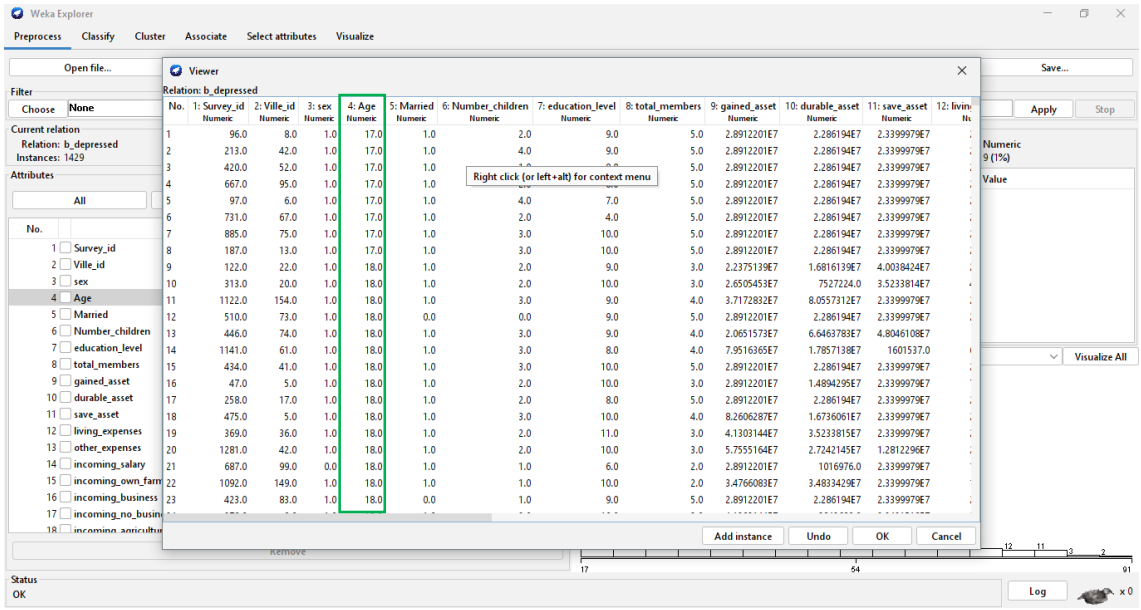
UNIVERSIDAD MAYOR DE SAN ANDRÉS FACULTAD DE CIENCIAS PURAS Y NATURALES CARRERA DE INFORMÁTICA INF-354		A
Nombre: Mery Zulema Apaza Calderon	**CI:** 11068229 LP	
Docente: Moises Silva	**Paralelo:** A	**INICIAL A. P.**

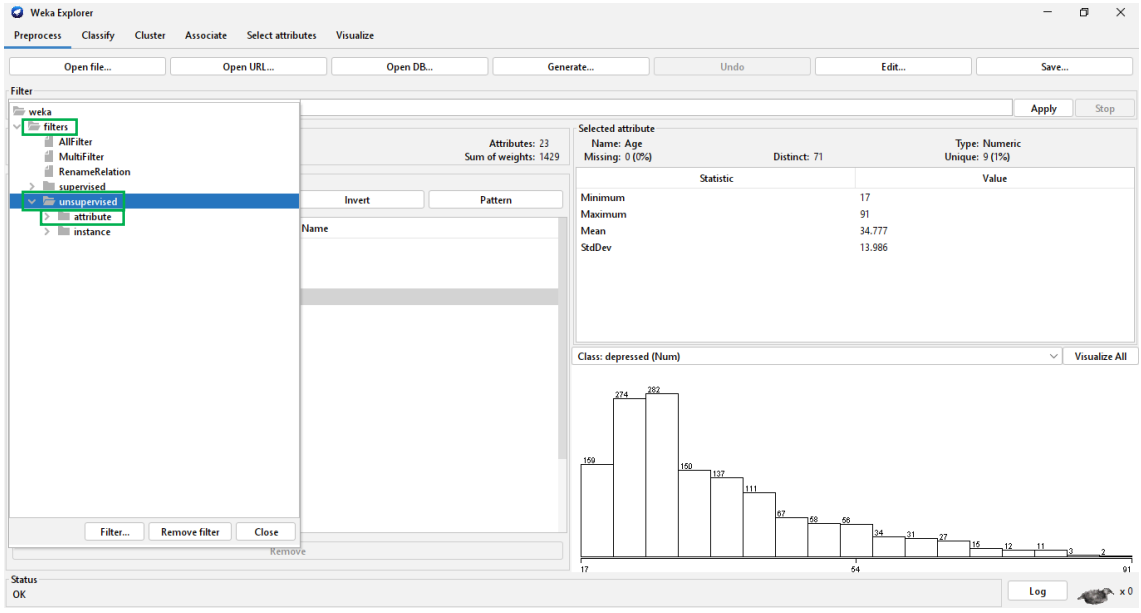
CUATRO TÉCNICAS DE PREPROCESAMIENTO

1. Discretización

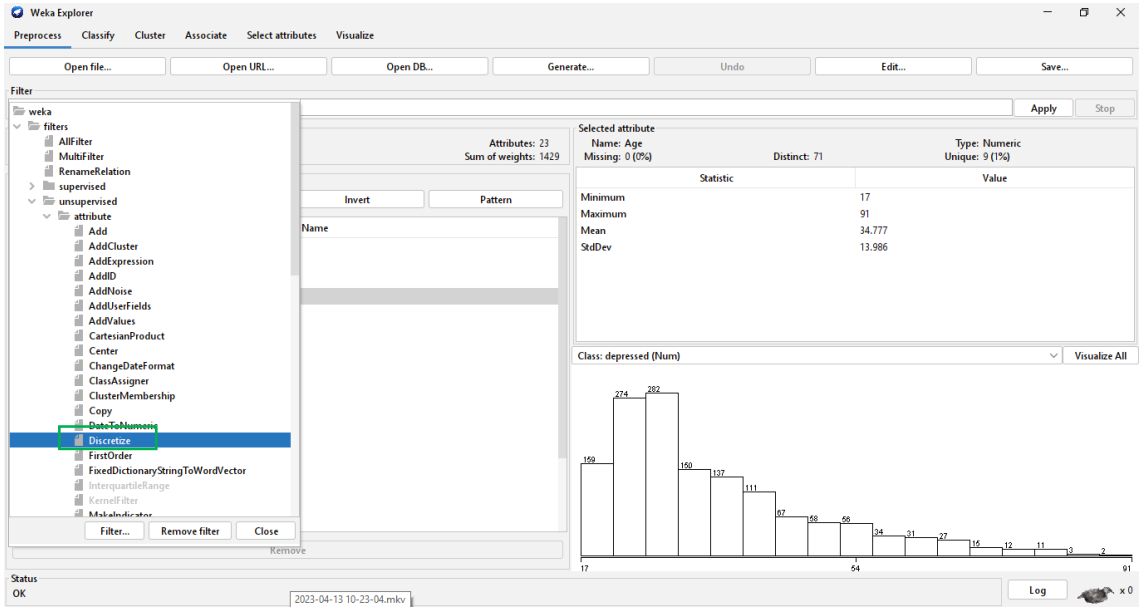
- Para ver los cambios primero veremos los datos iniciales en la columna de AGE (Edad), cuyos datos son de tipo numérico.



