**MAKALAH**

**PEMANFAATAN PROFIL PELANGGAN RESTORAN UNTUK PENGEMBANGAN PELAYANAN RESTORAN**

**Dosen pengampu: FHIRA NHITA, S.T., M.T.**

**Disusun Oleh :**

**- Aditiyan Iswahyudin | 1301154167**

**- R. Hatim Muhammad Amin | 1301154223**

**- Dhuki Dwi R | 1301154265**

**- Odia Pratama | 1301154405**

**- Fero Resyanto | 1301154318**

****

**FAKULTAS TEKNIK INFORMATIKA**

**TELKOM UNIVERSITY**

**2018**

Daftar Isi

[KATA PENGANTAR 3](#_Toc524520929)

[BAB I PENDAHULUAN 4](#_Toc524520930)

[Latar Belakang 4](#_Toc524520931)

[Rumusan Masalah 4](#_Toc524520932)

[Tujuan 4](#_Toc524520933)

[BAB II PEMBAHASAN 5](#_Toc524520934)

[Dataset 5](#_Toc524520935)

[Preprocessing 5](#_Toc524520936)

[Tahapan Preprocessing 5](#_Toc524520937)

[Kasus: Dataset Profil Pelanggan Sebuah Restoran 6](#_Toc524520938)

[Sumber Data 6](#_Toc524520939)

[Deskripsi Data 6](#_Toc524520940)

[Masalah Data 6](#_Toc524520941)

[Solusi Masalah Data 7](#_Toc524520942)

[Penggunaan WEKA 7](#_Toc524520943)

[REFERENSI 12](#_Toc524520944)

# KATA PENGANTAR

Dengan menyebut nama Allah SWT yang Maha Pengasih lagi Maha Penyayang, kami panjatkan puji syukur atas kehadirat-Nya, yang telah melimpahkan rahmat-Nya kepada kami, sehingga kami dapat menyelesaikan makalah tentang *data mining* yang kami beri judul "Pemanfaatan Profil Pelanggan Restoran Untuk Pengembangan Pelayanan Restoran".

Adapun makalah ini telah kami usahakan semaksimal mungkin dan tentunya dengan bantuan dari banyak pihak, sehingga dapat memperlancar proses pembuatan makalah ini. Oleh sebab itu, kami juga ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu kami dalam pembuatan makalah ini.

Akhirnya penulis berharap semoga makalah ilmiah *data mining* tentang “Pemanfaatan Profil Pelanggan Restoran Untuk Pengembangan Pelayanan Restoran" ini dapat diambil manfaatnya sehingga dapat memberikan manfaat terhadap pembaca. Selain itu, kritik dan saran dari pembaca sangat kami harapkan guna perbaikan makalah di masa yang akan dating.

# BAB I PENDAHULUAN

## Latar Belakang

Dijaman era digital ini, hampir semua orang melibatkan teknologi kedalam pekerjaan mereka, baik dari sisi pelayanan ataupun dari sisi pengolahan data. Hal ini bukan hanya dirasakan di tempat-tempat yang berfokus di bagian teknologi saja, melainkan berkembang juga di bagian-bagian yang lain seperti restoran.

Pada saat ini kita dapat menjumpai restoran dimana-mana. Persaingan antara tempat makan ini sangatlah ketat, bahkan kita bisa melihat jarak antar restoran tersebut saling berdekatan. Hal ini juga disebabkan karena adanya keterbatasan tempat dan tingkat keramaian tempat tersebut. Oleh hal itu, restoran harus saling membuat strategi satu dengan yang lainnya agar bisa terus bersaing dan tidak mengalami kebangkrutan.

Melibatkan teknologi pada restoran adalah suatu hal yang sangat penting disini, karena dengan adanya teknologi kita bisa mengolah data yang ada. Restoran bisa saja mendata karakter pembeli yang datang dan mencari tau karakteristik secara umum agar dapat menentukan strategi yang jauh lebih baik kedepannya. Data yang dibicarakan seperti makanan yang paling banyak terjual, harga yang paling banyak dipilih oleh pembeli, gaya pembeli yang datang dan masih banyak hal lainnya. Dengan restoran mendata semua hal ini maka restoran dapat dengan mudah untuk menentukan tema restoran mereka, bahan makanan yang harus diperbanyak atau mungkin penentuan harga yang harus diteliti lebih lanjut. Dengan adanya teknologi ternyata terbukti mampu membantu bukan hanya dibagian yang berfokus dengan teknologi saja tapi dapat membantu pihak-pihak seperti restoran ini.

## Rumusan Masalah

1. Apa itu dataset?
2. Apa itu *missing values* dalam dataset?
3. Apa itu *preprocessing*?
4. Bagaimana cara mengidentifikasi sebuah dataset?
5. Bagaimana cara memperbaiki *missing values*?
6. Bagaimana menggunakan *tools data mining*?

## Tujuan

1. Mengetahui dan memahami dataset.
2. Memahami masalah yang dapat timbul dalam sebuah dataset.
3. Dapat memberikan solusi terhadap masalah yang ada pada dataset.
4. Memahami dan mengaplikasikan tahapan *preprocessing* dalam mengolah sebuah dataset.
5. Eksplorasi *software* dan *tools* yang digunakan untuk *data mining*.

# BAB II PEMBAHASAN

## Dataset

Data dalam data mining merupakan kumpulan objek data dan atributnya. Dataset merupakan sekumpulan banyak data dalam table database. Data-data ini memiliki property atau karakteristik yang dinamakan atribut. Pada dataset mungkin timbul sebuah anomaly data, contohnya adalah *missing values*. *Missing values* adalah sebuah keadaan dataset dimana nilai data kosong atau tidak ada. Beberapa sebab terjadinya *missing values* antara lain, informasi tidak terkumpul secara lengkap atau atribut tidak bisa diterapkan untuk semua kasus. Salah satu solusi untuk penanganan *missing values* adalah dengan estimasi nilai *missing values* menggunakan mean (untuk data berbentuk angka) atau modus (untuk data berbentuk label/symbol) dari data tersebut.

## Preprocessing

Strategi dan teknik yang saling berkaitan untuk membuat data lebih mudah/cocok untuk digunakan pada *data mining*. Tujuannya adalah untuk meningkatkan hasil analisis data mining terkait masalah waktu, cost dan kualitas.

### Tahapan Preprocessing

1. Perhatikan mana yang atribut rasio, mana yang internval, ordinal dan nominal.
2. Perhatikan, apakah ada data/instance/sample yang ganda ? Kalau ada hapuslah dari training set.
3. Apakah ada missing data ? Kalau missing data-nya relatif sedikit, hapus saja dari training set. Tetapi kalau banyak nilai *missing values-*nya lebih baik di estimasi menggunakan mean.
4. Apakah ada data *outlier* ? Untuk kehati-hatian sebaiknya dihapus saja *outlier*-nya.
5. Perhatikan range tiap atribut. Konversikan range atribute ke interval yang sama. Biasanya dikonversikan ke [0,1] artinya dinormalisasi ke minimal 0 dan maksimal 1.

