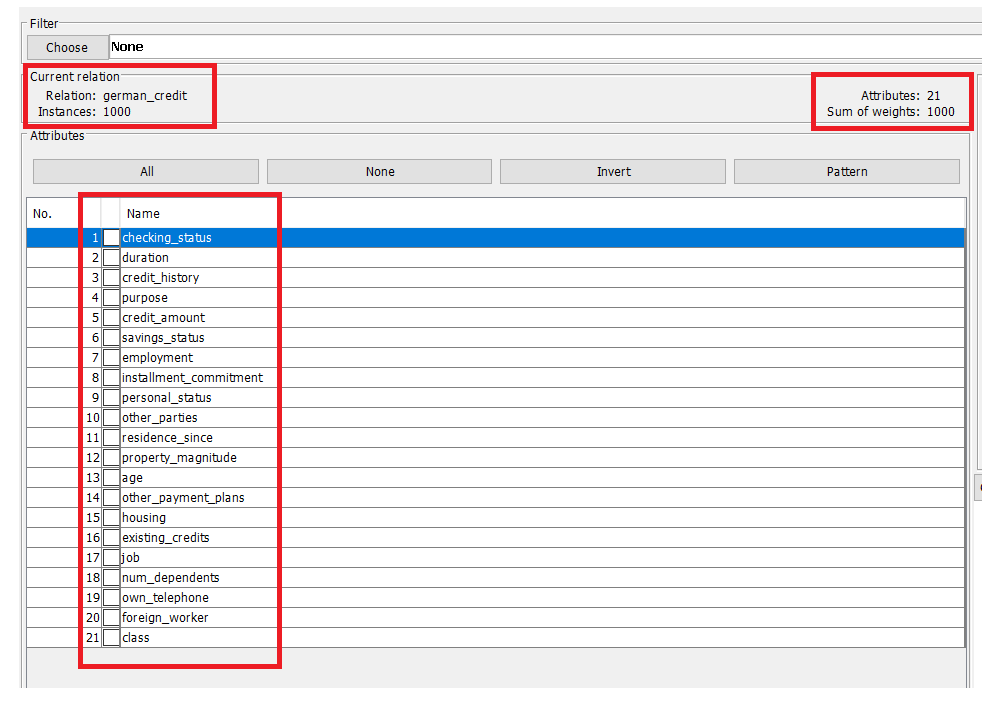
**DATA MINING LAB ASSIGNMENT – 1**

**GERMAN CREDIT DATASET:-**

**OBSERVATIONS:-**

The german\_credit dataset classifies people to either good or bad credit risks.

