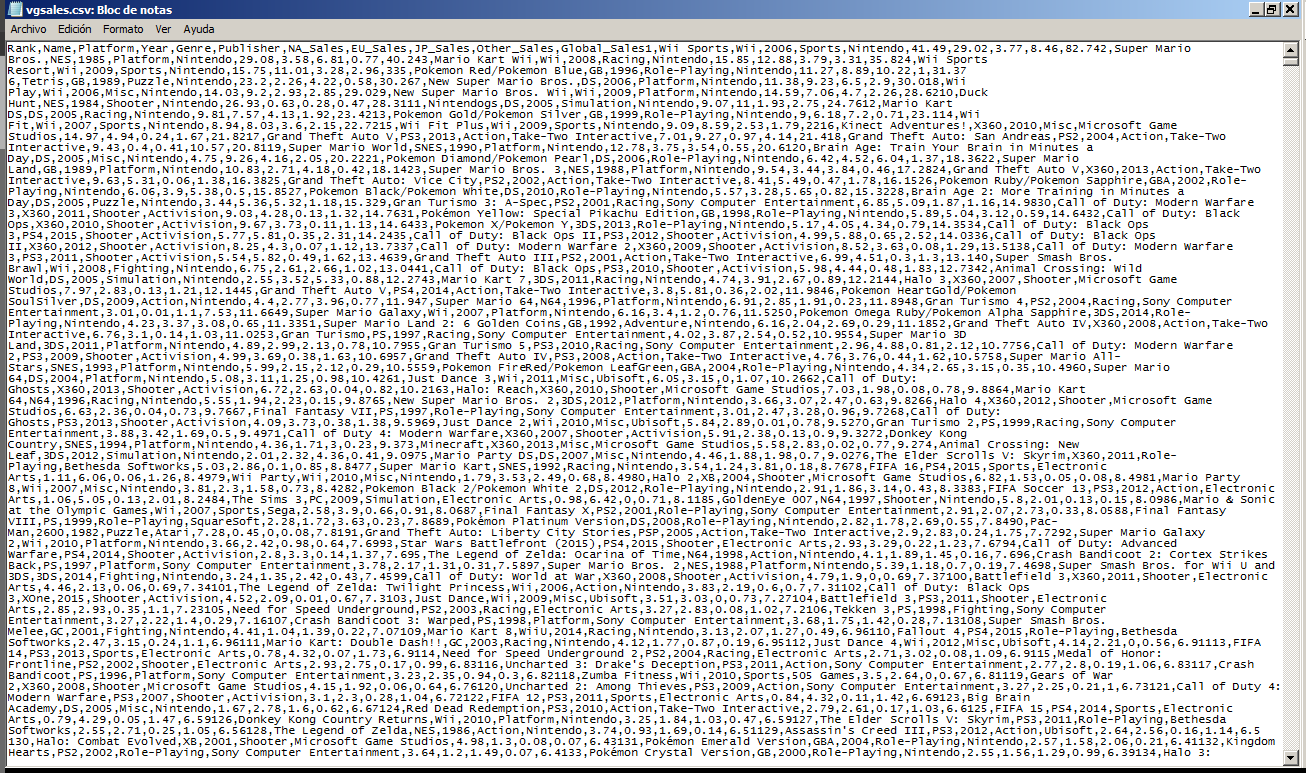
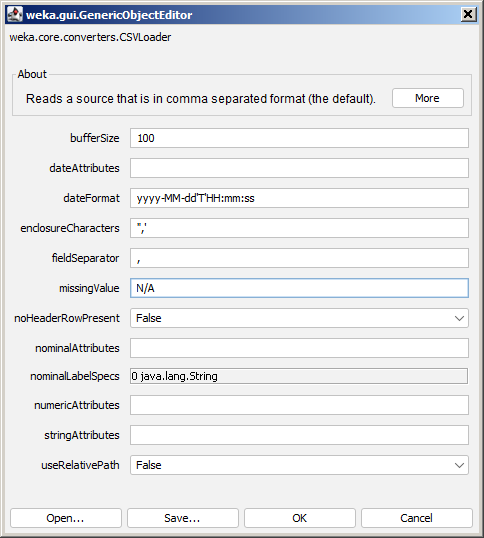
**2. Del dataset anterior realice en WEKA, tres algoritmos de preprocesamiento.**

