

Features= 15
Rows = 32561

- **age**: continuous.
- **workclass**: nominal(missing=1836)
- **fnlwgt**: continuous.
- **education**:ordinal.
- **education-num**: continuous.
- **marital-status**: nominal
- **occupation**: nominal(missing=1843)
- **relationship**: nominal
- **race**: nominal
- **sex**: nominal
- **capital-gain**: continuous.
- **capital-loss**: continuous.
- **hours-per-week**: continuous.
- **native-country**: nominal(missing=583)

CLASS: ordinal, two calsses(<50K,>=50K)

1.No inconsistent spellings

2.Data to WEKA should be formatted properly before iputiing

3.No Dates

4.Missing values

Filter(Choose)>Unsupervised>Attribute>**ReplaceMissingValues**

Weka Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Open fil... | Open U... | Open D... | Generat... | Undo | Edit... | Save...

Filter

- PartitionedMultiFilter
- PKIDiscretize
- PrincipalComponents
- PropositionalToMultiInstance
- RandomProjection
- RandomSubset
- RELAGGS
- Remove
- RemoveType
- RemoveUseless
- Reorder
- ReplaceMissingValues**
- Standardize
- StringToNominal
- StringToWordVector
- SwapValues
- TimeSeriesDelta
- TimeSeriesTranslate
- Wavelet
- instance

Selected attribute

Name: workclass | Type: Nominal
Missing: 1836 (6%) | Distinct: 8 | Unique: 0 (0%)

No.	Label	Count
1	Private	22696
2	Self-emp-not-inc	2541
3	Self-emp-inc	1116
4	Federal-gov	960
5	Local-gov	2093
6	State-gov	1298
7	Without-pay	14
8	Never-worked	7

ReplaceMissingValues

Replaces all missing values for nominal and numeric attributes in a dataset with the mode

CAPABILITIES

Class -- Unary class, Date class, String class, Relational class, Nominal class, No class, Bin

Attributes -- Empty nominal attributes, Binary attributes, Date attributes, Relational attrib