## Kasus: Dataset Profil Pelanggan Sebuah Restoran

### Sumber Data

[https://archive.ics.uci.edu/ml/datasets/Restaurant+%26+consumer+data#](https://archive.ics.uci.edu/ml/datasets/Restaurant+%26+consumer+data)

### Deskripsi Data

* Instances: 138
* Attributes: 19
* userID: Nominal
* latitude: Ratio (Kontinu,)
* longitude: Ratio (Kontinu)
* the\_geom\_meter: Nominal (Geospatial)
* smoker: Nominal, Missing: 3, 2 [false,true]
* drink\_level: Nominal, 3 [abstemious,social drinker,casual drinker]
* dress\_preference:Nominal, Missing: 5, 4 [informal,formal,no preference,elegant]
* ambience: Nominal, Missing: 6, 3 [family,friends,solitary]
* transport: Nominal, Missing: 7, 3 [on foot,public,car owner]
* marital\_status: Nominal, Missing: 4, 3 [single,married,widow]
* hijos: Nominal, Missing: 11, 3 [independent,kids,dependent]
* birth\_year: Nominal
* interest: Nominal, 5 [variety,technology,none,retro,eco-friendly]
* personality: Nominal, 4 [thrifty-protector,hunter-ostentatious,hard-worker,conformist]
* religion: Nominal, 5 [none,Catholic,Christian,Mormon,Jewish]
* activity: Nominal, Missing: 7, 4 [student,professional,unemployed,working-class]
* color: Nominal, 8 [black,red,blue,green,purple,orange,yellow,white]
* weight: Numeric
* budget: Nominal, Missing: 7, 3 [medium,low,high]
* height: Numeric

### Masalah Data

Beberapa atribut diatas memiliki *missing values*, seperti di atribut smoker, dress\_preference, ambiance, transport, marital\_status, hijos, activity, dan budget. Pada dataset diatas juga ditemukan adanya *outlier*, walaupun kecil (hanya 2 dari 300an data). Lalu untuk mengoptimalkan data sehingga lebih mudah diintepretasi maka metode preprocessing yang digunakan adalah diskretisasi. Tujuan dari diskritisasi ini adalah untuk membagi nilai data kedalam beberapa range data sehingga bisa di kategorisasi dalam bentuk label.

### Solusi Masalah Data

Solusi dalam penanganan *missing values* adalah dengan estimasi nilai menggunakan mean dari dataset yang ada sehingga didapat sebuah nilai yang akan digunakan pada atribut yang memiliki *missing values.* Metode ini bisa digunakan langsung pada software WEKA dengan menggunakan *filter* “ReplaceMissingValues”. Untuk menghilangkan *outlier* bisa menggunakan *filter* “InterquartieRange” untuk menampilkan jumlah *outlier* dan dilanjutkan dengan *filter ”*RemoveWithValues*”* untuk menghilangkan data *outlier*.