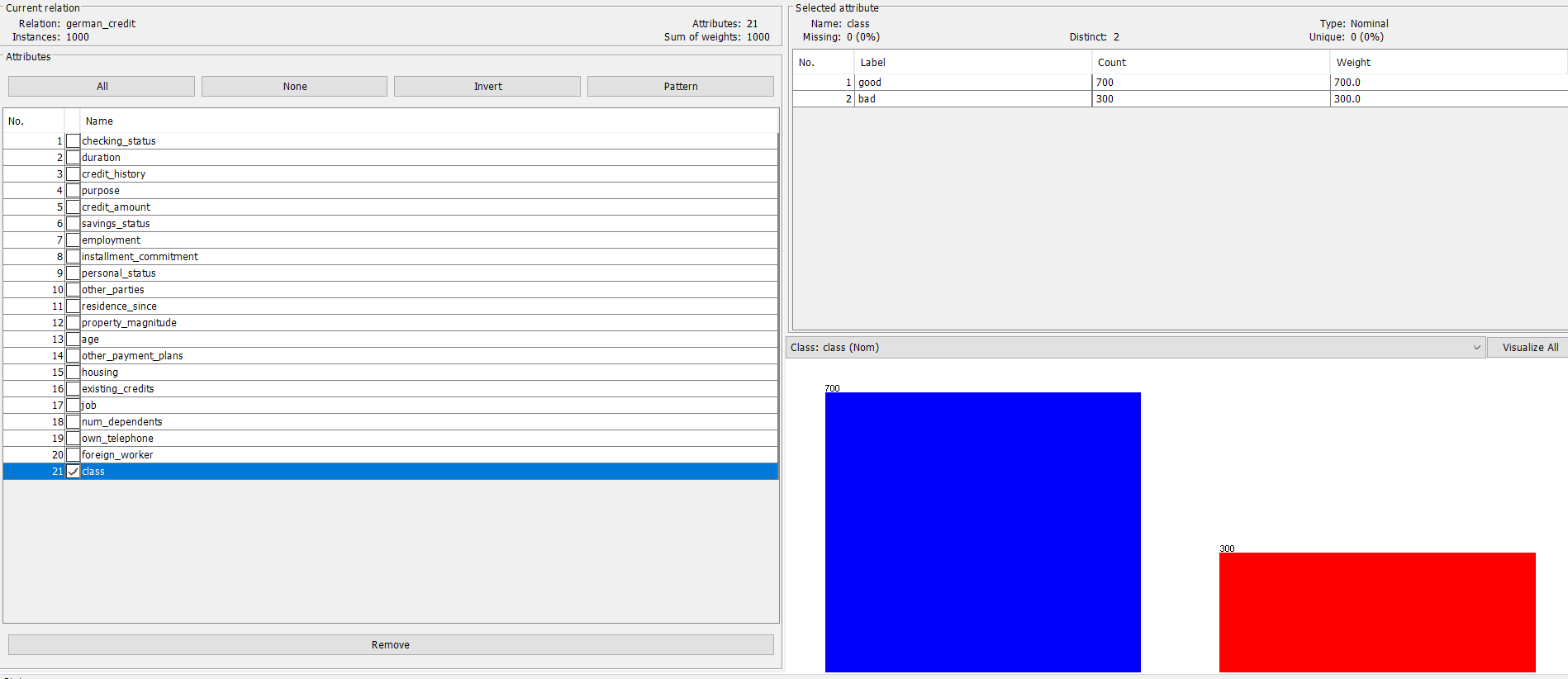
****

**Number of records, attribute names and types:-**

The german credit dataset consist of 1000 records with 21 attributes.

|  |  |
| --- | --- |
| **Attribute Name** | **Attribute Type** |
| checking\_status | Nominal |
| duration | Numeric |
| credit\_history | Nominal |
| purpose | Nominal |
| credit\_amount | Numeric |
| savings\_status | Nominal |
| employment | Nominal |
| Installment\_commitment | Numeric |
| Personal\_status | Nominal |
| Other\_parties | Nominal |
| residence\_since | Numeric |
| property\_magnitude | Nominal |
| age | Numeric |
| other\_payment\_plans | Nominal |
| housing | Nominal |
| existing\_credits | Numeric |
| job | Nominal |
| num\_dependents | Numeric |
| own\_telephone | Nominal |
| foreign\_worker | Nominal |
| class | Nominal |