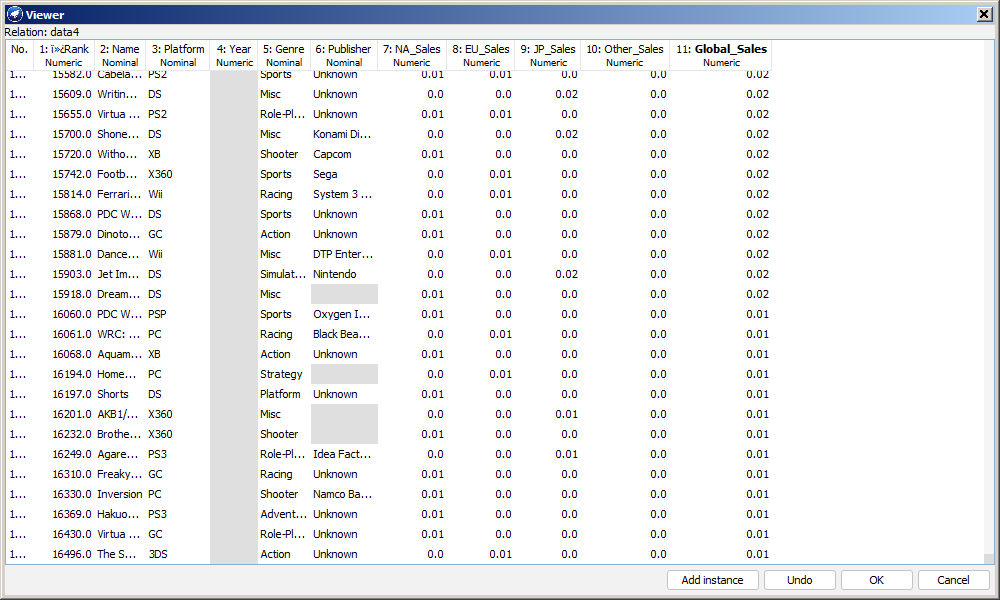
En primer lugar limpiamos los caracteres que estan provocando errores asi como son ";, que se presentan en los nombres de los video juegos de (vgsales.csv) y lo guardamos en (data4.csv)



Seguidamente cargamos el archivo (data4.csv) en WEKA, y asignamos el valor de missingValue a los datos de algunas columnas que presentan datos como “N/A”



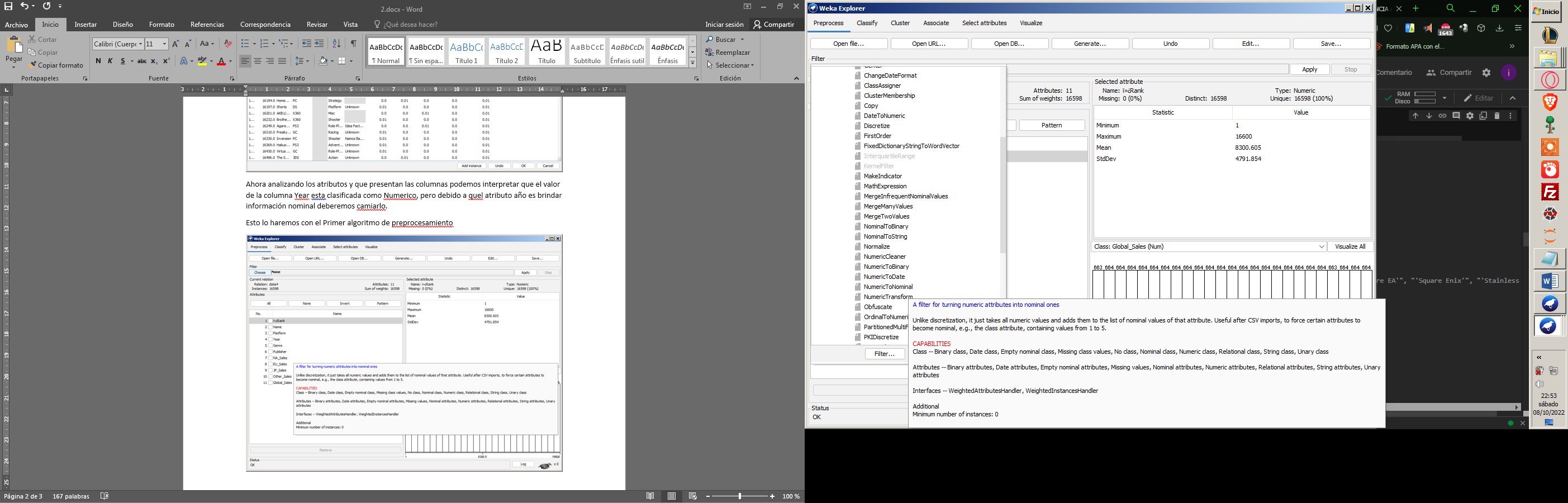
De esta forma los valores nulos se presentaran como espacios vacíos, tal es el ejemplo de la columna Year.



