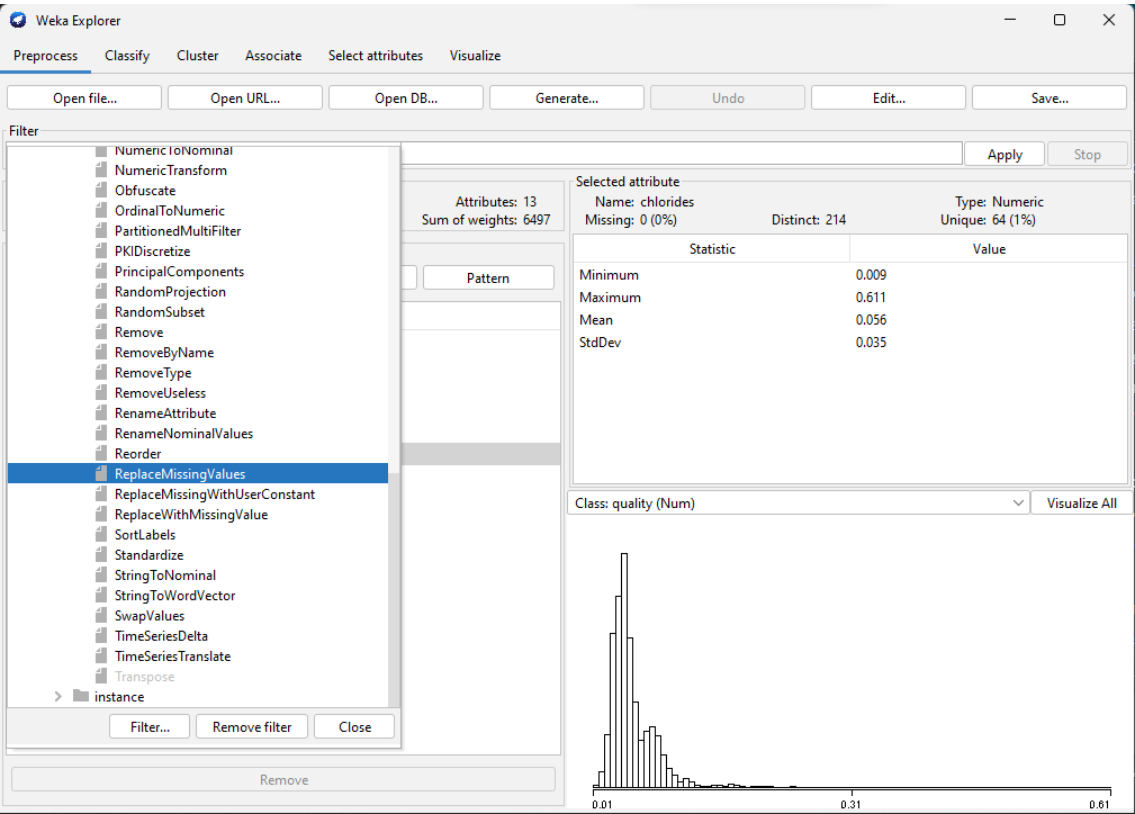
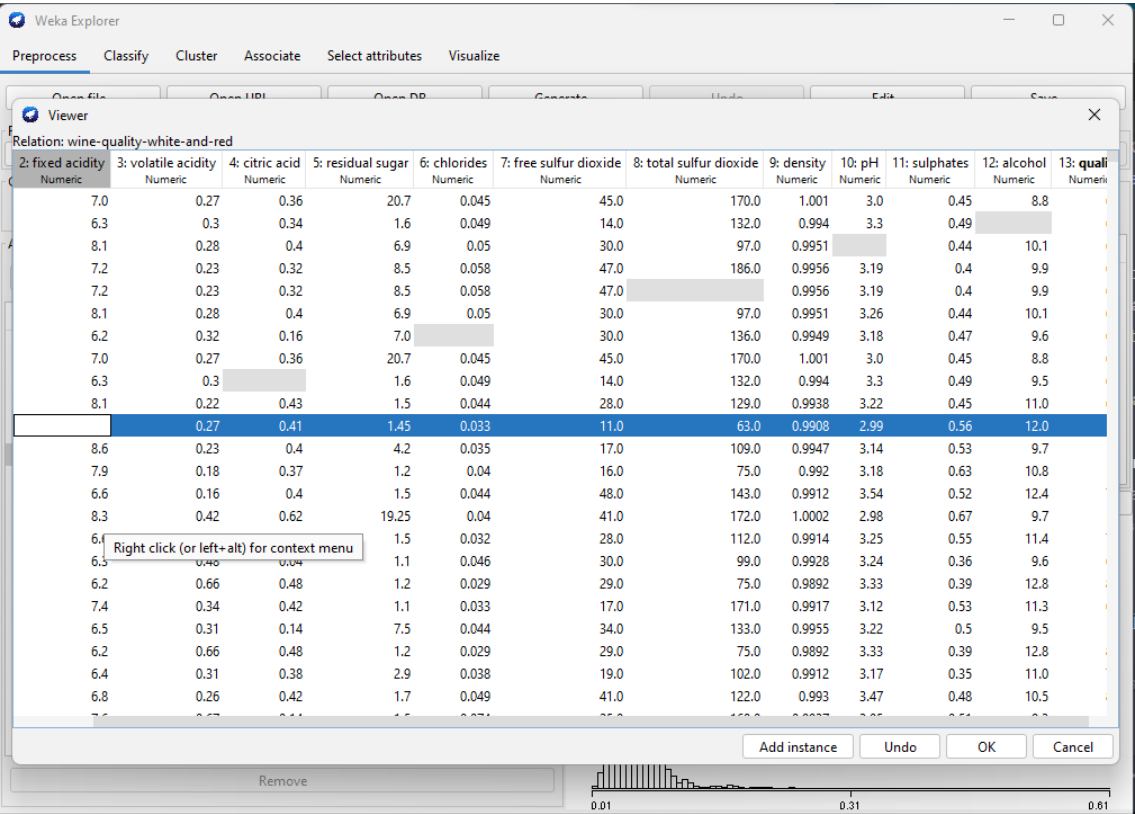