## Penggunaan WEKA

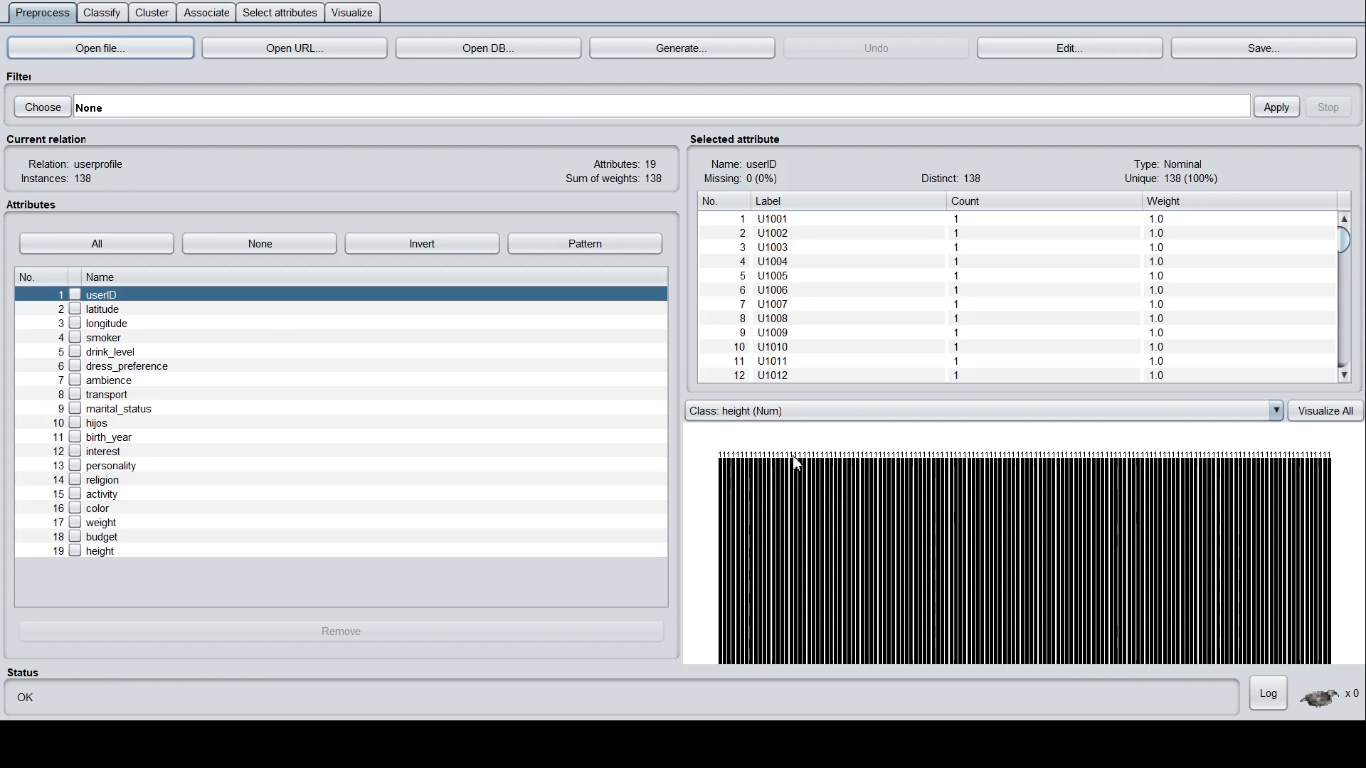


Figure 1 Tampilan awal WEKA

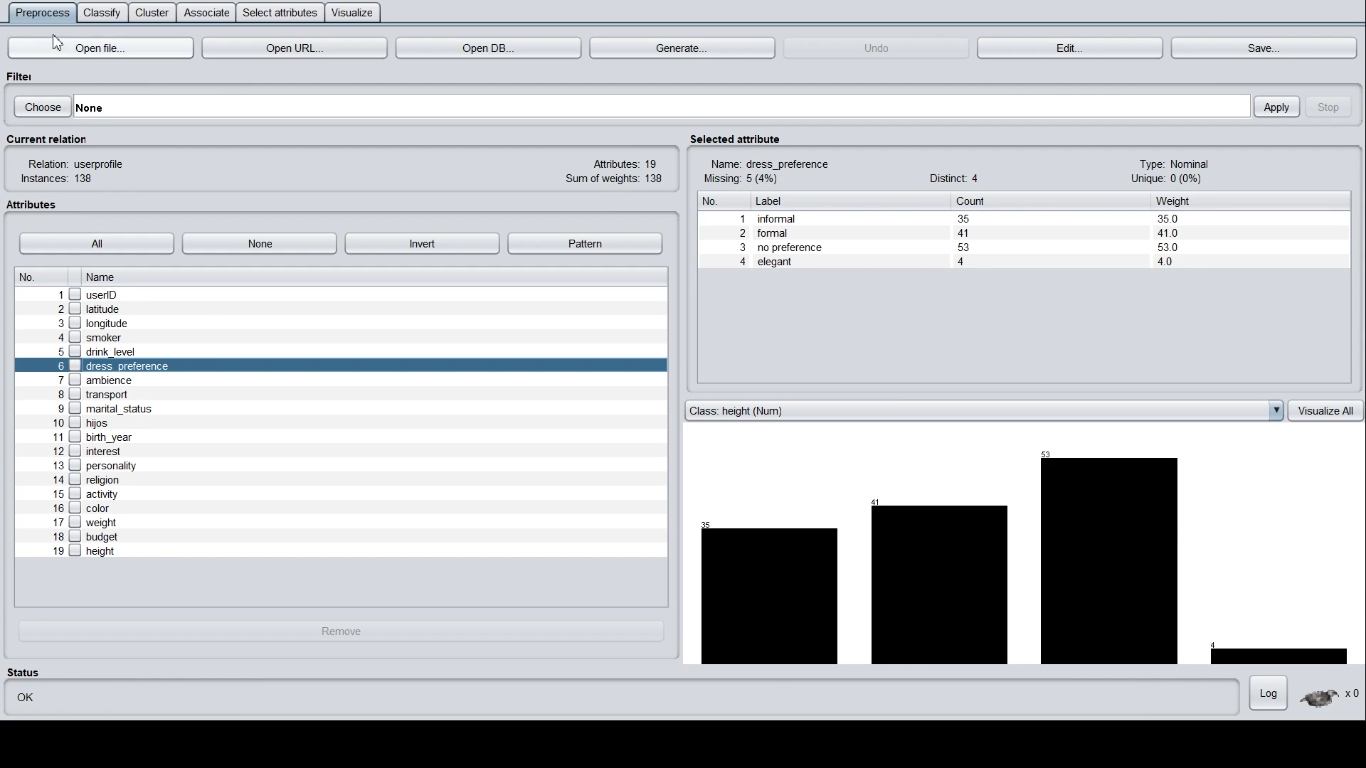


Figure 2 Atribut yang missing values

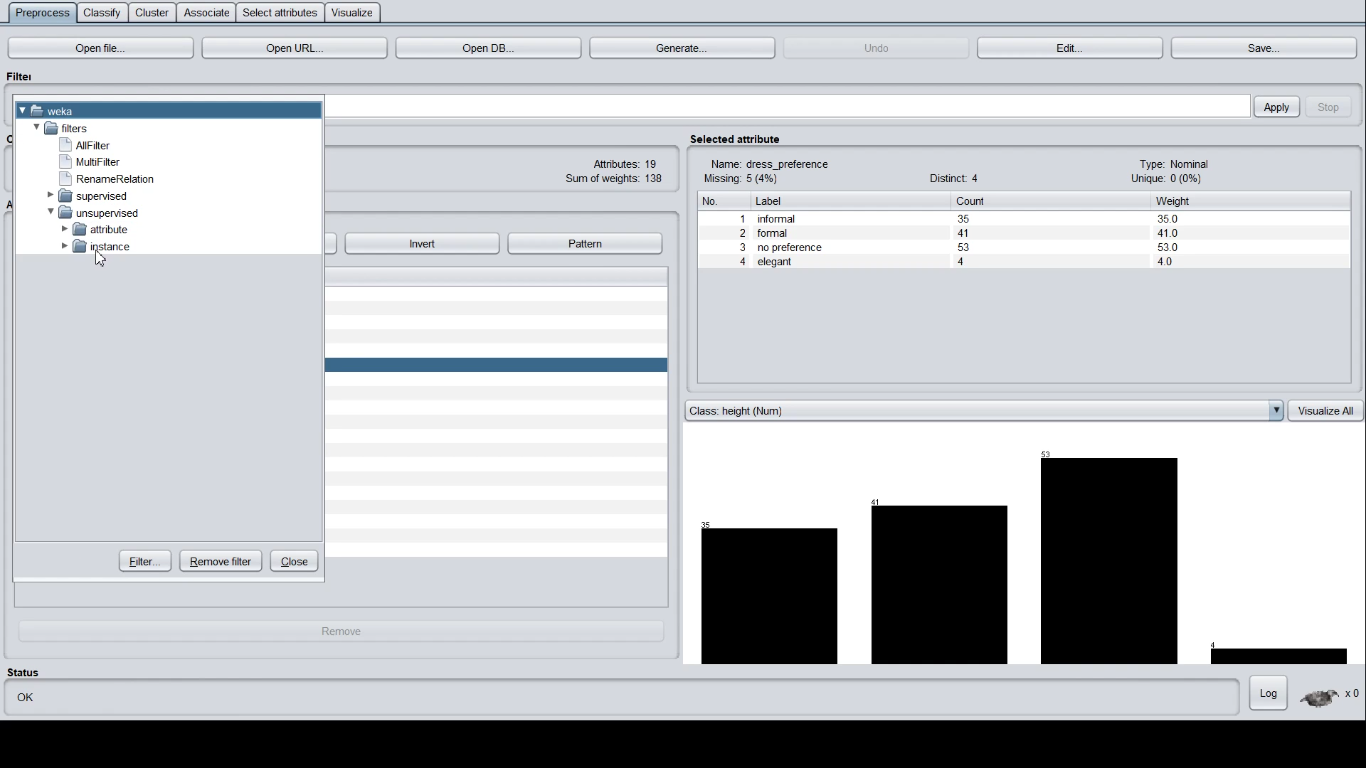


