Jenis Kelamin. Diketahui bahwa dari 450 data Penderita Penyakit Stroke terdapat hasil 189 penderita stroke yang berjenis kelamin “Female (Perempuan)” dan 261 penderita stroke kelamin “Male (Laki-Laki)”, Nilai rata-rata (mean) yang didapat adalah 1,42 dan standar deviasi adalah 0,494. maka dapat disimpulkan bahwa jumlah penderita yang berjenis laki-laki perempuan lebih banyak daripada jumlah penderita yang berjenis kelamin perempuan.

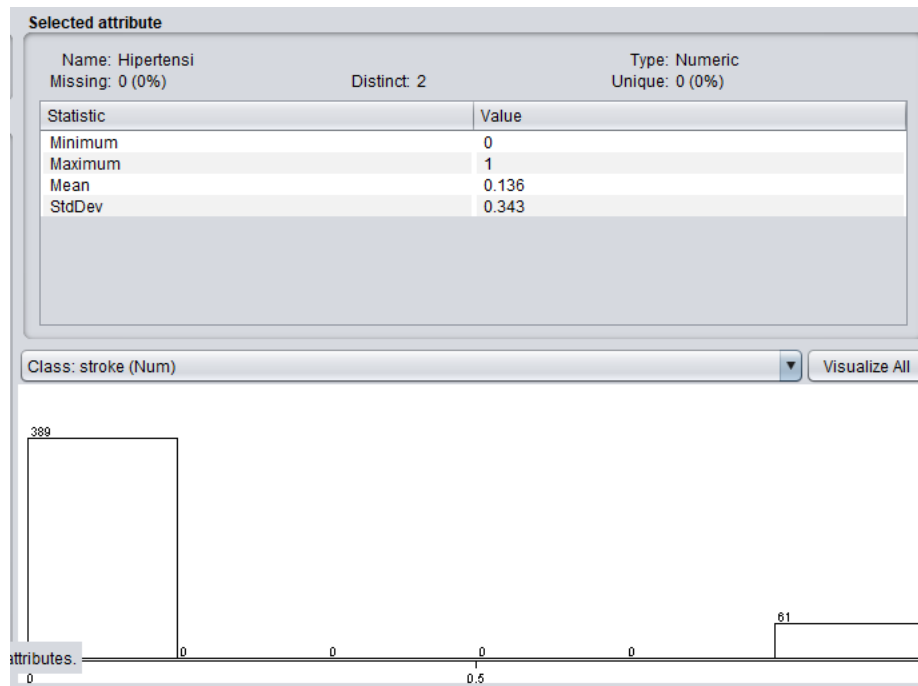
## 2. Visualisasi Atribut Umur



**Gambar 5.4 Visualisasi Atribut Jenis Kelamin**

Gambar 5.4 adalah visualisasi dari atribut Umur. Diketahui bahwa dari 450 data Penderita Penyakit Stroke terdapat hasil umur terbanyak adalah umur “83” pada data penderita stroke, Nilai rata-rata (mean) yang didapat adalah 49.271 dan standar deviasi adalah 17.375.