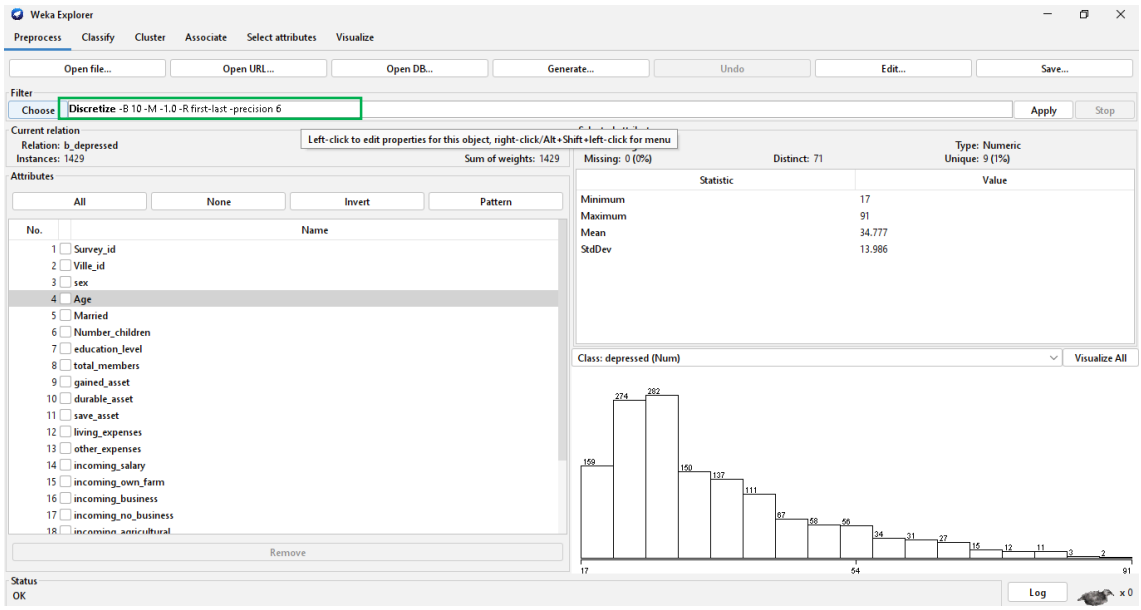
- Para encontrar el filtro *Discretize* debemos ir al botón *Choose* en la parte superior, donde se desplegará un menú en donde seleccionaremos *filters*, para luego seleccionar *unsupervised*, donde seleccionaremos *attribute*.



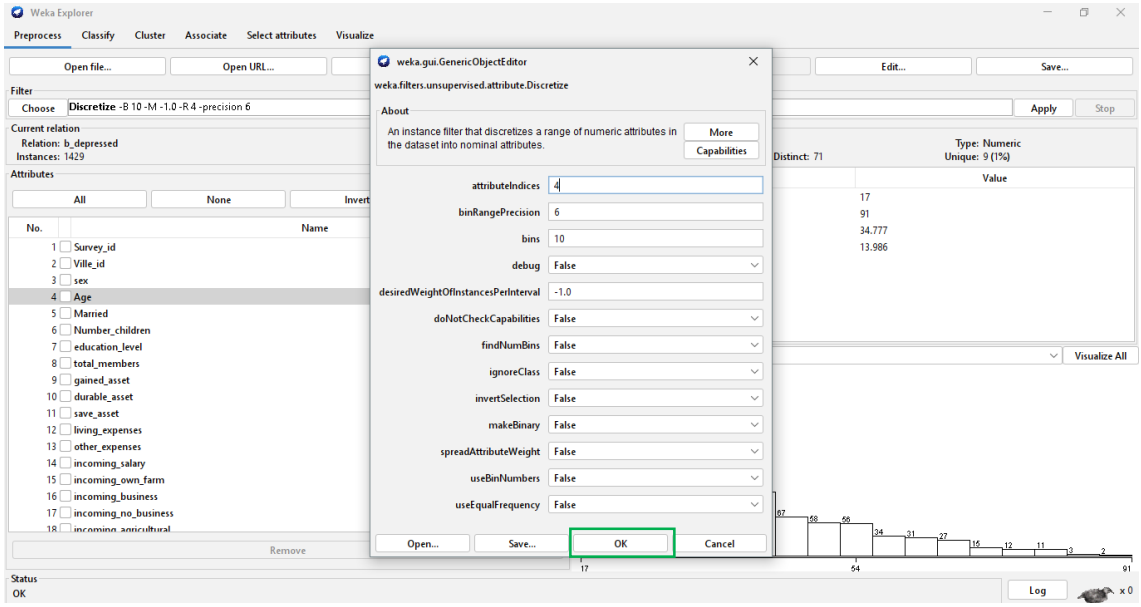
- Una vez seleccionado *attribute*, se desplegara un lista en donde seleccionaremos *Discretize*



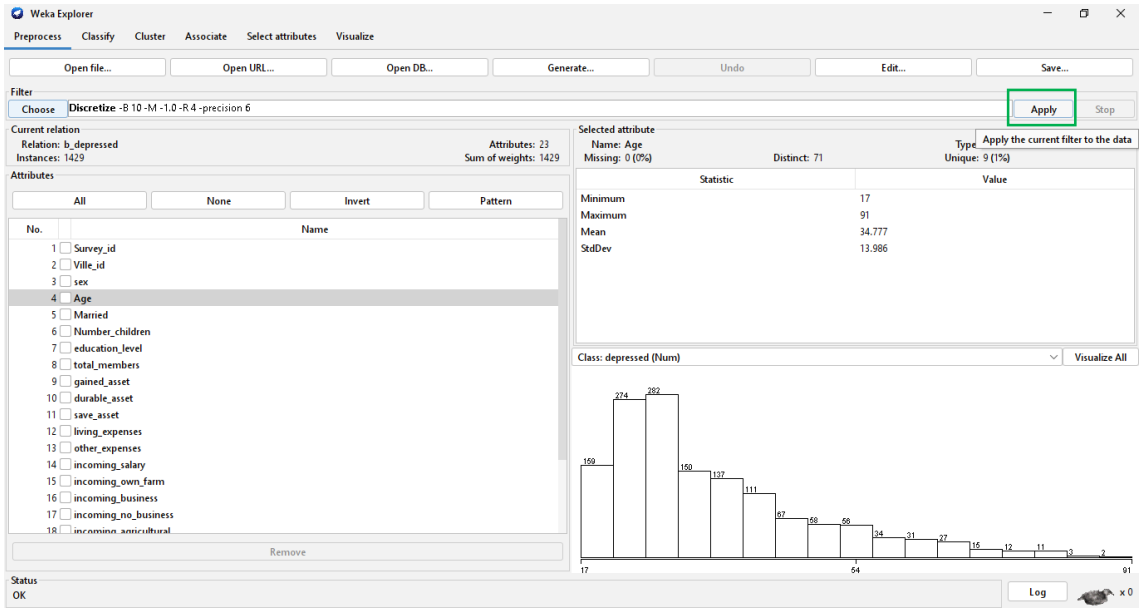
- Hecho eso, en la parte superior se podrá ver que se seleccionó el filtro *Discretize*, en donde haremos click sobre el mismo para configurarlo.



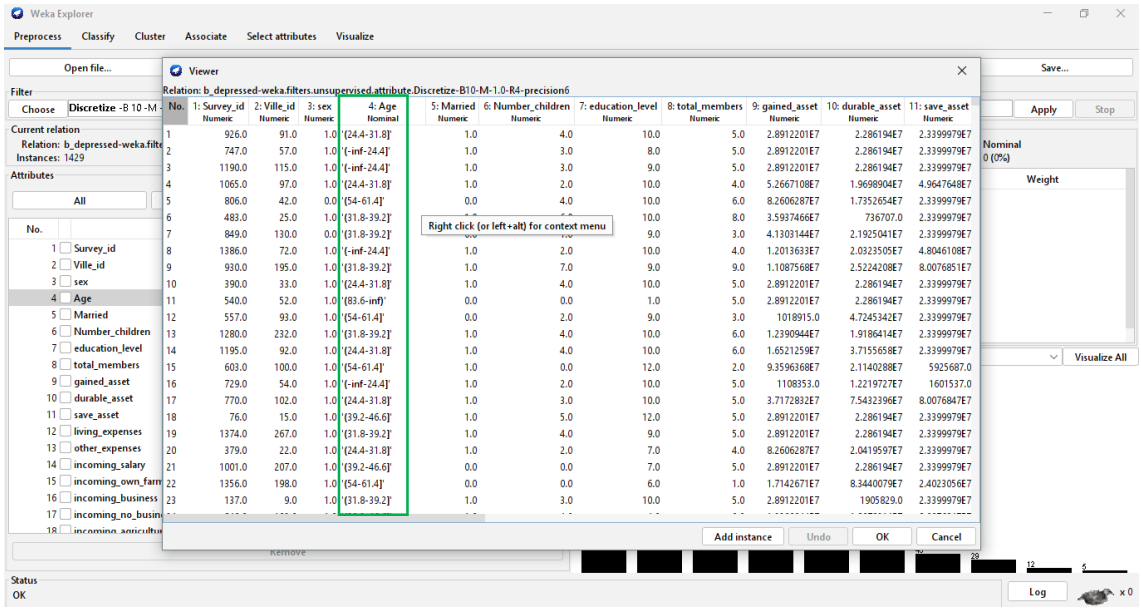
- Se abrirá un menú donde indicaremos la posición del atributo que deseamos modificar, en mi caso será el 4 (Age), con un rango de precisión 6 y la cantidad de contenedores 10. Una vez listo esto se presionan *OK*