Unary attributes

Additional
min # of instances: 0

Status: OK

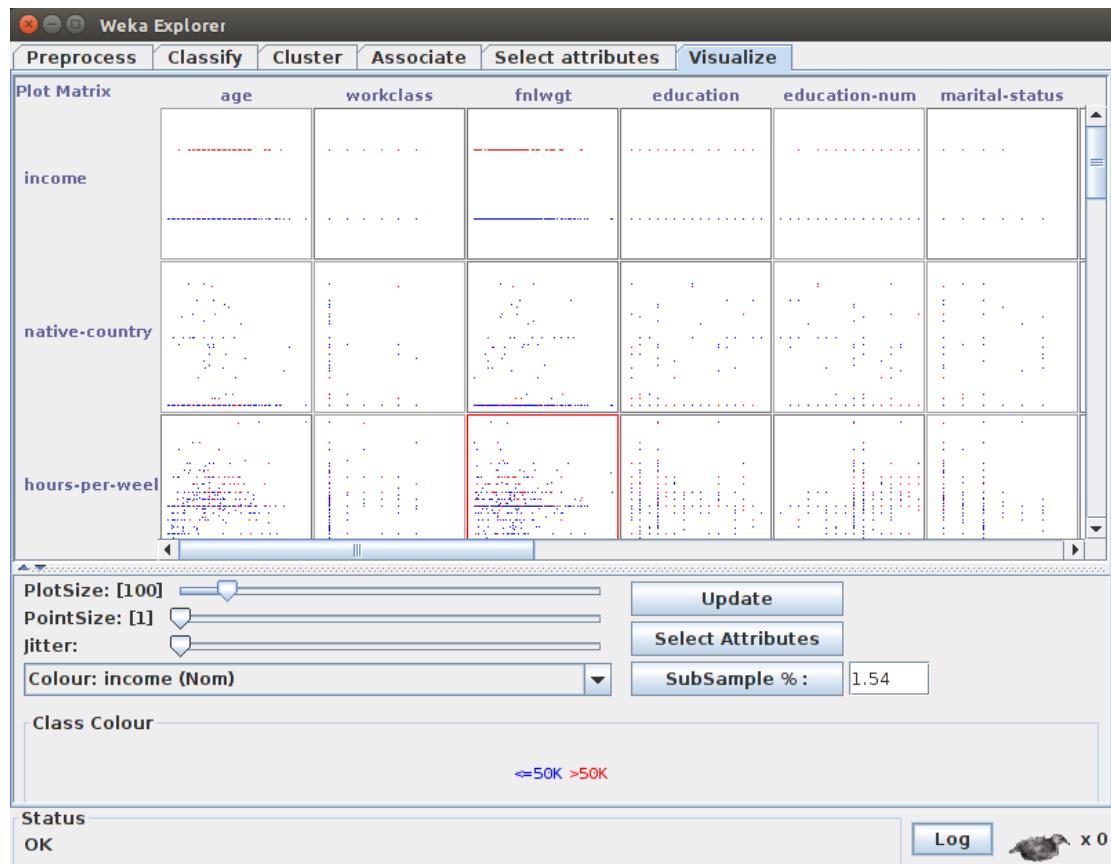
Log

THIS IS NOT SUPPORTED BY OPEN-REFINE.