Figure 3 Tahapan untuk memperbaiki missing values

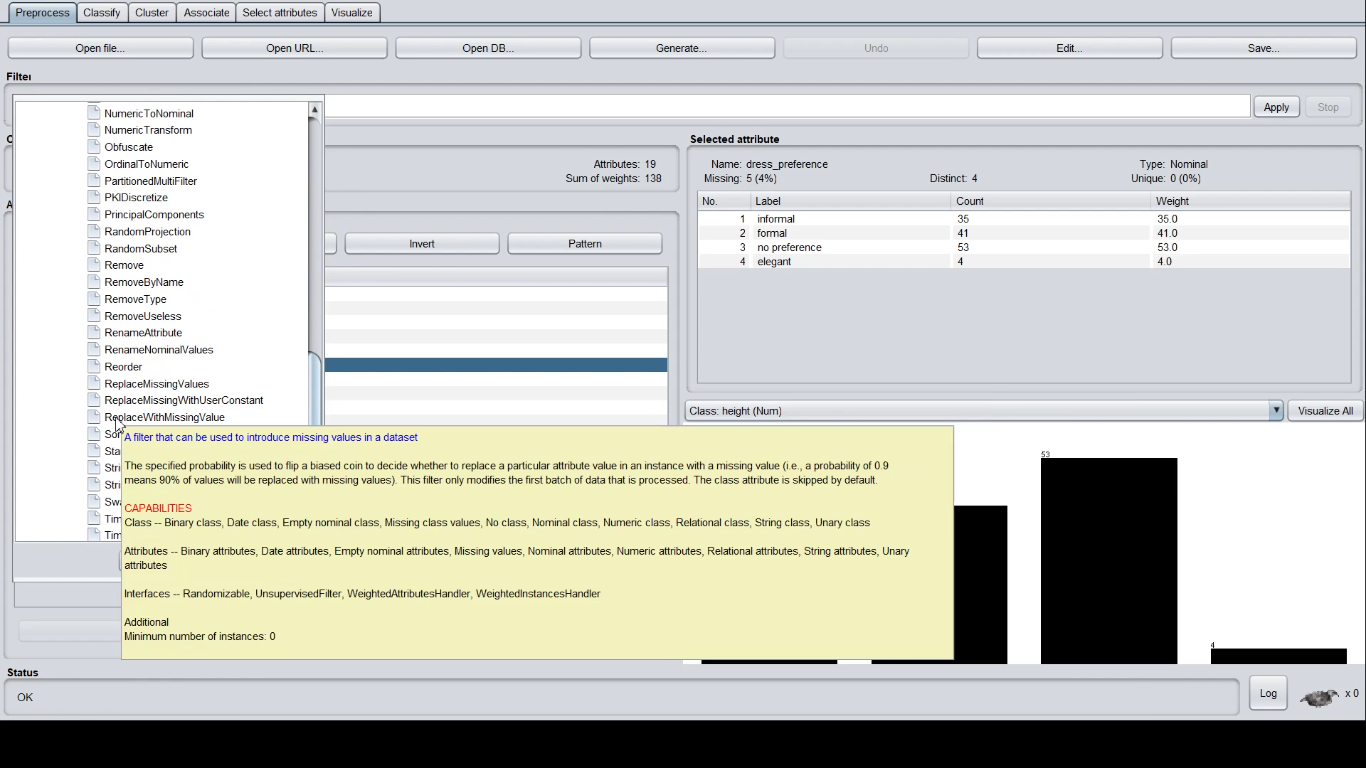


Figure 4 Menggunakan tools ReplaceMissingValues

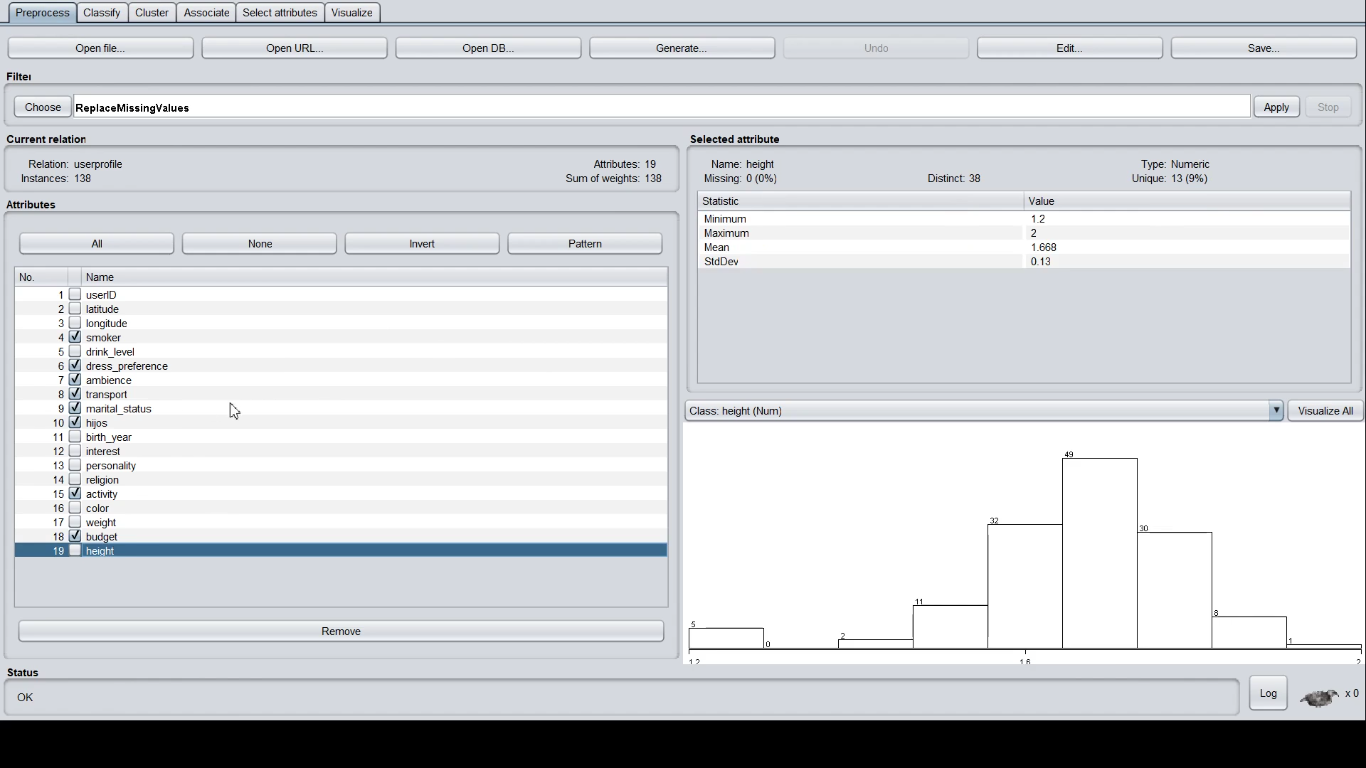


Figure 5 Centang semua data yang memiliki missing values