### 3. Visualisasi Atribut Hipertensi

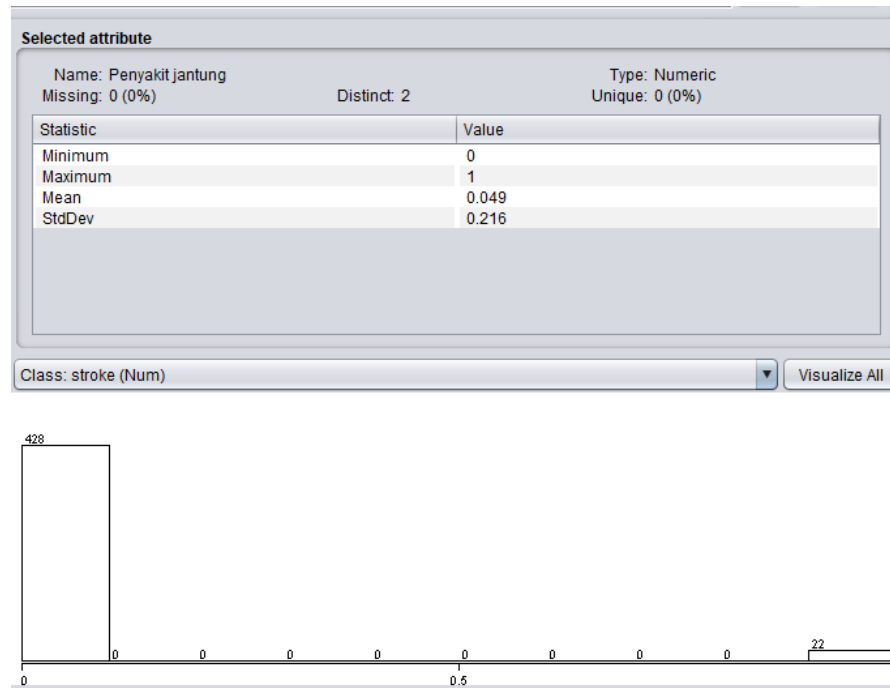


**Gambar 5.5 Visualisasi Atribut Hipertensi**

Gambar 5.5 adalah visualisasi dari atribut Hipertensi. Diketahui bahwa dari 450 data Penderita Penyakit Stroke Terdapat 389 penderita dengan Status Tingkat “0” dan 61 penderita stroke pada tingkat “1”, Nilai rata-rata (mean) yang didapat adalah 0.136 dan standar deviasi adalah 0.343.

Maka dapat disimpulkan bahwa jumlah penderita dengan Status Tingkat “0” lebih banyak daripada jumlah penderita dengan Status Tingkat lainnya.

#### 4. Visualisasi Atribut Penyakit Jantung

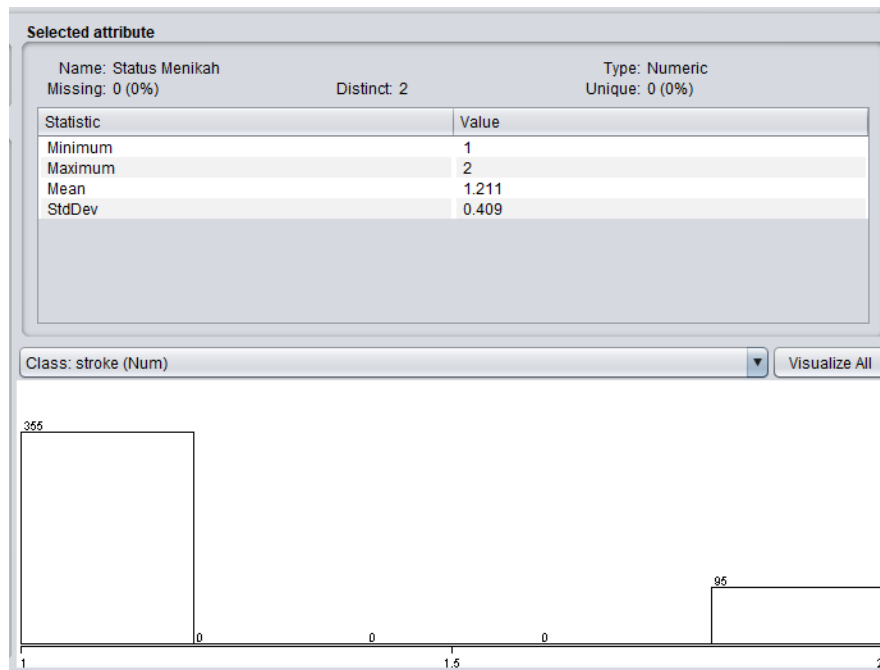


**Gambar 5.6 Visualisasi Atribut Penyakit Jantung**

Gambar 5.6 adalah visualisasi dari atribut Penyakit Jantung. Diketahui bahwa dari 450 data Penderita Penyakit Stroke Terdapat 428 penderita dengan Status Tingkat “0” dan 22 penderita stroke pada tingkat “1”, Nilai rata-rata (mean) yang didapat adalah 0.049 dan standar deviasi adalah 0.216.