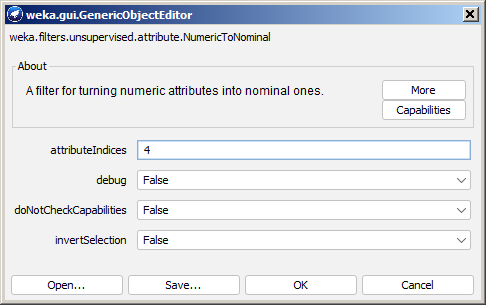
Ahora analizando los atributos y que presentan las columnas podemos interpretar que el valor de la columna Year esta clasificada como Numerico, pero debido a quel atributo año es brindar información categorica deberemos cambiarlo, y como el valor de La columna Publisher fue designado como “N/A” y “Unknown” estadarizamos manualmente a que el valor por defecto se “Unknown”

**1.**

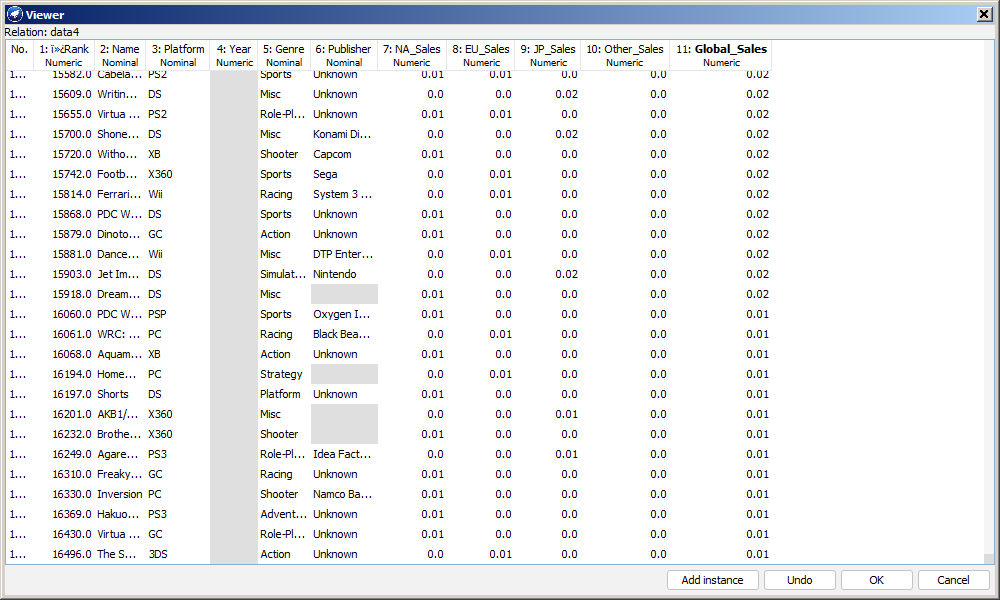
El Primer algoritmo de preprocesamiento “NumericToNominal” que nos ayuda a cambiar el atributo que tenga una columna numérica de Year