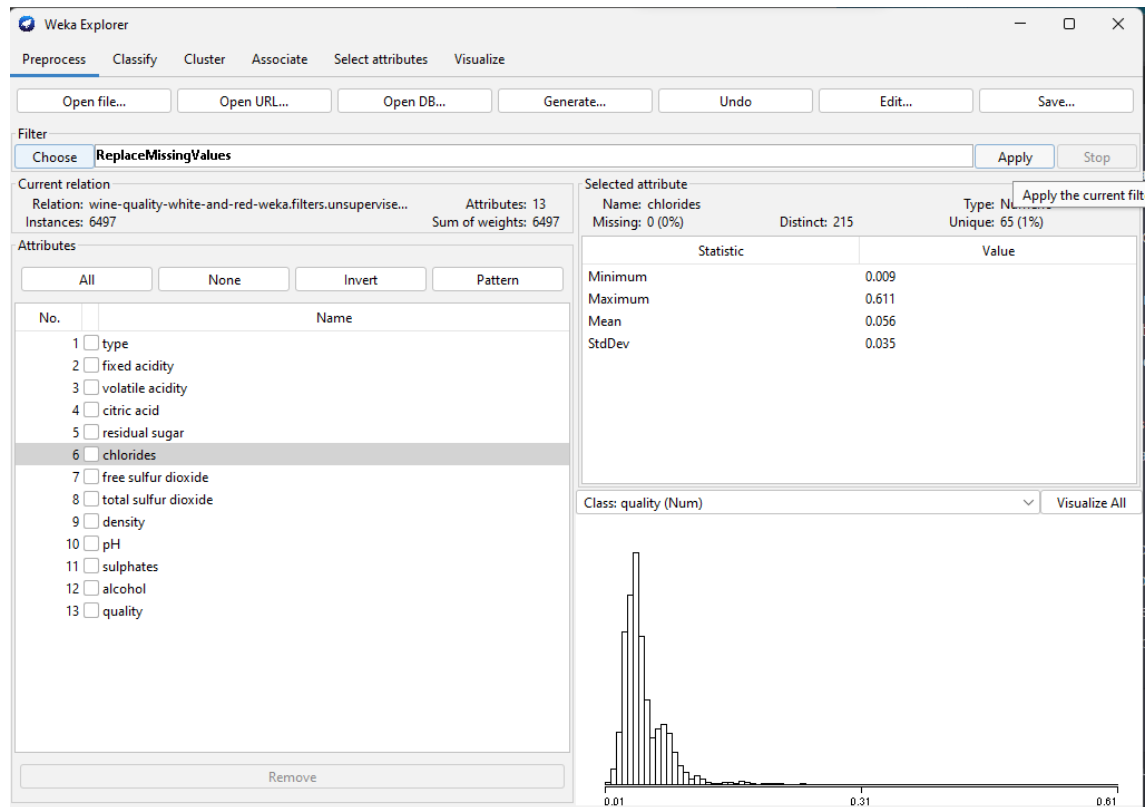