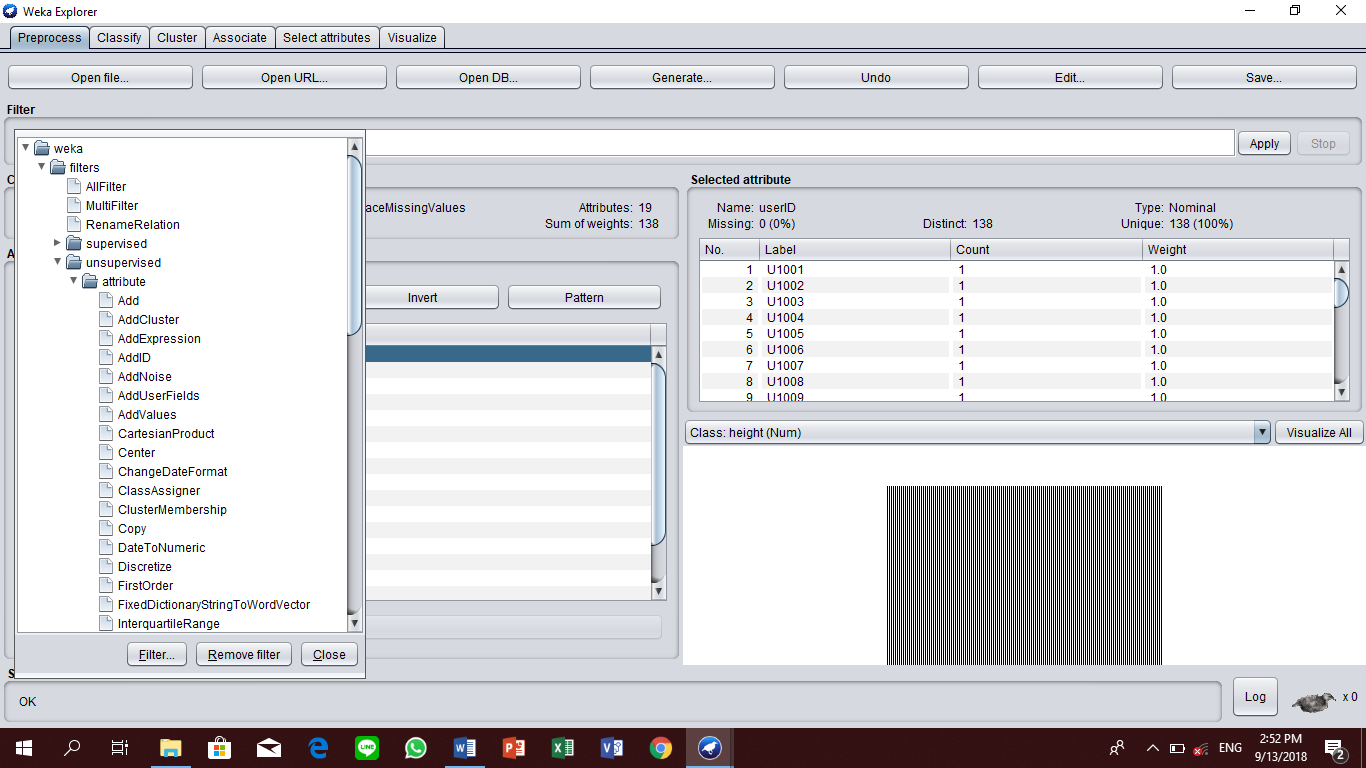


Figure 6 Penggunanaan filter InterquartileRange

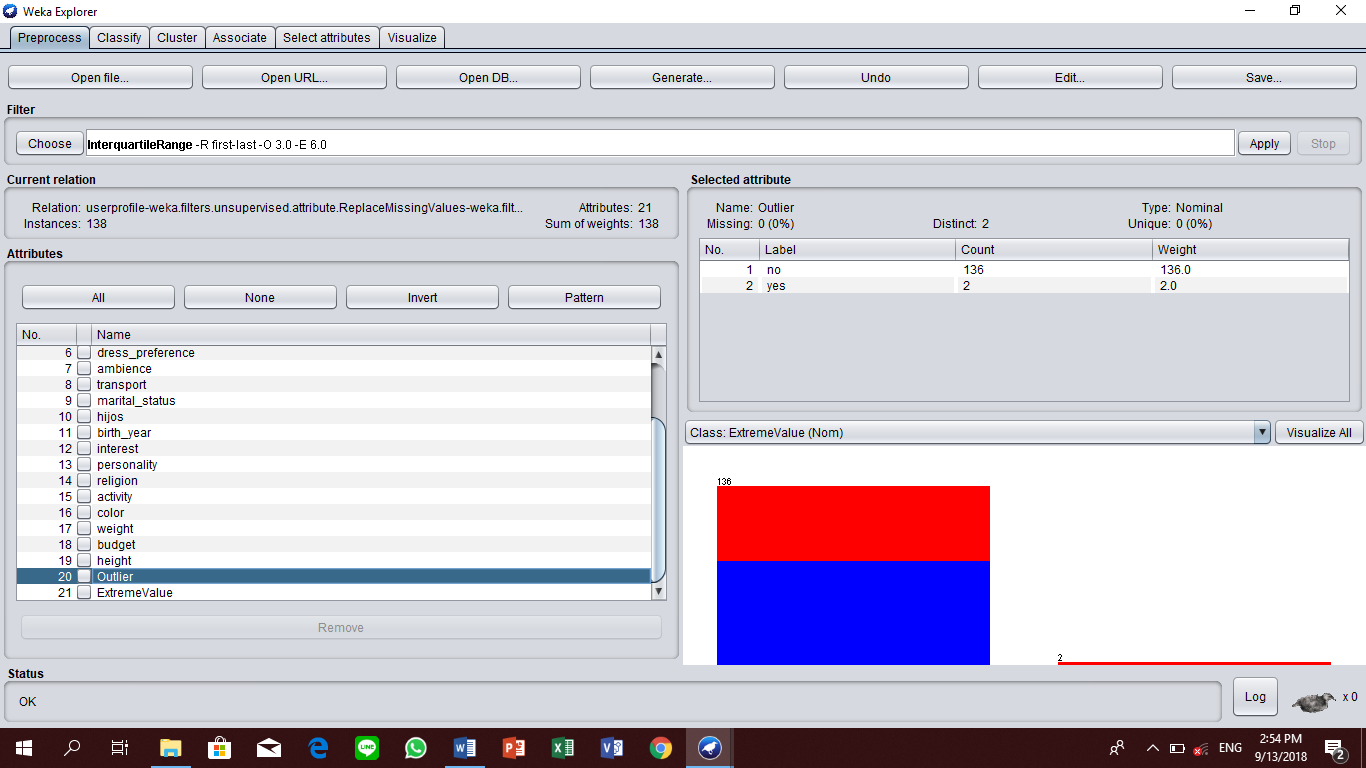


Figure 7 Tampilan outlier data

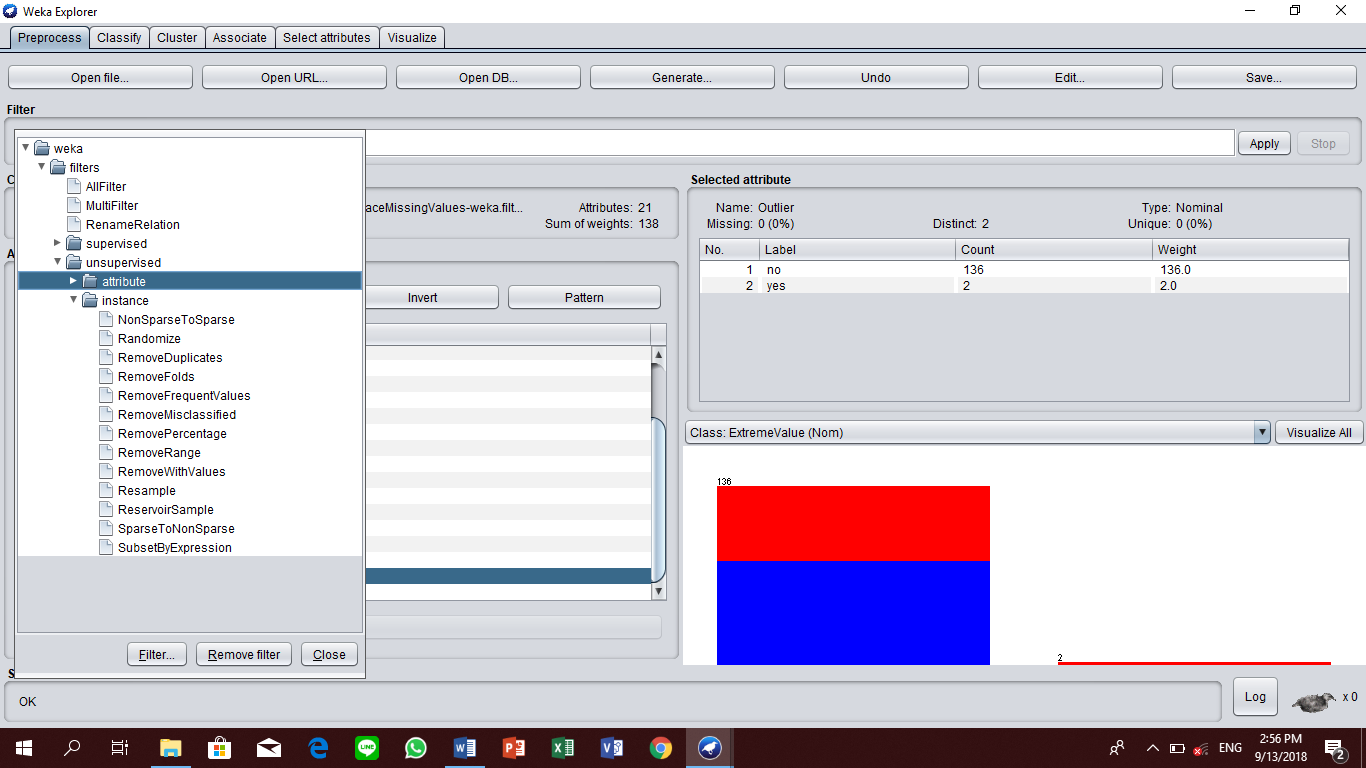


Figure 8 Proses menghilangkan outlier menggunakan filter RemoveWithValues