Maka dapat disimpulkan bahwa jumlah penderita dengan Status Tingkat “0” lebih banyak daripada jumlah penderita dengan Status Tingkat lainnya.

## 5. Visualisasi Atribut Status Menikah



**Gambar 5.7 Visualisasi Atribut Status Menikah**

Gambar 5.7 adalah visualisasi dari atribut Status Menikah. Diketahui bahwa dari 450 data Penderita Penyakit Stroke terdapat hasil 355 penderita stroke telah menikah dan 95 penderita stroke belum menikah. Nilai rata-rata (mean) yang didapat adalah 1.211 dan standar deviasi adalah 0.409.

maka dapat disimpulkan bahwa jumlah status menikah pada penderita banyak yang telah memiliki menikah dari pada status belum menikah.

## 6. Visualisasi Atribut Jenis Pekerjaan



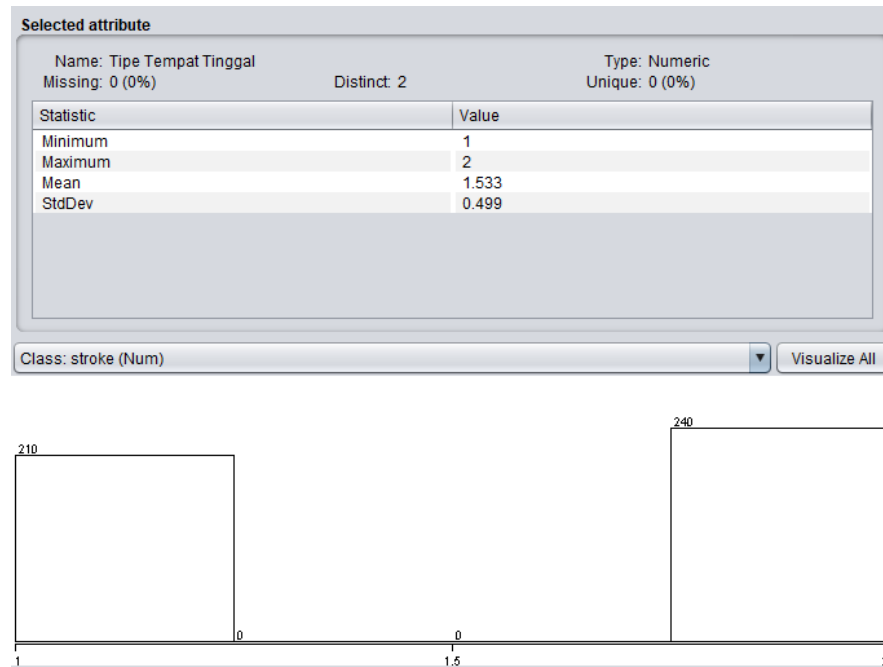
**Gambar 5.8 Visualisasi Atribut Jenis Pekerjaan**

Gambar 5.8 adalah visualisasi dari atribut Status Menikah. Diketahui bahwa dari 450 data Penderita Penyakit Stroke terdapat hasil 355 penderita stroke telah menikah dan 95 penderita stroke belum menikah. Nilai rata-rata (mean) yang didapat adalah 1.573 dan standar deviasi adalah 1.03.

maka dapat disimpulkan bahwa jumlah jenis pekerjaan pada penderita banyak yang telah memiliki menikah dari pada status belum menikah.