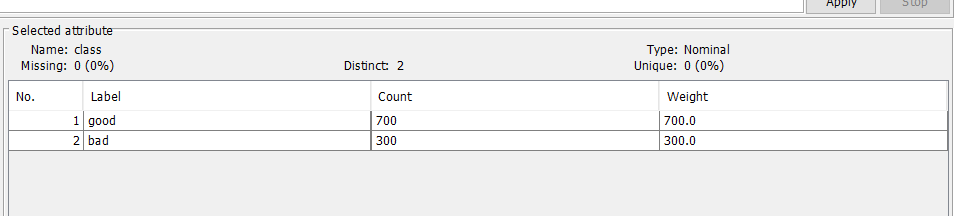
**Class attribute in the dataset:**

****

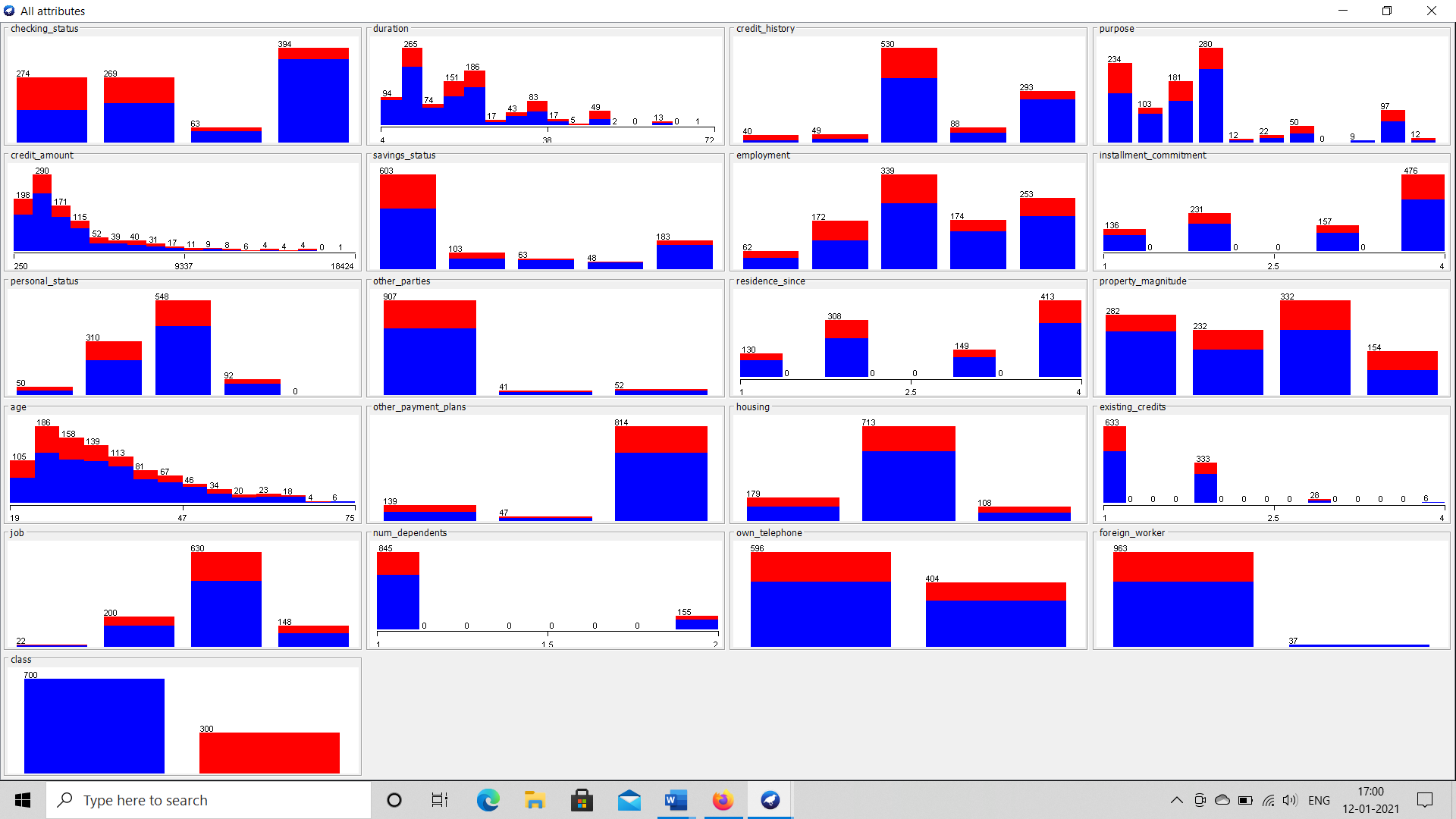
**Number of records for class:**

Good – 700

Bad – 300



**Plot histogram:-**



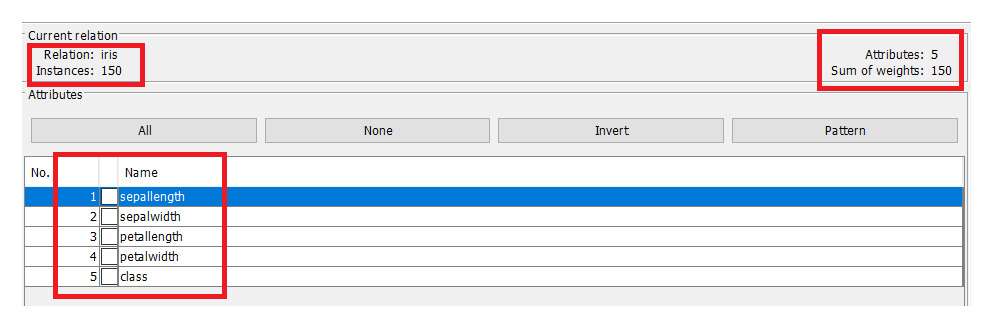
Blue colour represents good and red colour represents bad class.

**IRIS DATASET:-**

**OBSERVATIONS:-**

It is a dataset through which we classify flowers into 3 categories.