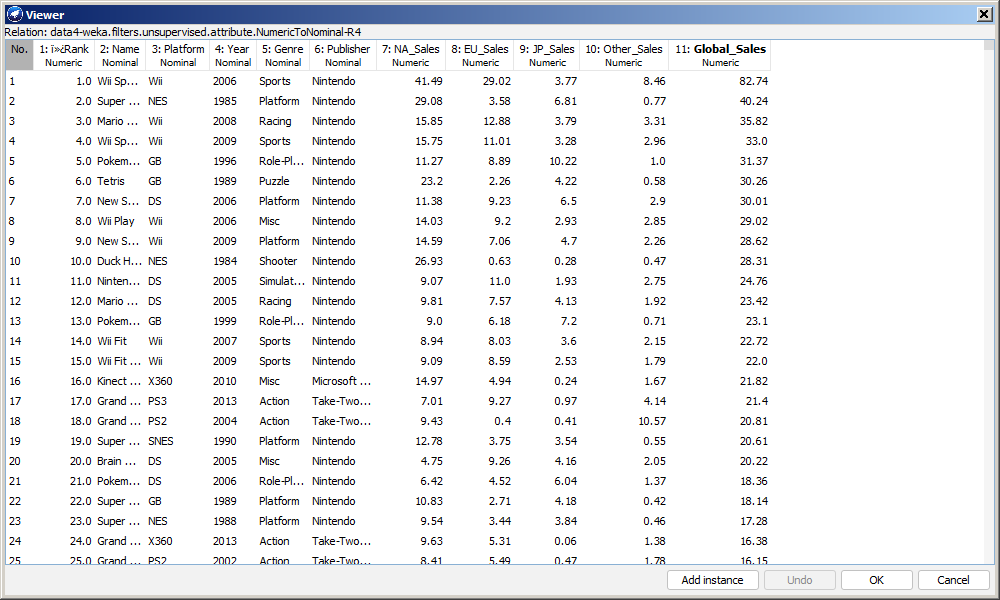


En su configuración solo tomamos la columna 4 que es la que necesitamos



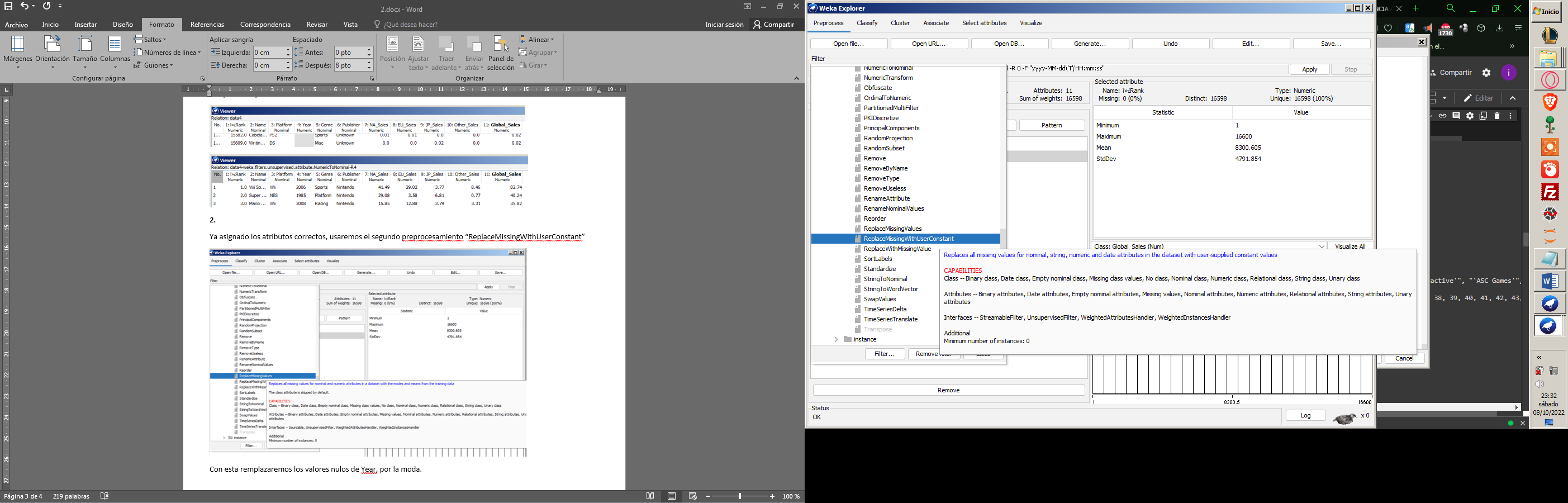
Lo aplicamos y vemos el cambio



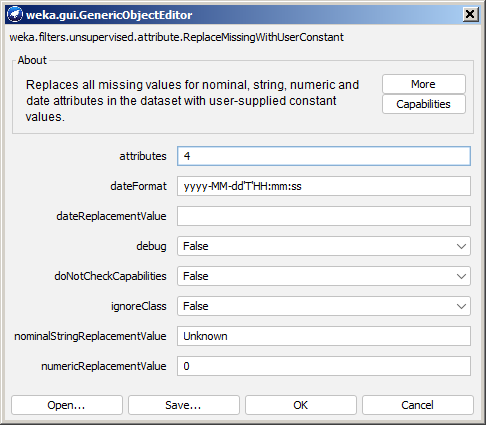