- Finalmente lista la anterior configuración se presiona *Apply* para aplicar dicha configuración

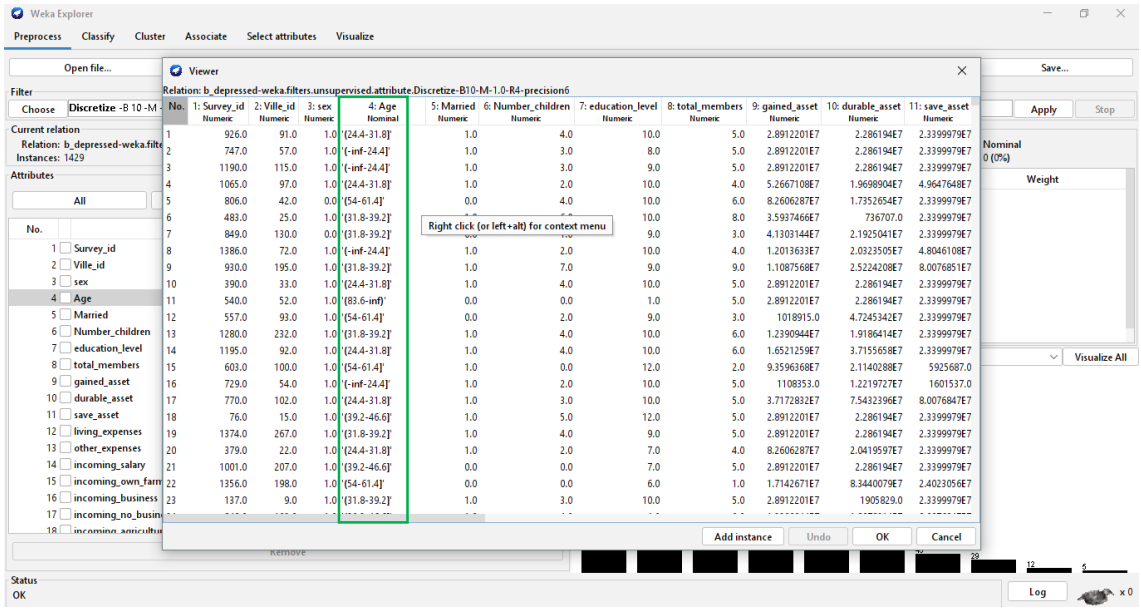


- Y para ver estos cambios podemos ir a Editar, donde la columna Age ahora es de tipo Nominal, donde se han establecido los rango.

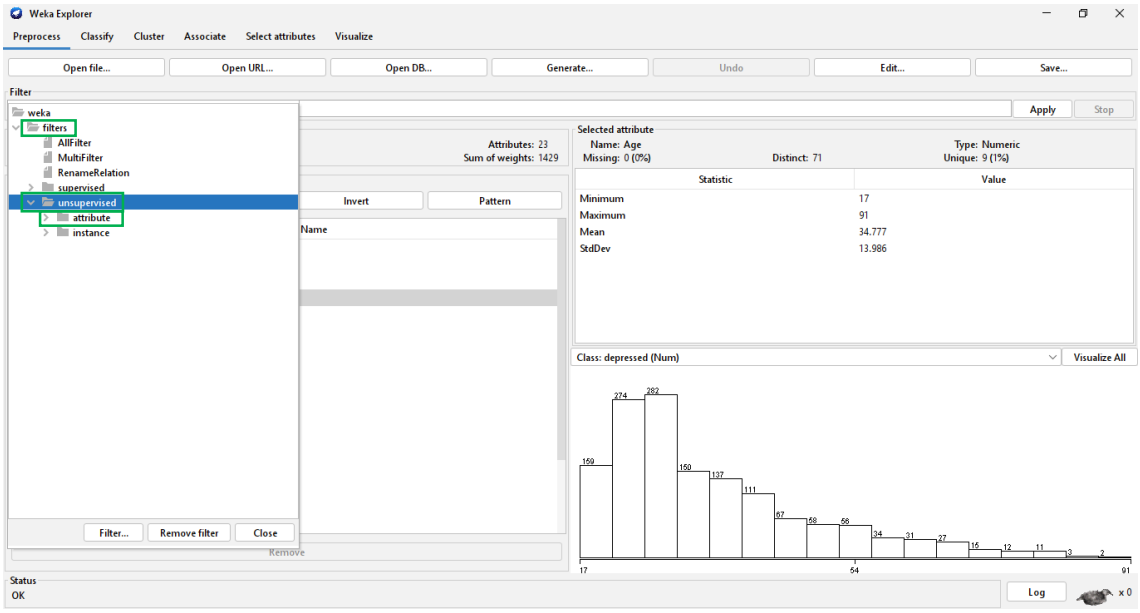


2. AddNoise

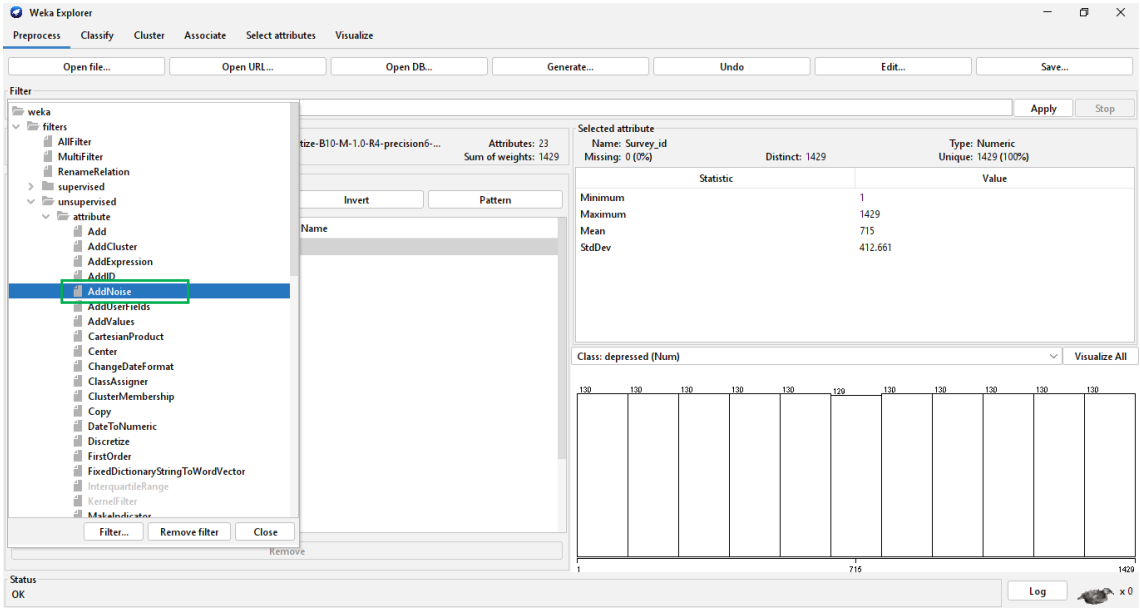
- Para ver los cambios primero veremos los datos iniciales en la columna de AGE (Edad), cuyos datos son de tipo Nominal.