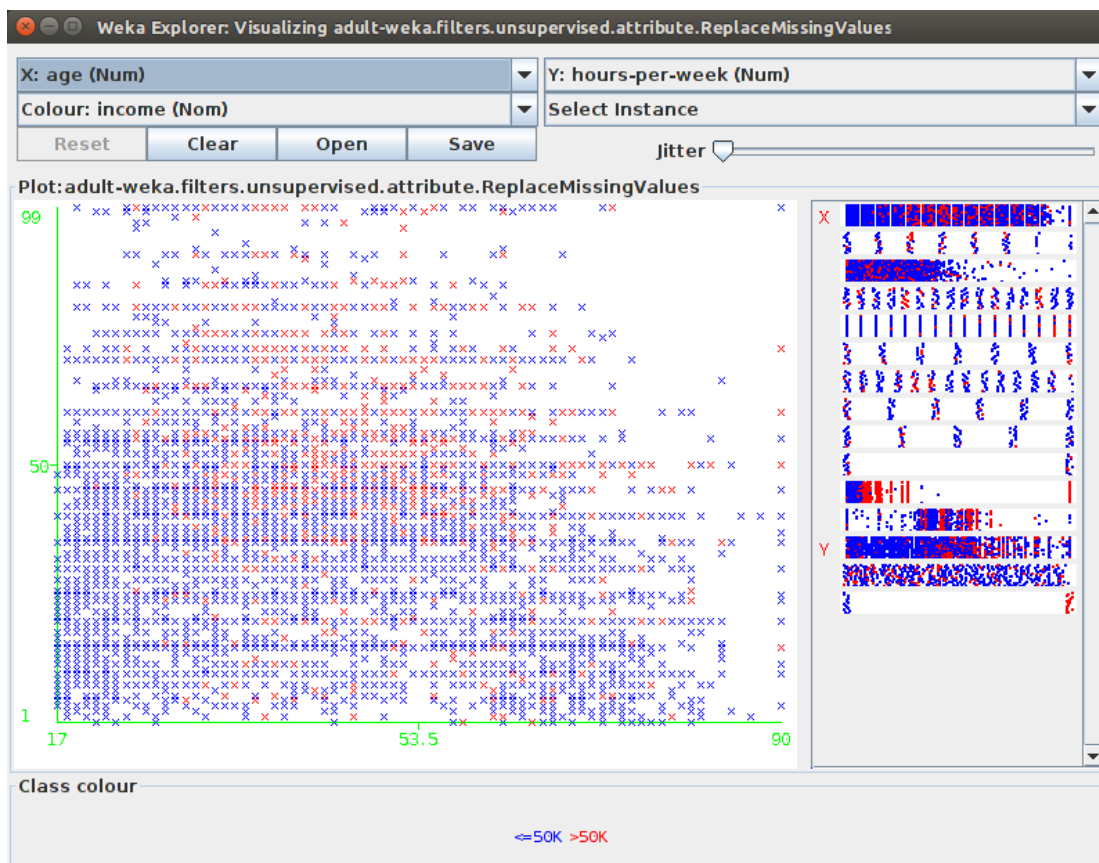
5.No duplicates

6.Scatter-Plot

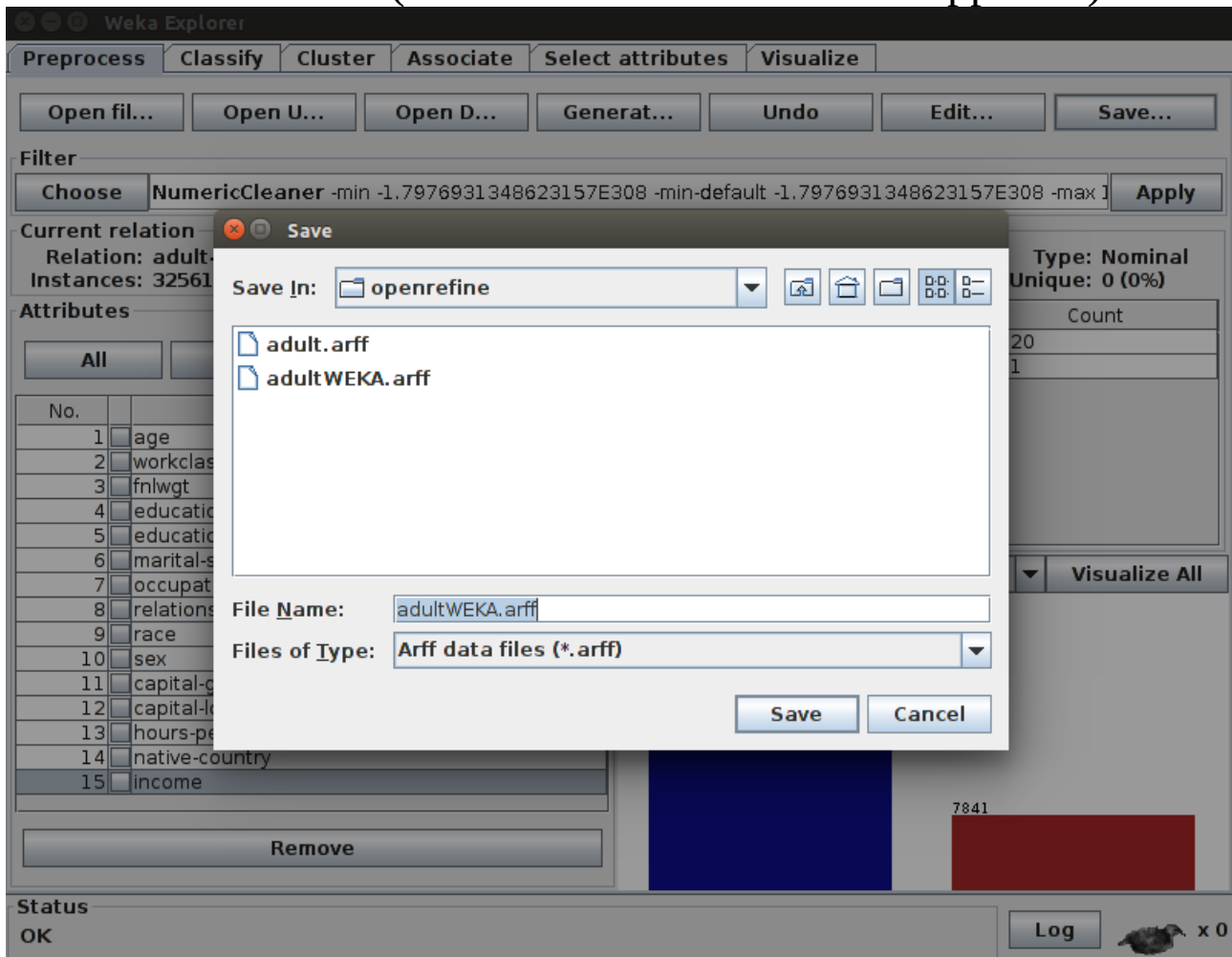
Select **Vizualize** tab



Selecting anyone, we get a clear view and flexibility to change X,Y axes as we wish



7. Save to an **arff** file(csv and some other are too supported)



SUMMARY:

Open-Refine **cleans** data

Weka **processes** data