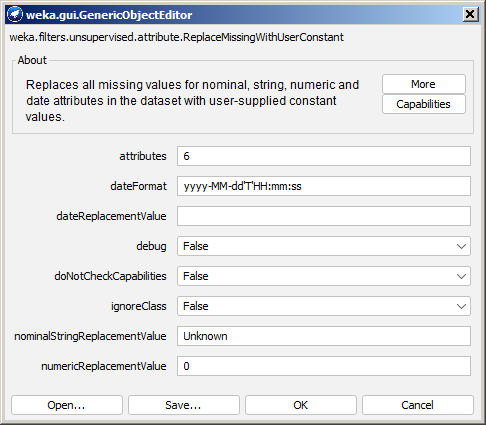


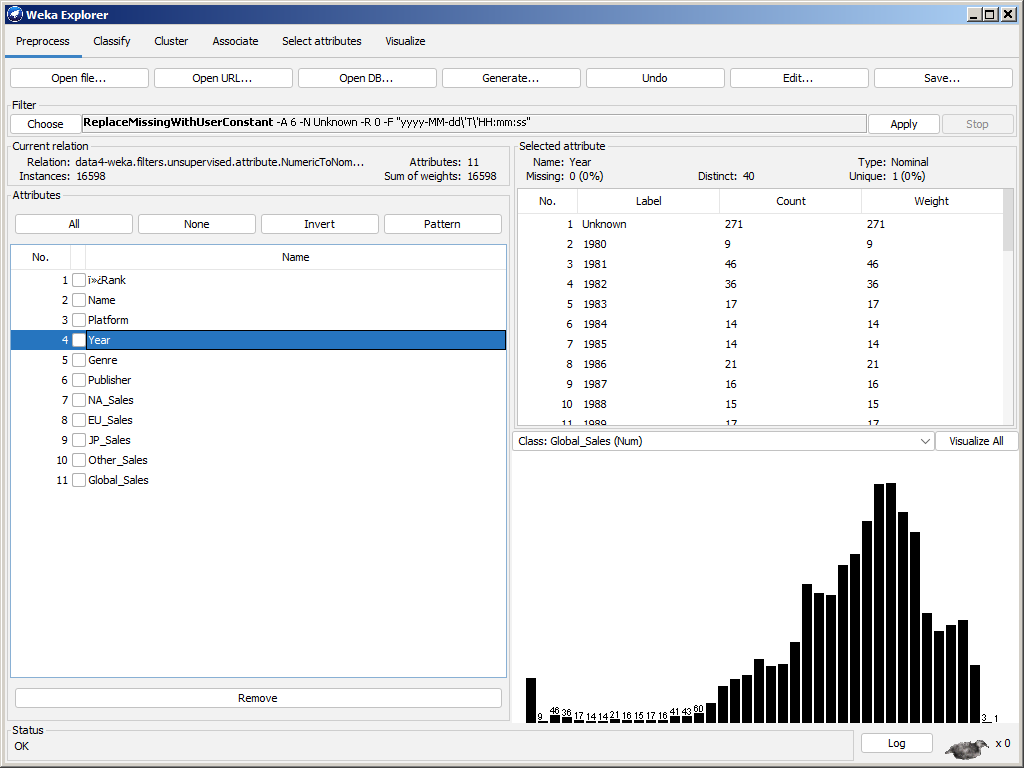
**2.**

Ya asignado los atributos correctos, usaremos el segundo preprocesamiento “ReplaceMissingWithUserConstant”