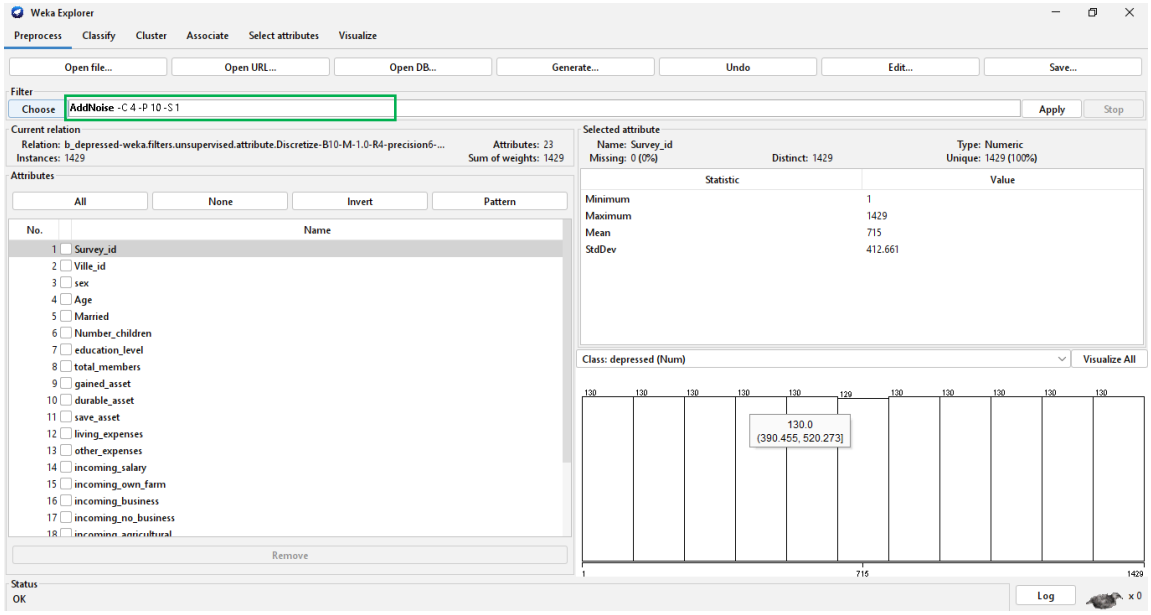
- Para encontrar el filtro *AddNoise* debemos ir al botón *Choose* en la parte superior, donde se desplegará un menú en donde seleccionaremos *filters*, para luego seleccionar *unsupervised*, donde seleccionaremos *attribute*.



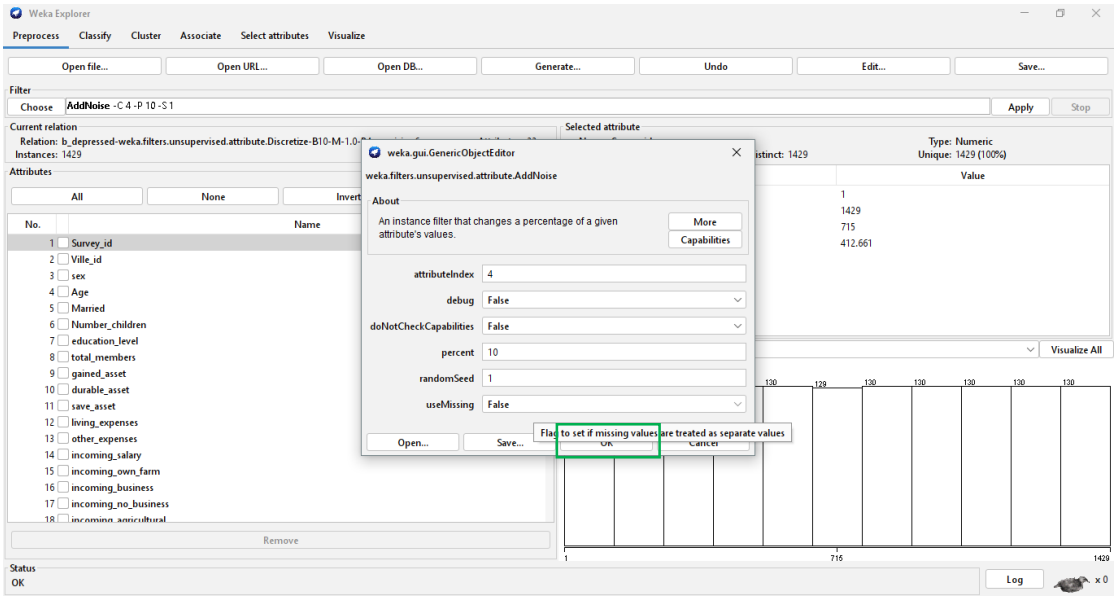
- Una vez seleccionado *attribute*, se desplegara un lista en donde seleccionaremos *AddNoise*



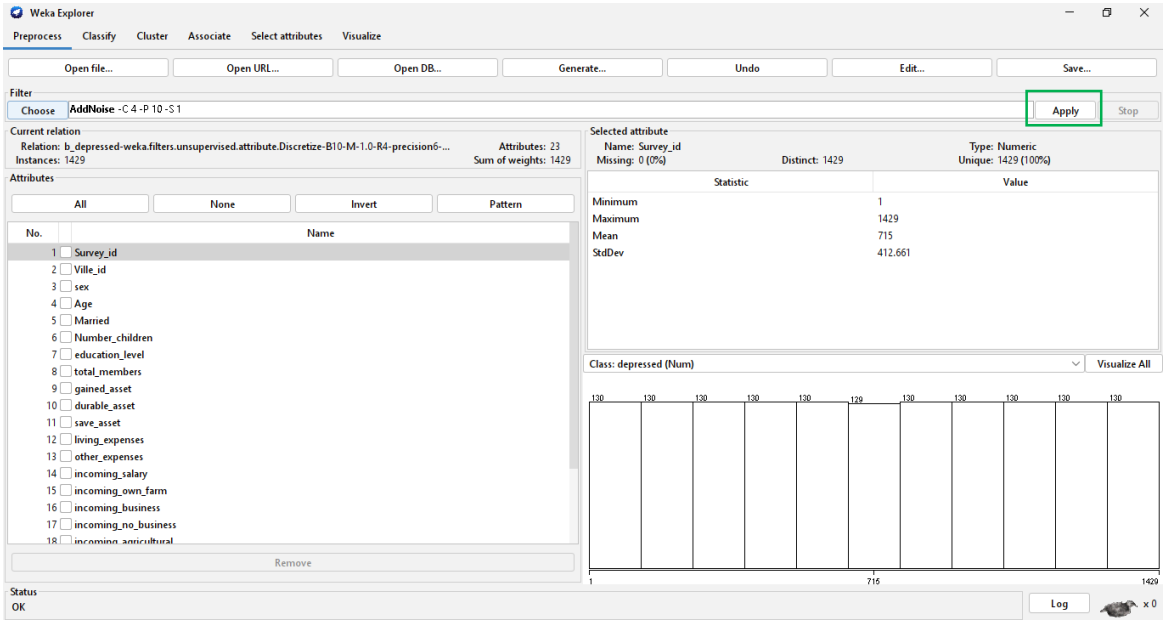
- Hecho eso, en la parte superior se podrá ver que se seleccionó el filtro *AddNoise*, en donde haremos click sobre el mismo para configurarlo.



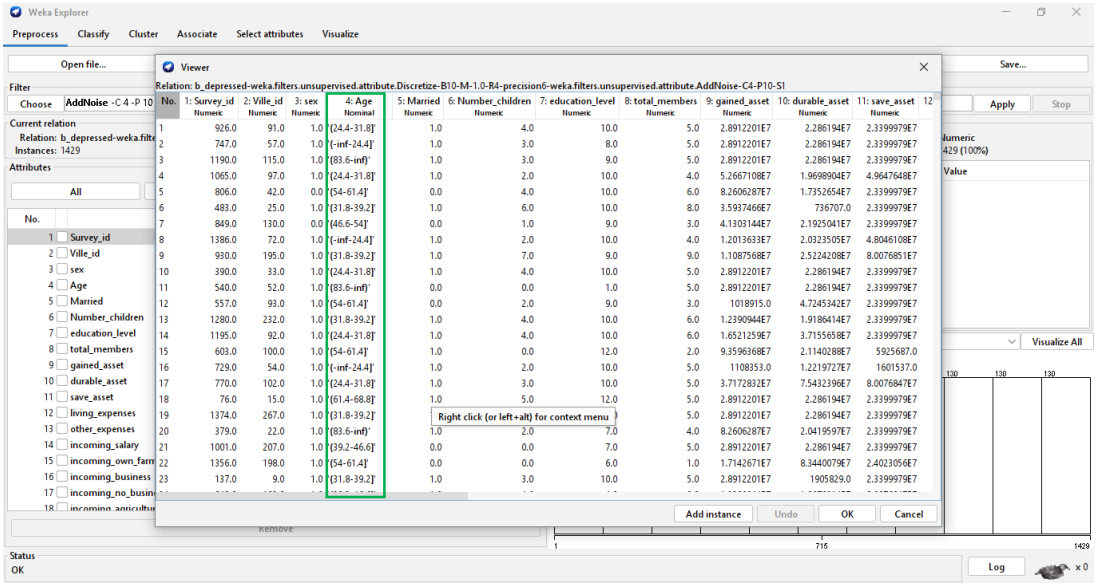
- Se abrirá un menú donde indicaremos la posición del atributo que deseamos modificar, en mi caso será el 4 (Age), con un porcentaje de 10. Una vez listo esto se presionan *OK*