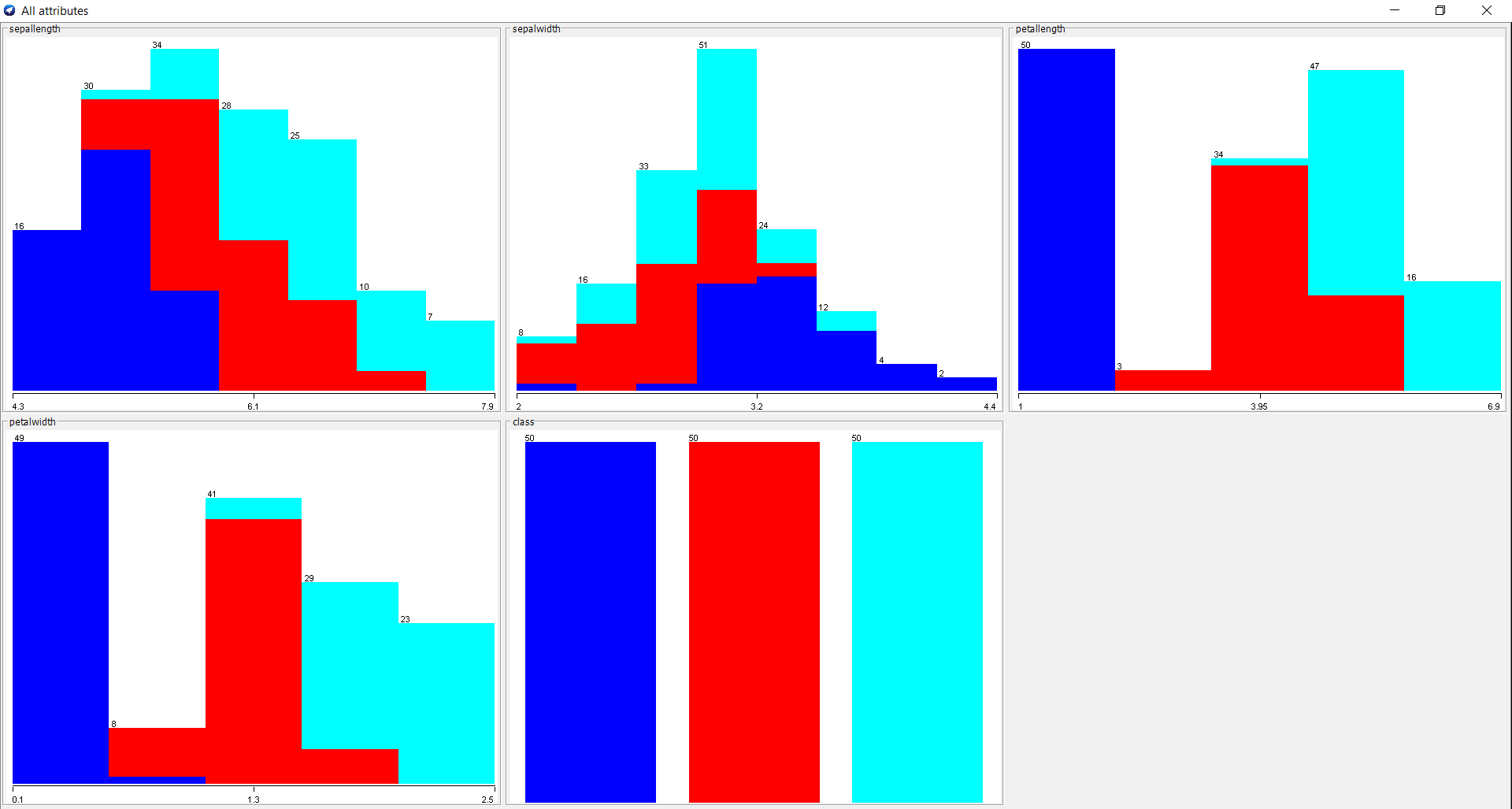
**Number of records, attribute names and types:**

****

The iris dataset consists of 150 records with 5 attributes.

|  |  |
| --- | --- |
| **Attribute Name** | **Attribute Type** |
| sepallength | Numeric |
| sepalwidth | Numeric |
| petallength | Numeric |
| petalwidth | Numeric |
| class | Nominal |

**Plot Histogram:**



There are 3 classes in the dataset. Where blue represents Iris-setosa, red represents Iris-versicolor and light blue represents Iris-virginica.