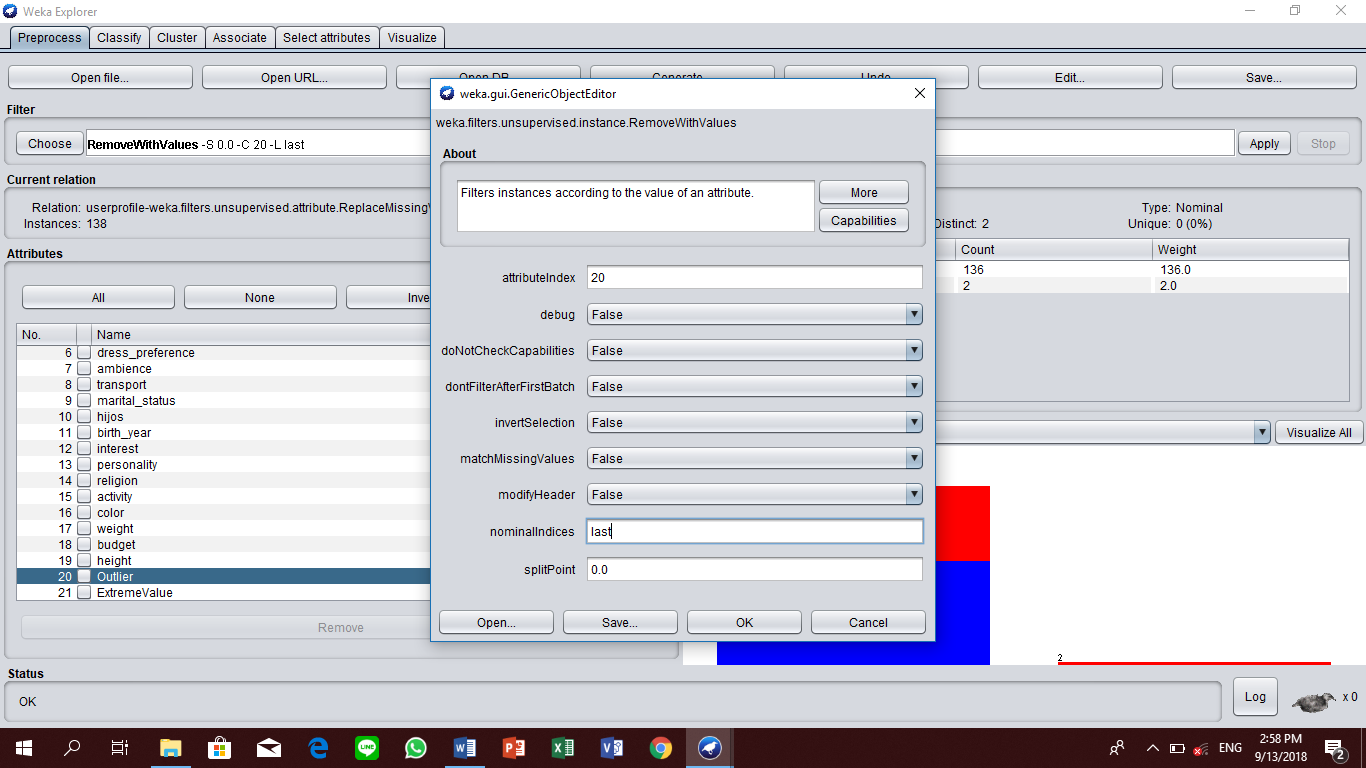


Figure 9 Parameter untuk filter RemoveWithValues

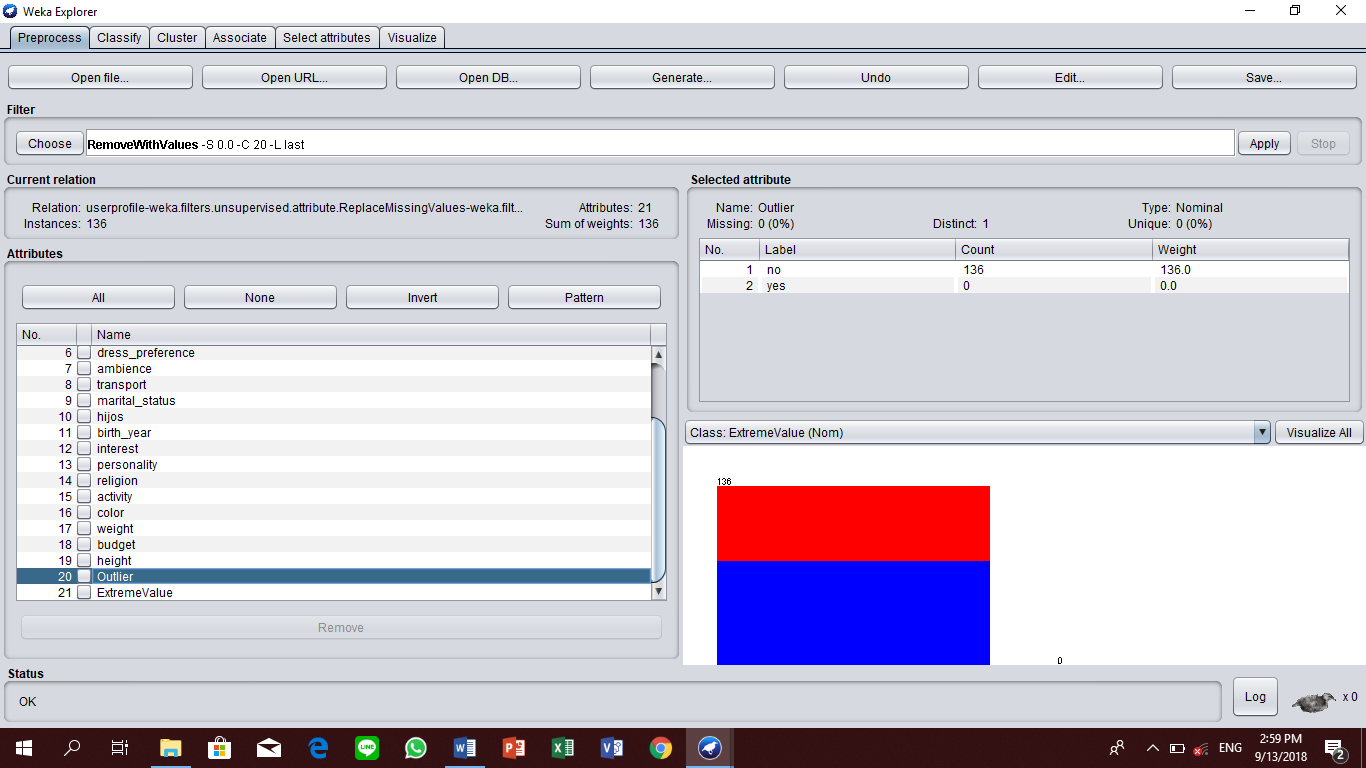


Figure 10 Data setelah menggunakan filter RemoveWithValues