- Finalmente lista la anterior configuración se presiona *Apply* para aplicar dicha configuración

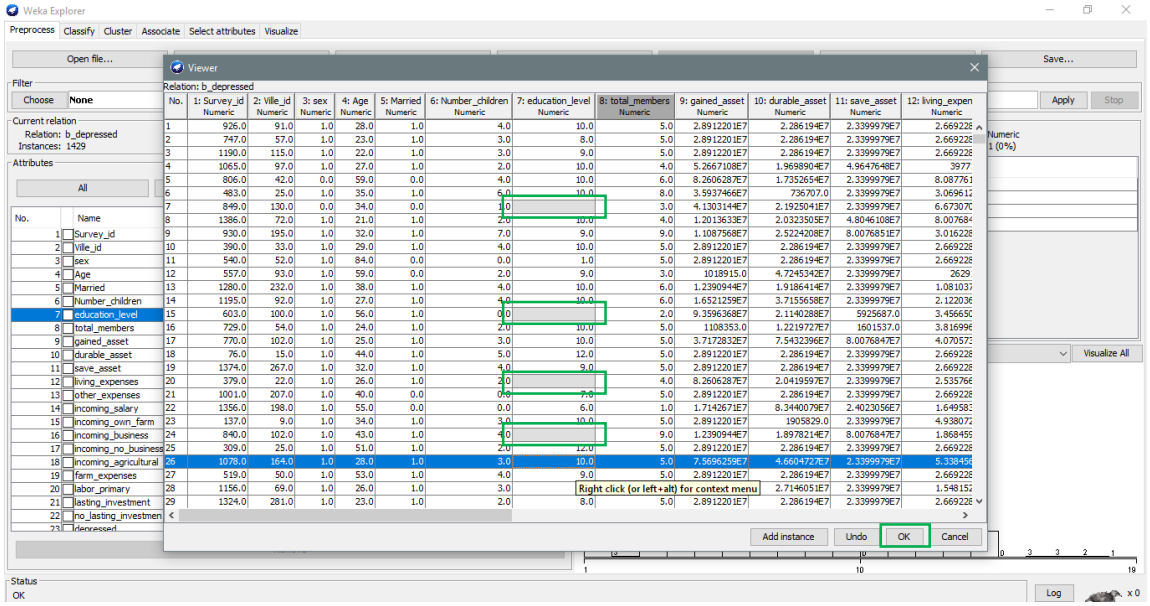


- Y para ver estos cambios podemos ir a Editar, donde en la columna Age, se han cambiado el 10 % de los valores.

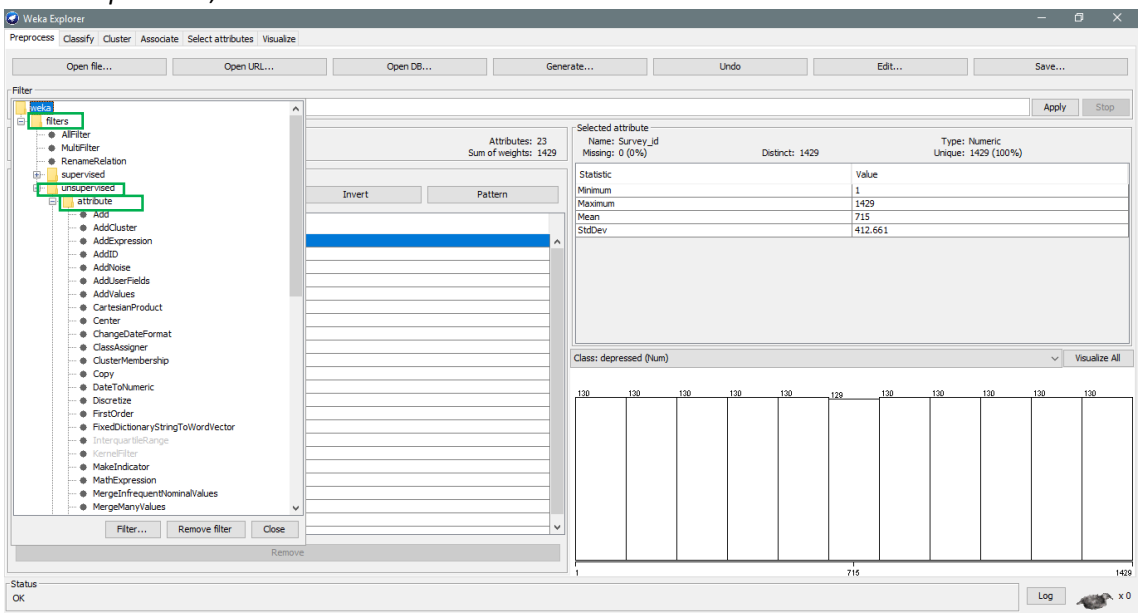


3. ReplaceMissingValues

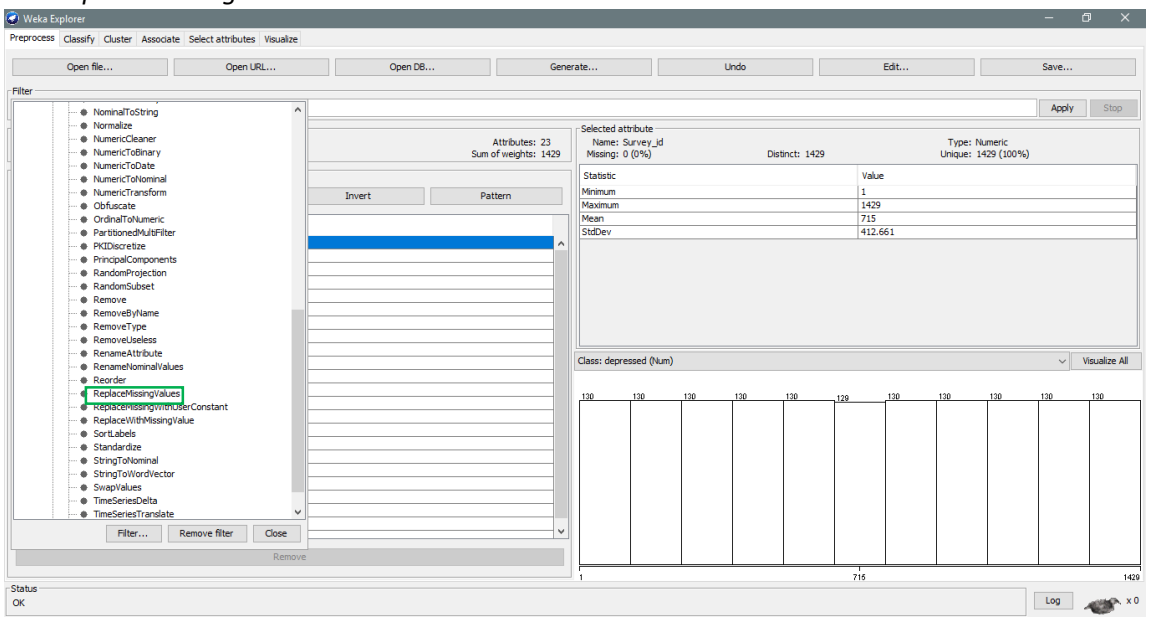
- Para ver los cambios primero veremos los datos iniciales con *Edit* en la columna de *education_level* (Nivel de Educación), en donde previamente se han eliminado 4 datos para luego oprimir *OK*.