## 7. Visualisasi Atribut Tipe Tempat Tinggal

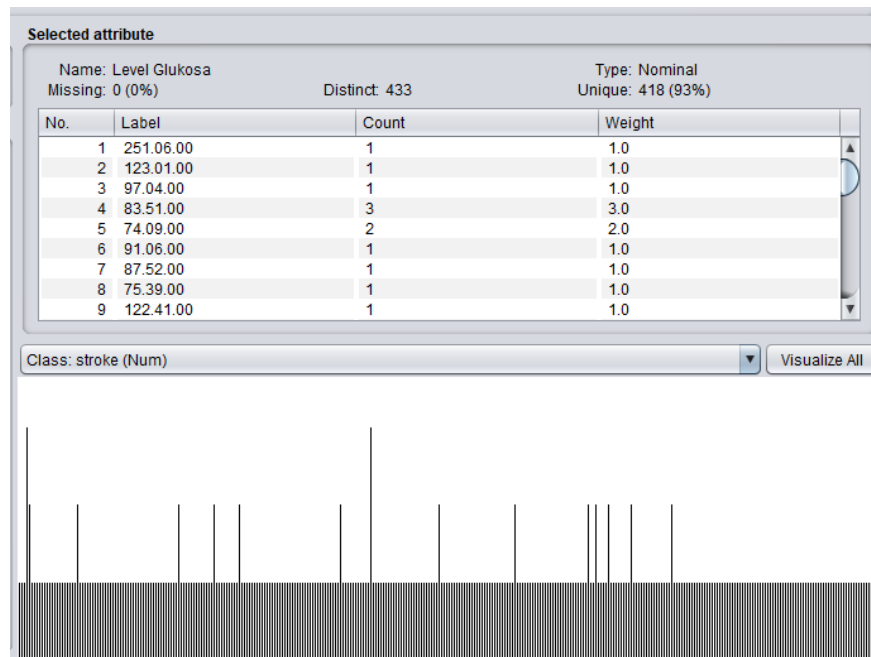


**Gambar 5.9 Visualisasi Atribut Tipe Tempat Tinggal**

Gambar 5.8 adalah visualisasi dari atribut Penyakit Jantung. Diketahui bahwa dari 450 data Penderita Penyakit Stroke Terdapat 210 penderita dengan Status Tingkat “1” dan 240 penderita stroke pada tingkat “2”, Nilai rata-rata (mean) yang didapat adalah 1.533 dan standar deviasi adalah 0.499.

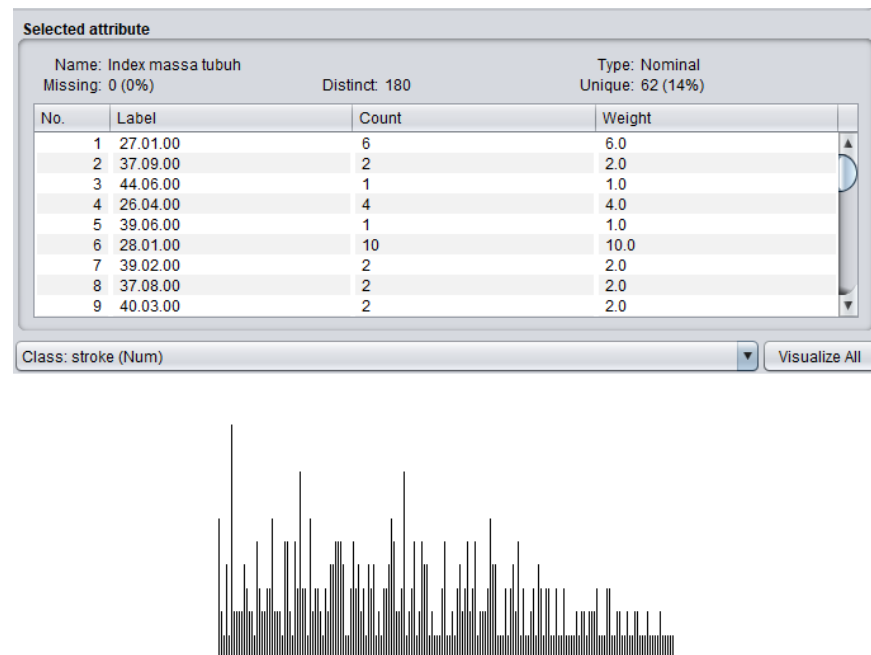
Maka dapat disimpulkan bahwa jumlah penderita dengan Status Tingkat “2” lebih banyak daripada jumlah penderita dengan Status Tingkat lainnya.

