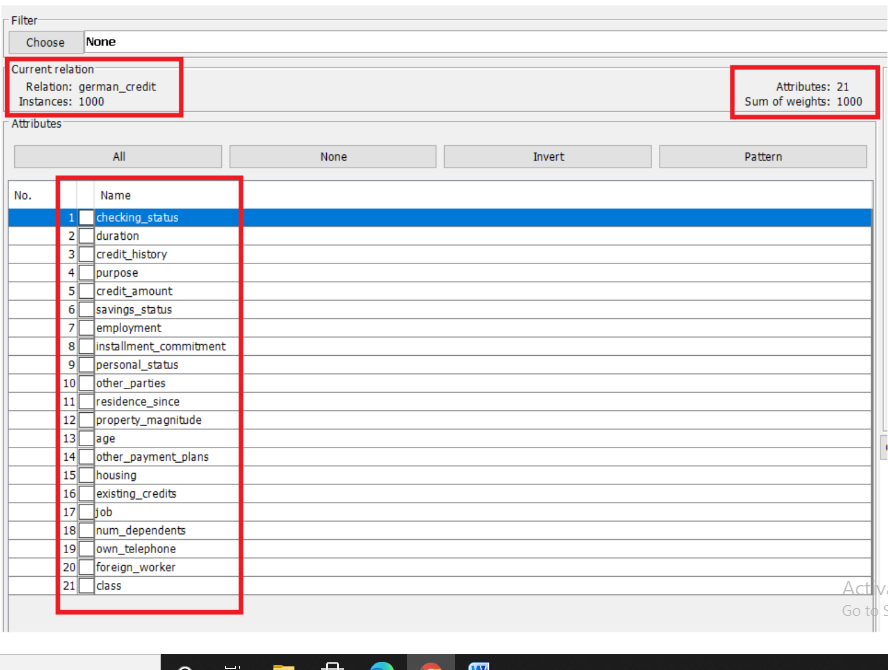
**DATA MINING LAB ASSIGNMENT - 1**

**OBSERVATIONS FOR GERMAN CREDIT DATASETS:**

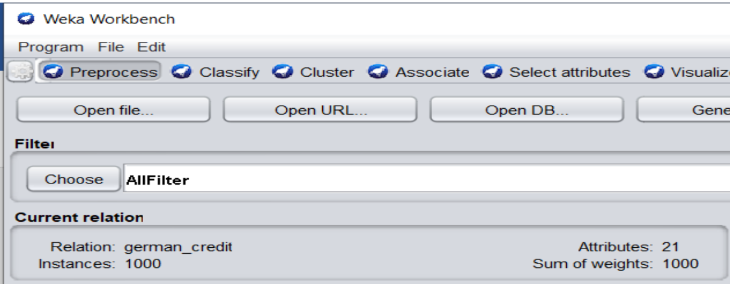
1. List the attribute names and their types?



|  |  |
| --- | --- |
| **ATTRIBUTE NAME** | **ATTRIBUTE TYPE** |
| Checking\_status | Nominal |
| Duration | Numeric |
| Credit\_history | Nominal |
| Purpose | Nominal |
| Credit\_amount | Numeric |
| Saving status | Nominal |
| Employment | Nominal |
| Installement\_commitment | Numeric |
| Personal\_status | Nominal |
| Other parties | Nominal |
| Residence\_since | Numeric |
| Property\_magnitude | Nominal |
| Age | Numeric |
| Other\_payment\_plans | Nominal |
| Housing | Nominal |
| Exisiting\_credits | Numeric |
| Job | Nominal |
| Num\_dependents | Numeric |
| Own\_telephone | Nominal |
| Foreign\_workers | Nominal |
| Class | Nominal |

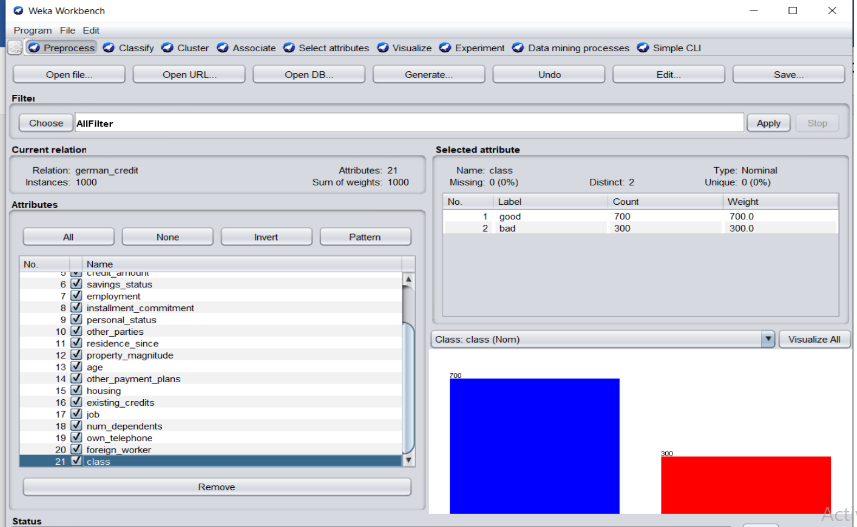
1. Number of records in each dataset?