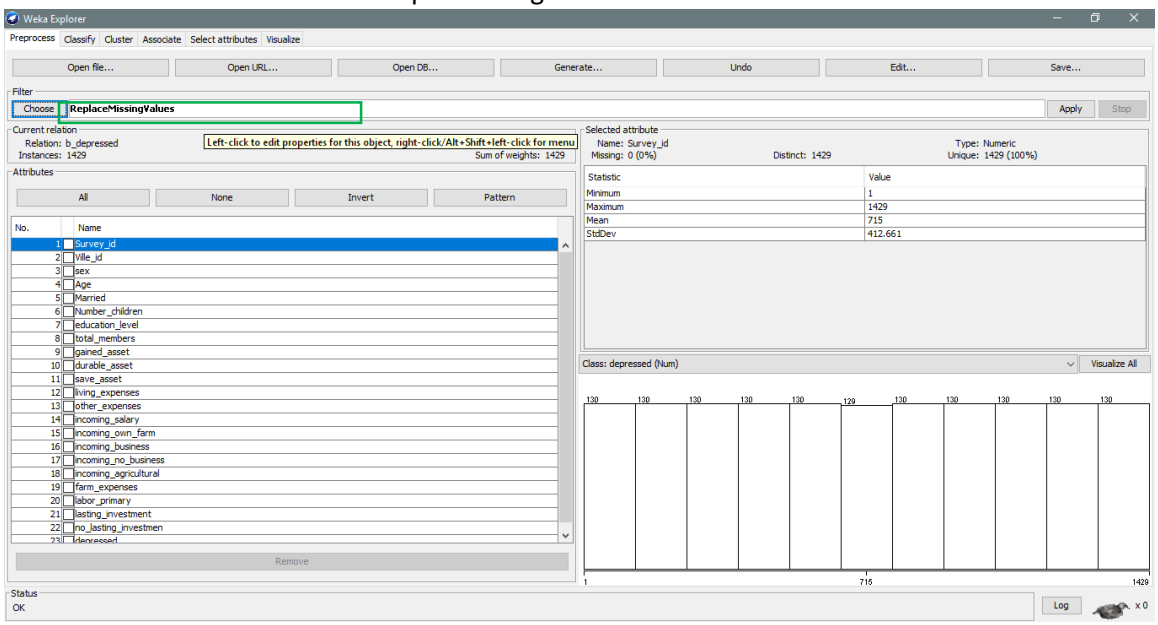
- Para encontrar el filtro *ReplaceMissingValues* debemos ir al botón *Choose* en la parte superior, donde se desplegará un menú en donde seleccionaremos *filters*, para luego seleccionar *unsupervised*, donde seleccionaremos *attribute*.



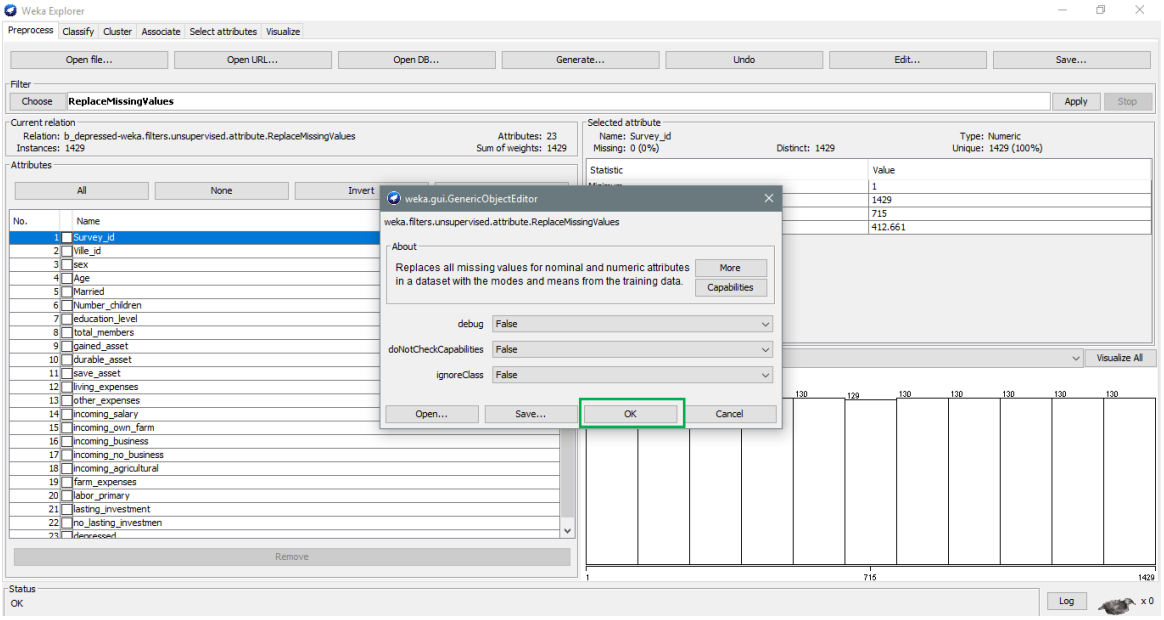
- Una vez seleccionado *attribute*, se desplegara un lista en donde seleccionaremos *ReplaceMissingValues*



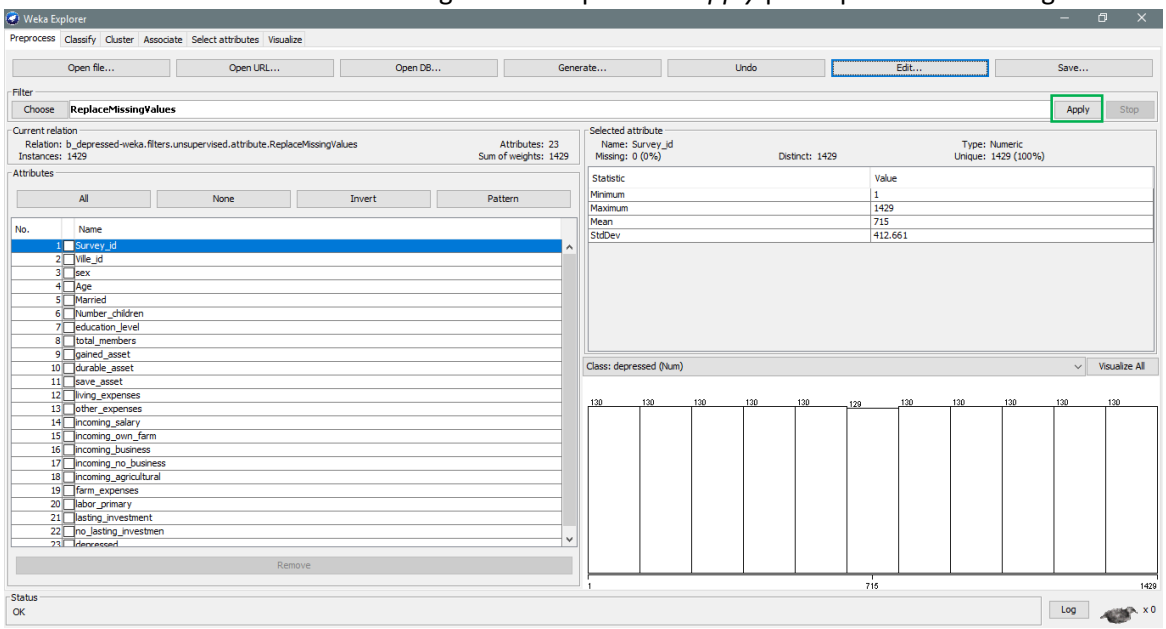
- Hecho eso, en la parte superior se podrá ver que se seleccionó el filtro *Discretize*, en donde haremos click sobre el mismo para configurarlo.