Con el cual en los valores perdidos los rellenaremos con el valor estándar Unknown

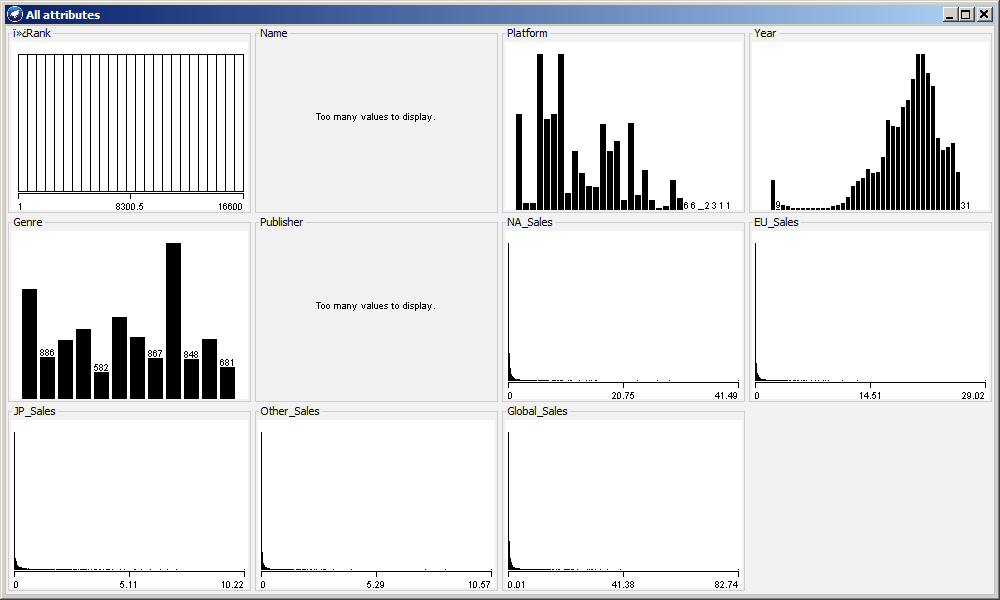
 



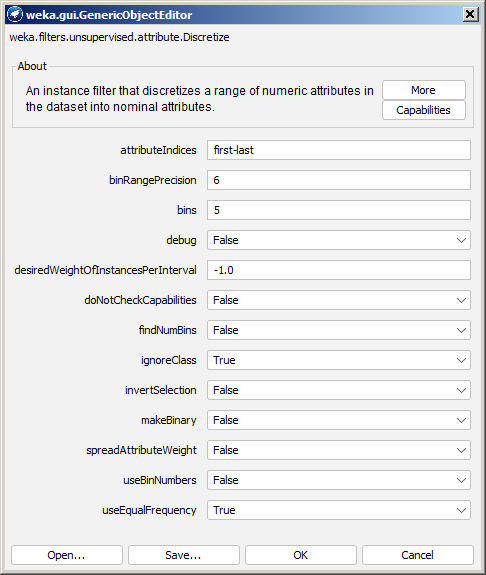
Con esta remplazaremos los valores nulos de Year y Publisher.

**3.**

Analizando un poco la visualización de los valores numericos podemos ver que los valores se encuentran demasiado dispersos por lo cual podemos normalizar los valores con “Discretize”



Configuraremos el valor a 5 rangos, manejaremos el uso de frecuencias iguales y ignoremos la clase puesto que con el dataset podemos asignar diferentes columnas como columna de clase



Volviendo a visualizar veremos que los valores se encuentran mejor ordenado

