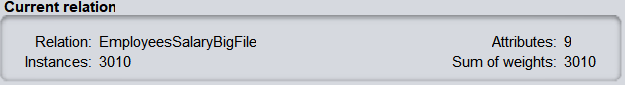
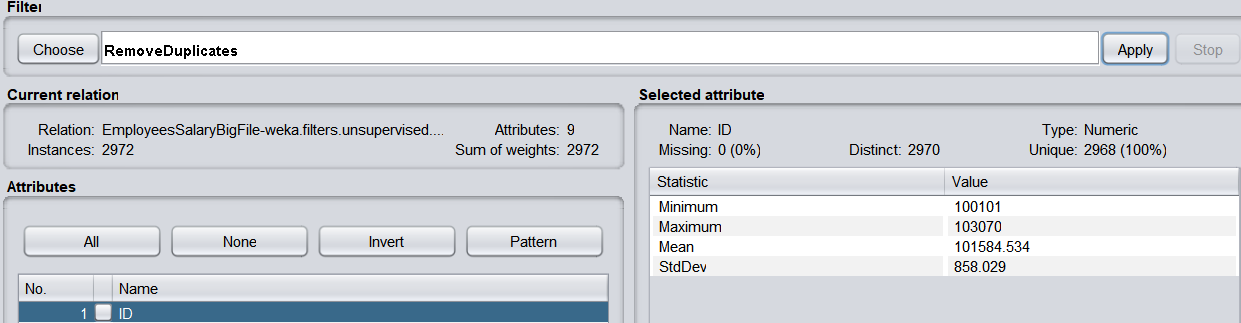
1.Basic operations:

Instances:3010



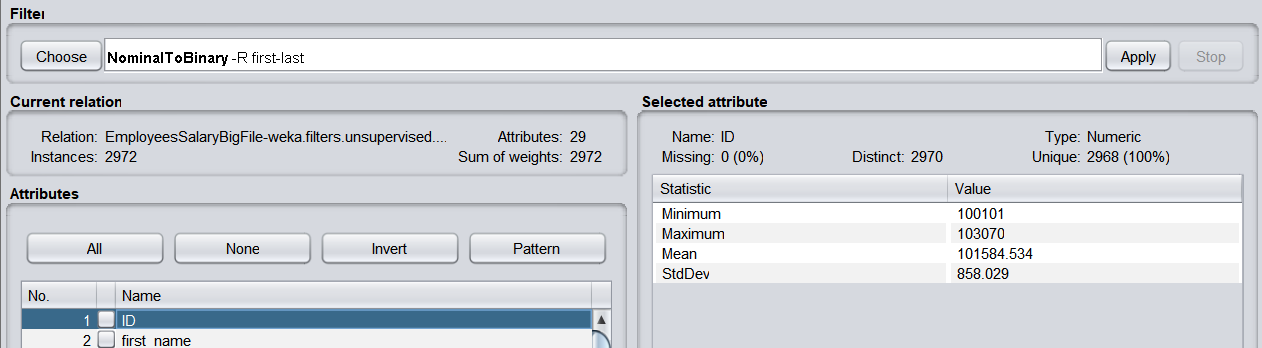
2.Duplicates removing

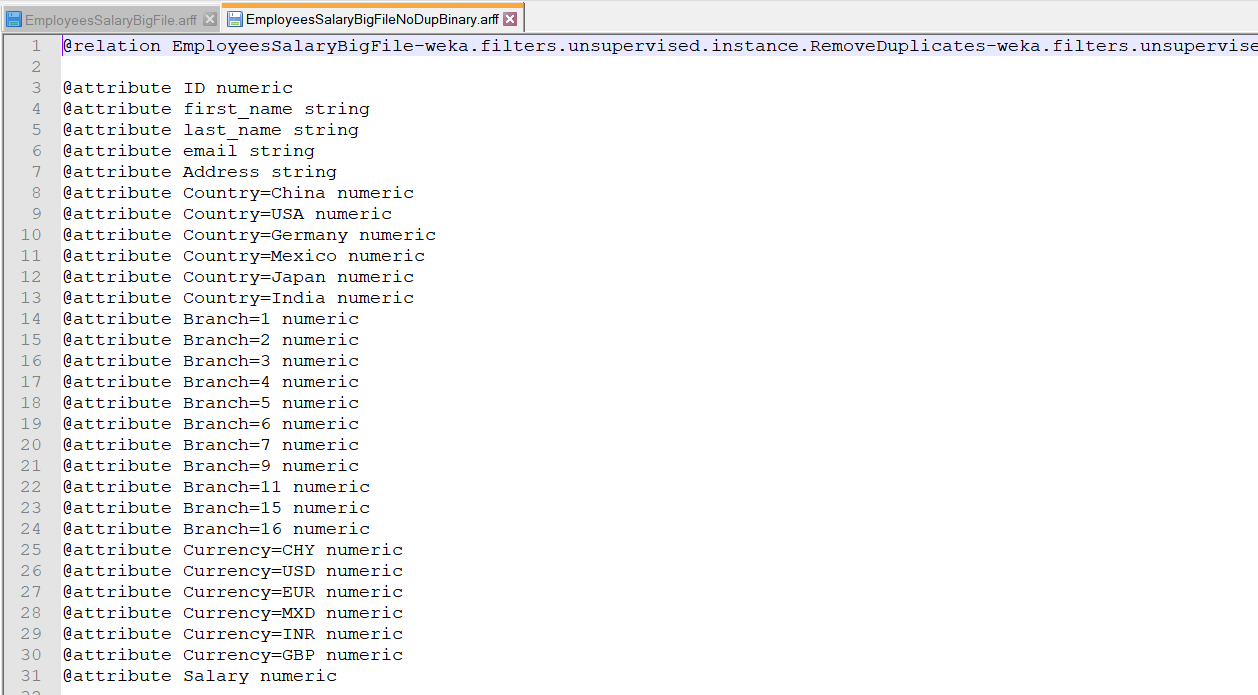
Instances: 2972 Duplicates:38 (3010 the number before removing duplicates-2972)



3.Converting attributes:

3 nominal attributes: Country, Branch, Currency





For my analysis:

For example, after removing data there are 38 duplicates in this EmployeesSalaryBigFile.csv, that means one employee has more than 1 salary, this is not reasonable.

Through transforming attribute from Nominal to Binary, we will do some algorithm later, for example, clustering.