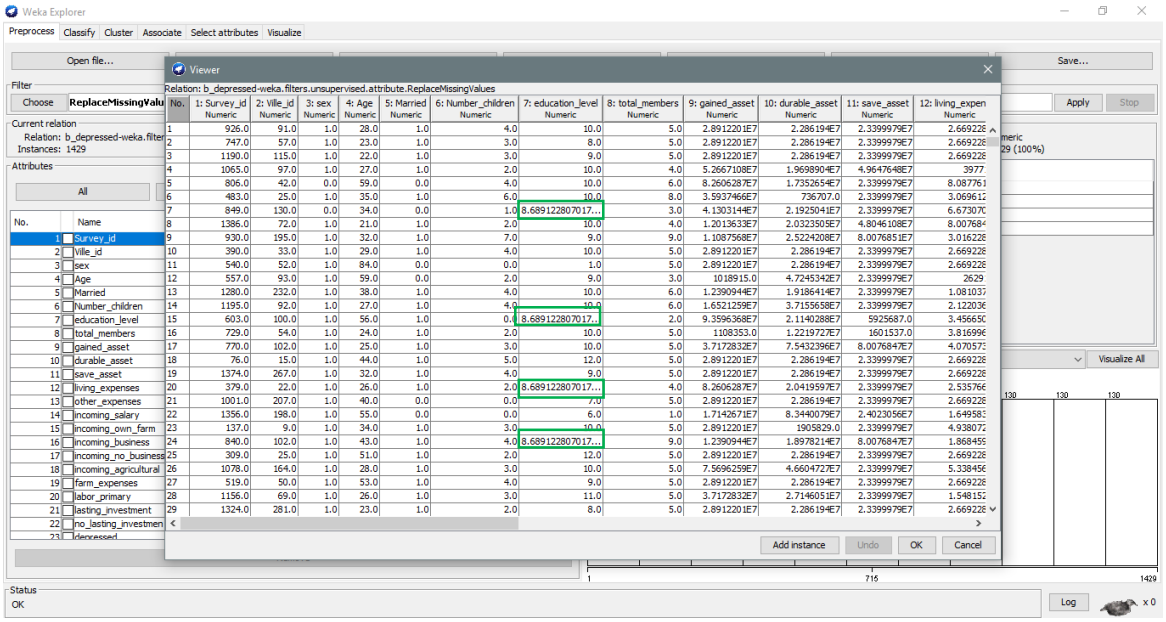
- Se abrirá un menú donde y presionaremos *OK*



- Finalmente lista la anterior configuración se presiona *Apply* para aplicar dicha configuración

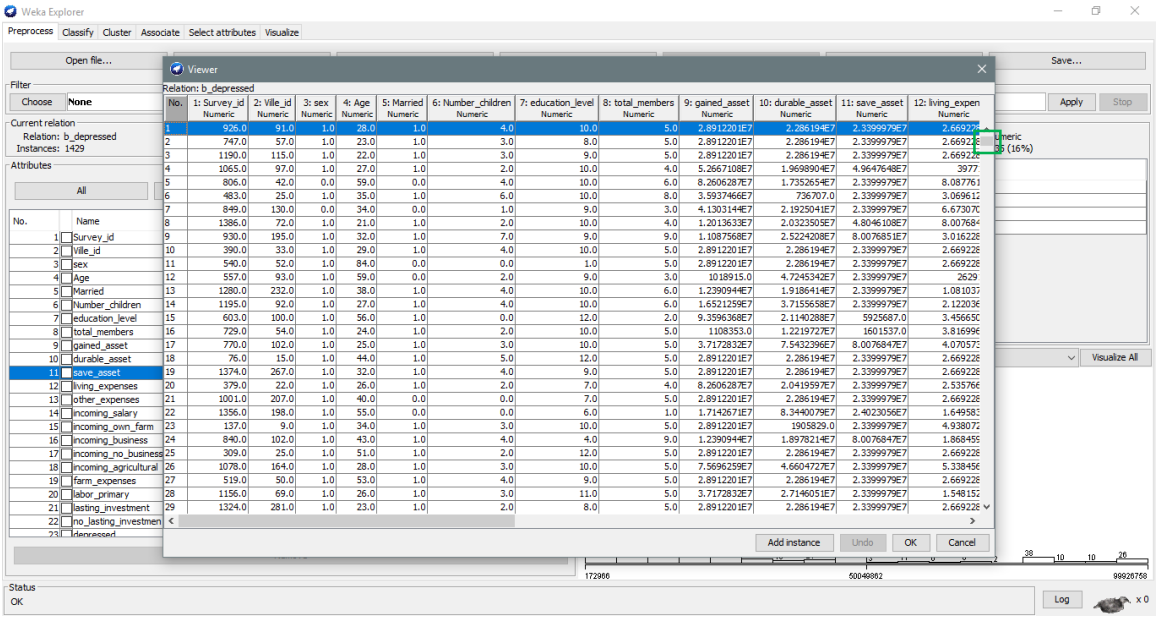


- Y para ver estos cambios podemos ir a Editar, donde en la columna *education_level* ahora hay valores en donde se habían eliminado en un principio.

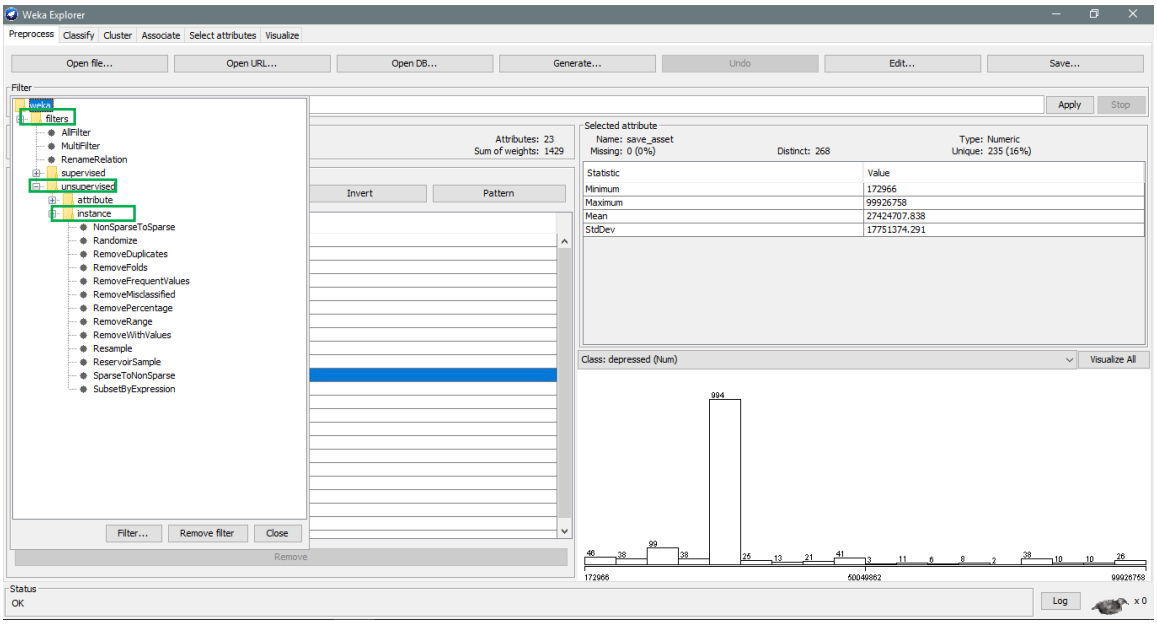


4. Resample

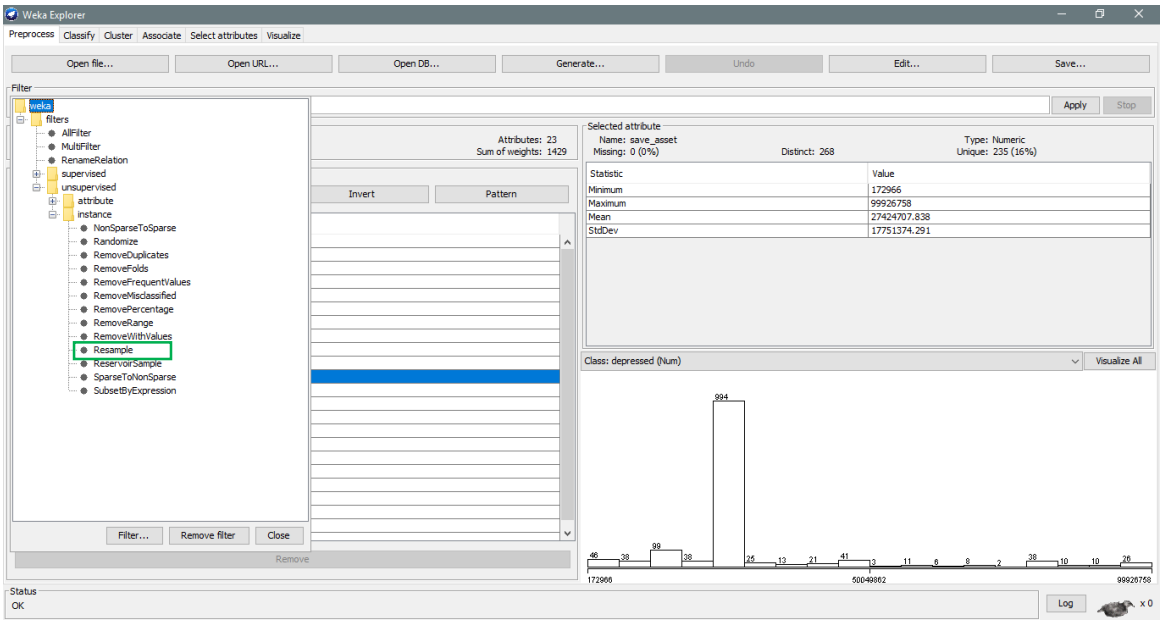
- Para ver los cambios primero veremos los datos iniciales con *Edit* en donde se observa una enorme cantidad de datos