* The German credit dataset consists of 1000 records with 21 attributes

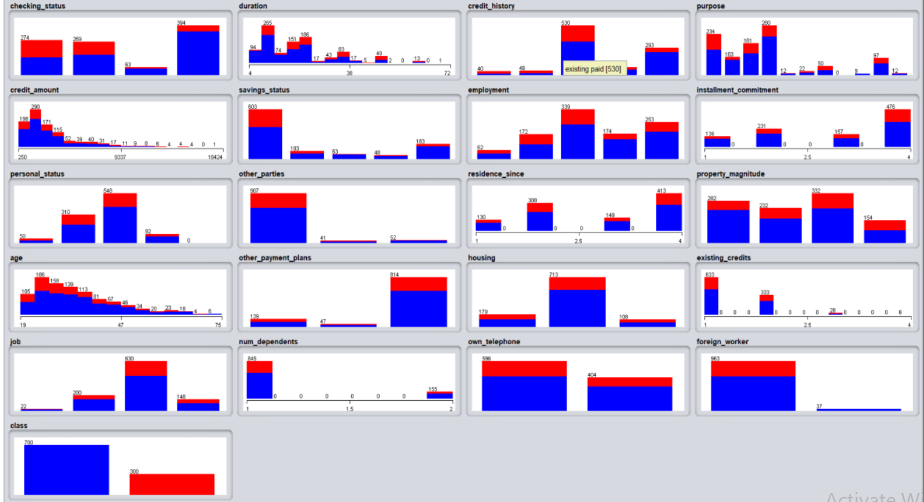


1. Identify the class attribute?

* Class Attribute in dataset = Class



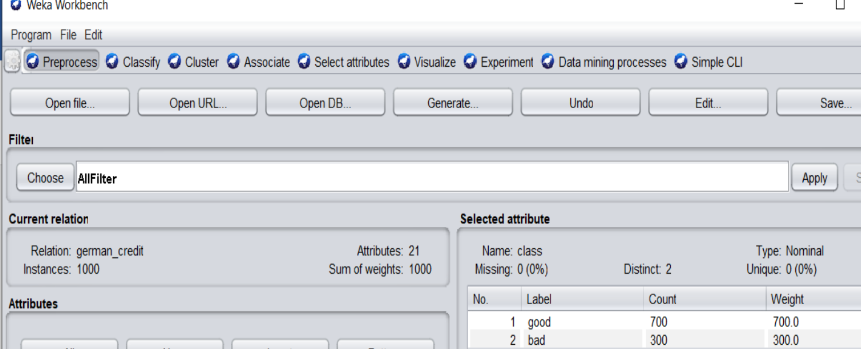
1. plot histogram:



* Here red color represents “bad class”.
* Blue color represents “good class”.

5.)\_Determine the number of records for each class.

* Here good --700 records
* Bad - 300 records



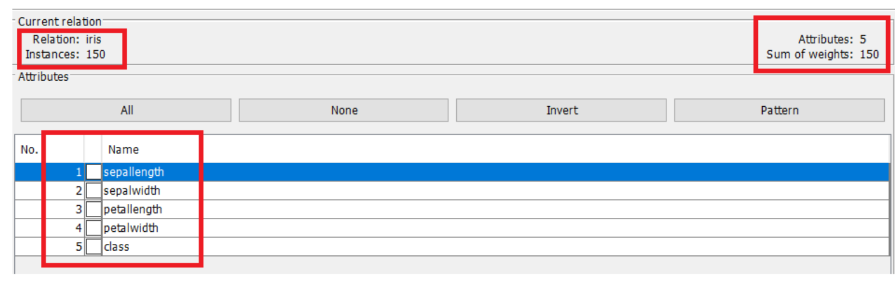
**IRIS DATASET:**

**OBSERVATIONS:**

In this dataset we observe that we can classify flowers into 3 categories.

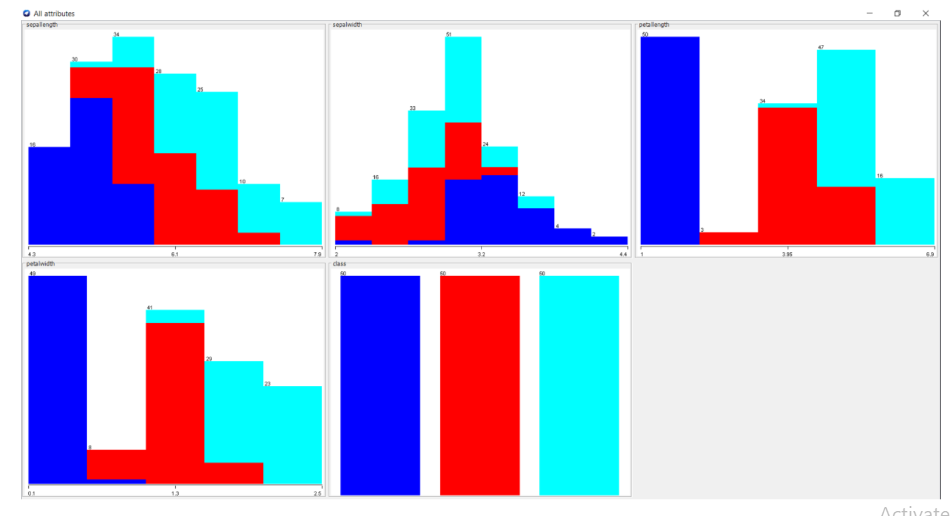
1. Numbers of records, attribute names and types:

* In this iris dataset consists of 150 records with 5 attributes, and two types (Numeric, Nominal).



|  |  |
| --- | --- |
| **ATTRIBUTE NAME** | **ATTRIBUTE TYPE** |
| sepallength | Numeric |
| sepalwidth | Numeric |
| petallength | Numeric |
| petalwidth | Numeric |
| class | Nominal |

1. PLOT HISTOGRAM:

****

* In this dataset we have 3 classes.
* Blue represents “Iris-setosa”.
* Red represents “Iris-versicolor”.
* Light-Blue represents “Iris-virginica”.