## 8. Visualisasi Atribut Level Glukosa



Gambar 5.10 Visualisasi Atribut Level Glukosa

## 9. Visualisasi Atribut Index Massa Tubuh



Gambar 5.11 Visualisasi Atribut Index Massa Tubuh

## 10. Visualisasi Atribut Status Merokok



**Gambar 5.12 Visualisasi Atribut Status Merokok**

Gambar 5.11 adalah visualisasi dari atribut Status Merokok. Diketahui bahwa dari 450 data Penderita Penyakit Stroke Terdapat 240 penderita dengan Status Tingkat “1” dan 104 penderita stroke pada tingkat “2”, dan 97 penderita stroke pada tingkat “3”, Nilai rata-rata (mean) yang didapat adalah 1.662 dan standar deviasi adalah 0.81.

Maka dapat disimpulkan bahwa jumlah penderita dengan Status Tingkat “1” lebih banyak daripada jumlah penderita dengan Status Tingkat lainnya.

## 11. Pengambilan Pusat *Cluster*

```

Cluster output
-----
Jenis Kelamin
Umur
Hipertensi
Penyakit jantung
Status Menikah
Jenis Pekerjaan
Tipe Tempat Tinggal
Level Glukosa
Index massa tubuh
Status Merokok
stroke
Test mode:  evaluate on training data

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

kMeans
=====

Number of iterations: 10
Within cluster sum of squared errors: 1246.000926373948

Initial starting points (random):

Cluster 0: 1,39,0,0,1,4,2,191.47.00,28.03.00,1,0
Cluster 1: 2,25,0,0,2,1,2,113.08.00,35.03.00,2,0
Cluster 2: 1,70,0,0,2,2,1,98.42.00,36.04.00,2,0

```

**Gambar 5.12 Pengambilan *Cluster***

Pada weka pengambilan pusat *cluster* diambil secara random dan disebutkan nilainya dari 11 atribut, pada weka dijelaskan jumlah iterasi yang di lakukan untuk mendapatkan sebanyak 3 *cluster* yaitu sebanyak 3 iterasi. Kemudian setiap atribut mendapatkan nilai masing-masing dapat dilihat pada gambar 5.10 dibawah ini :

Clusterer output

Attribute	Full Data (450.0)	Cluster#		
		0 (153.0)	1 (91.0)	2 (206.0)
Jenis Kelamin	1.42	1.0131	1.989	1.4709
Umur	49.2711	48.7712	48.8242	49.8398
Hipertensi	0.1356	0.1046	0.1758	0.1408
Penyakit jantung	0.0489	0.0327	0.0879	0.0437
Status Menikah	1.2111	1.2157	1.1978	1.2136
Jenis Pekerjaan	1.5733	1.6601	1.4945	1.5437
Tipe Tempat Tinggal	1.5333	1.9869	1.978	1
Level Glukosa	83.51.00	77.04.00	83.51.00	112.09.00
Index massa tubuh	28.01.00	26.06.00	27.04.00	28.01.00
Status Merokok	1.6622	1.6536	1.6813	1.6602
stroke	0.0244	0.0261	0.011	0.0291

Time taken to build model (full training data) : 0.07 seconds

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

Clustered Instances

0	153 ( 34%)
1	91 ( 20%)
2	206 ( 46%)

**Gambar 5.13 Centroid Cluster**

Kemudian didapatkan jumlah dalam setiap *clusternya*

*Cluster 0* sebanyak 153 data atau sebesar 34%

*Cluster 1* sebanyak 91 data atau sebesar 20%

*Cluster 2* sebanyak 206 data atau sebesar 46%

Dapat disimpulkan hasil *cluster* dengan tools *WEKA* adalah sebanyak 450 data penderita penyakit stroke dibagi menjadi 3 *cluster* dimana dari masing-masing *cluster* nantinya akan dikelompokkan menjadi 3 kelompok, yaitu pada *cluster 0* terdapat 153 penderita, *cluster 1* sebanyak 91 penderita, *cluster 2* sebanyak 206 penderita. Kelemahan dari *WEKA* ini yaitu tidak dijelaskan penderita yang mana berada pada setiap *clusternya*, *WEKA* hanya menjelaskan jumlah data keseluruhan dan kelompoknya saja.