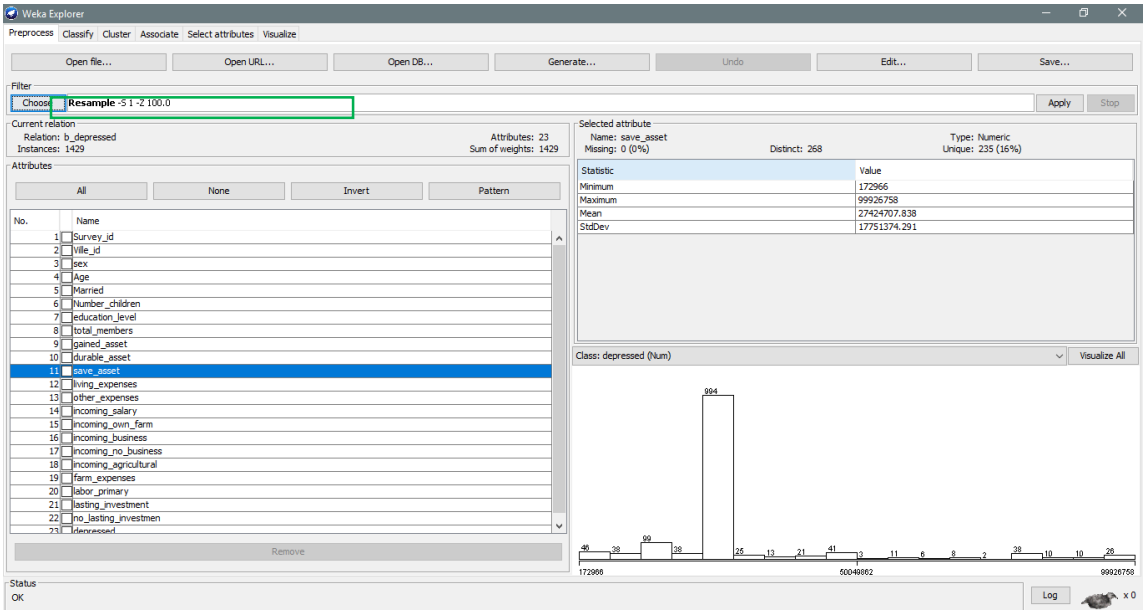
- Para encontrar el filtro *Resample* debemos ir al botón *Choose* en la parte superior, donde se desplegará un menú en donde seleccionaremos *filters*, para luego seleccionar *unsupervised*, donde seleccionaremos *instance*.



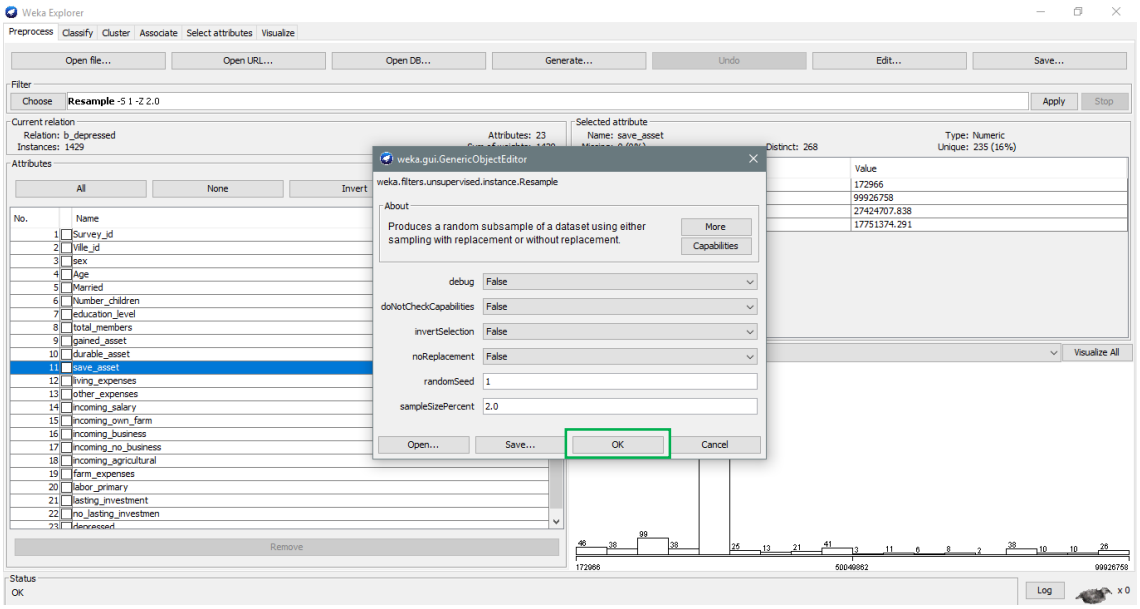
- Una vez seleccionado *instance*, se desplegara un lista en donde seleccionaremos *Resample*



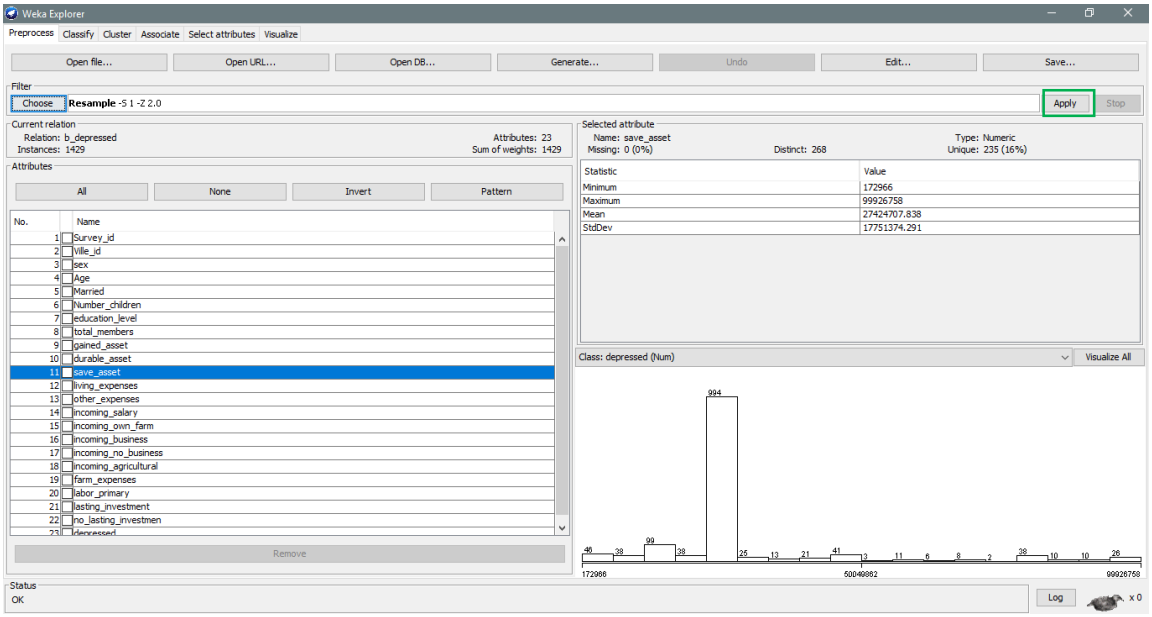
- Hecho eso, en la parte superior se podrá ver que se seleccionó el filtro *Resample*, en donde haremos click sobre el mismo para configurarlo.



- Se abrirá un menú donde pondremos en sampleSizePercent que será el 2% de los datos originales y presionaremos *OK*



- Finalmente lista la anterior configuración se presiona *Apply* para aplicar dicha configuración



- Y para ver estos cambios podemos ir a Editar, donde ahora hay solo 28 valores.