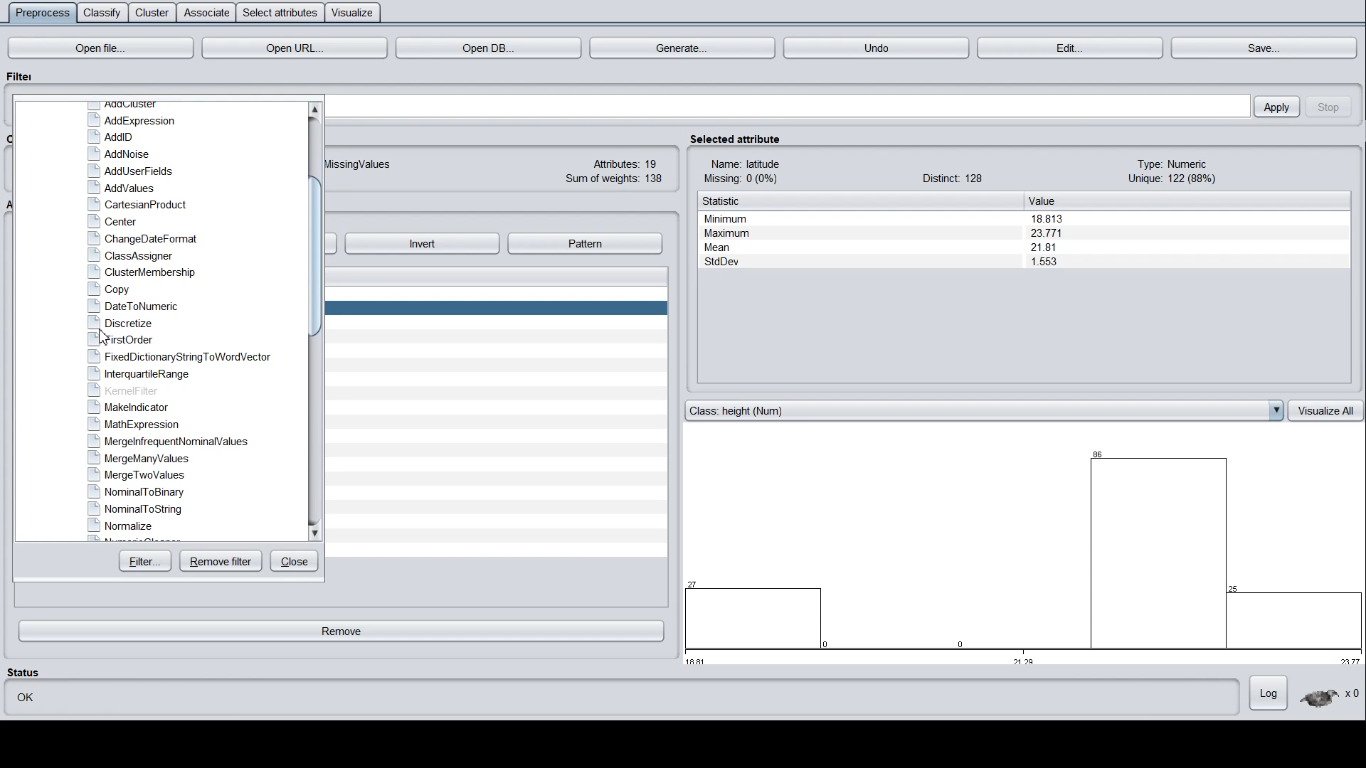


Figure 11 Tahapan diskretisasi

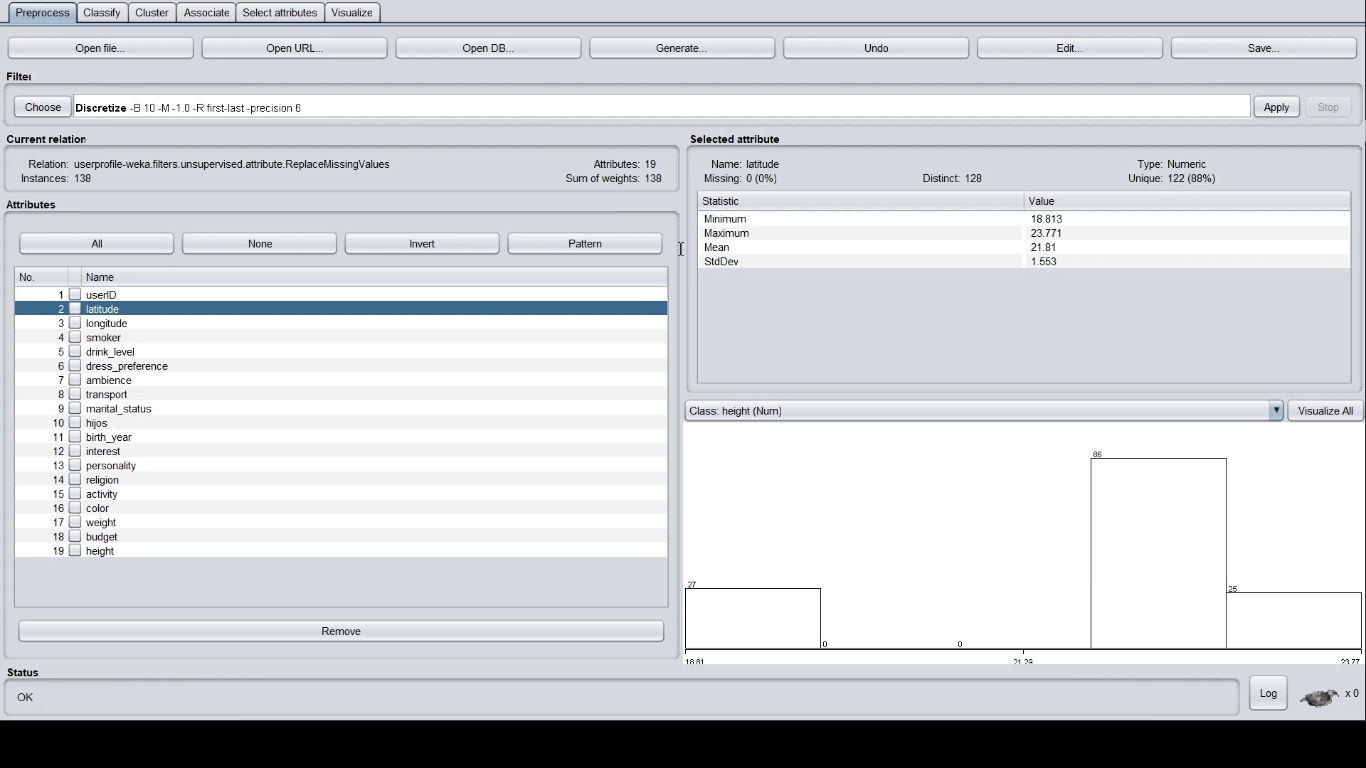


Figure 12 Data sebelum di diskretisasi

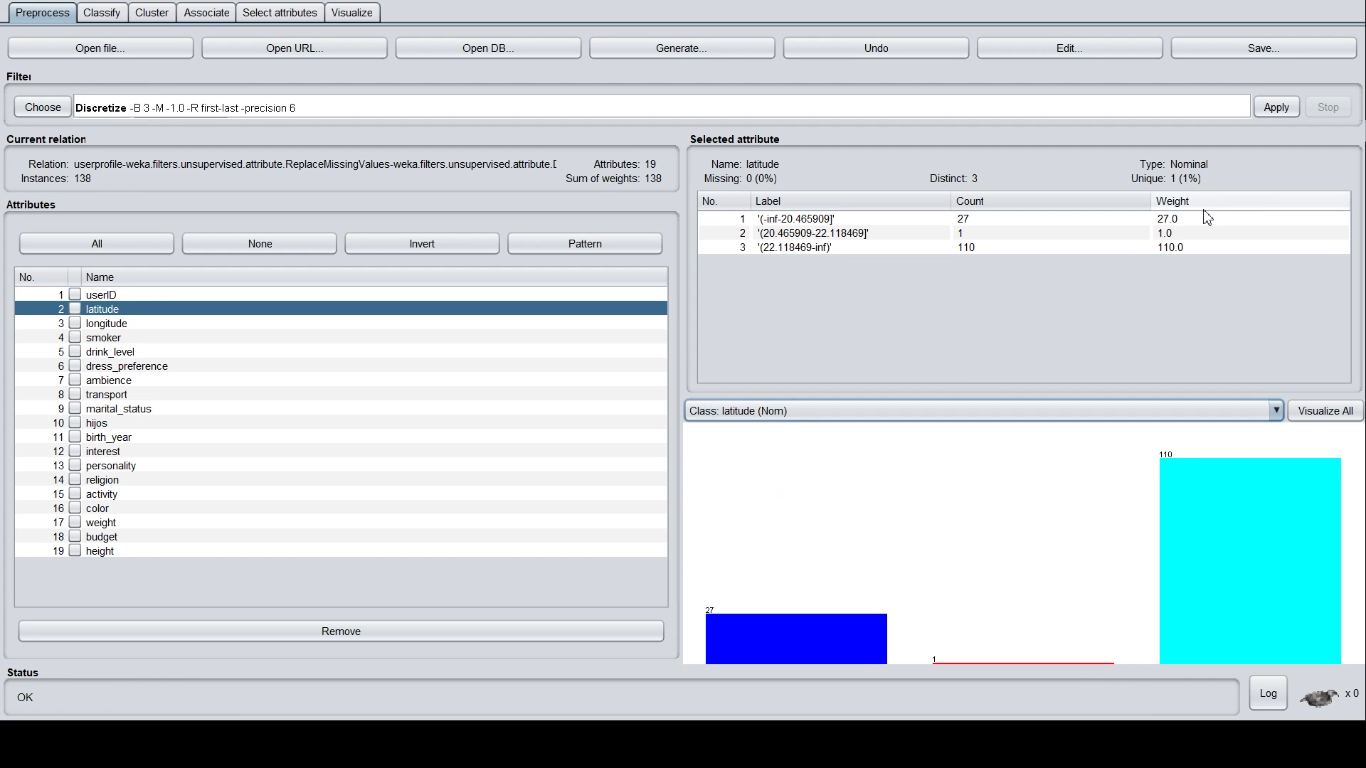


Figure 13 Data sesudah di diskretisasi

Link youtube : https://www.youtube.com/watch?v=YsSpR3EhnYc&feature=youtu.be

# REFERENSI

* <https://asnugroho.wordpress.com/2016/10/11/langkah-langkah-preprocessing/>
* Slide pembelajaran Data Mining IF
* https://www.youtube.com/watch?v=m7kpIBGEdkI