1.- Imputación o Manejo de datos faltantes:

Razón: Los datos faltantes pueden ser un problema en el análisis de datos y el aprendizaje automático, ya que muchos algoritmos no pueden manejar valores faltantes. La imputación de datos faltantes es crucial para mantener la integridad del conjunto de datos y evitar la pérdida de información valiosa.





Weka Explorer

Preprocess

Classify

Cluster

Associate

Select attributes

Visualize

Open File

Open URL

Open DB

Generate

Undo

Edit

Save

Viewer

Relation: wine-quality-white-and-red-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.ReplaceMissingValues

1: name	2: fixed acidity	3: volatile acidity	4: citric acid	5: residual sugar	6: chlorides	7: free sulfur dioxide	8: total sulfur dioxide	9: density	10: pH	11: sulphates	12: alcohol	13: quality
	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric
	7.0	0.27	0.36	20.7	0.045	45.0	170.0	1.001	3.0	0.45	8.8	6.0
	6.3	0.3	0.34	1.6	0.049	14.0	132.0	0.994	3.3	0.49	10.491953...	6.0
	8.1	0.28	0.4	6.9	0.05	30.0	97.0	0.9951	3.2184...	0.44	10.1	6.0
	7.2	0.23	0.32	8.5	0.058	47.0	186.0	0.9956	3.19	0.4	9.9	6.0
	7.2	0.23	0.32	8.5	0.058	47.0	115.7337592364532	0.9956	3.19	0.4	9.9	6.0
	8.1	0.28	0.4	6.9	0.05	30.0	97.0	0.9951	3.26	0.44	10.1	6.0
	6.2	0.32	0.16	7.0	0.0560355...	30.0	136.0	0.9949	3.18	0.47	9.6	6.0
	7.0	0.27	0.36	20.7	0.045	45.0	170.0	1.001	3.0	0.45	8.8	6.0
	6.3	0.3	0.31862992...	1.6	0.049	14.0	132.0	0.994	3.3	0.49	9.5	6.0
	8.1	0.22	0.43	1.5	0.044	28.0	129.0	0.9938	3.22	0.45	11.0	6.0
7.2151708743...		0.27	0.41	1.45	0.033	11.0	63.0	0.9908	2.99	0.56	12.0	5.0
	8.6	0.23	0.4	4.2	0.035	17.0	109.0	0.9947	3.14	0.53	9.7	5.0
	7.9	0.18	0.37	1.2	0.04	16.0	75.0	0.992	3.18	0.63	10.8	5.0
	6.6	0.16	0.4	1.5	0.044	48.0	143.0	0.9912	3.54	0.52	12.4	7.0
	8.3	0.42	0.62	19.25	0.04	41.0	172.0	1.0002	2.98	0.67	9.7	5.0
	6.6	0.17	0.38	1.5	0.032	28.0	112.0	0.9914	3.25	0.55	11.4	7.0
	6.3	0.48	0.04	1.1	0.046	30.0	99.0	0.9928	3.24	0.36	9.6	6.0
	6.2	0.66	0.48	1.2	0.029	29.0	75.0	0.9892	3.33	0.39	12.8	8.0
	7.4	0.34	0.42	1.1	0.033	17.0	171.0	0.9917	3.12	0.53	11.3	6.0
	6.5	0.31	0.14	7.5	0.044	34.0	133.0	0.9955	3.22	0.5	9.5	5.0
	6.2	0.66	0.48	1.2	0.029	29.0	75.0	0.9892	3.33	0.39	12.8	8.0
	6.4	0.31	0.38	2.9	0.038	19.0	102.0	0.9912	3.17	0.35	11.0	7.0
	6.8	0.26	0.42	1.7	0.049	41.0	122.0	0.993	3.47	0.48	10.5	8.0

Add instance

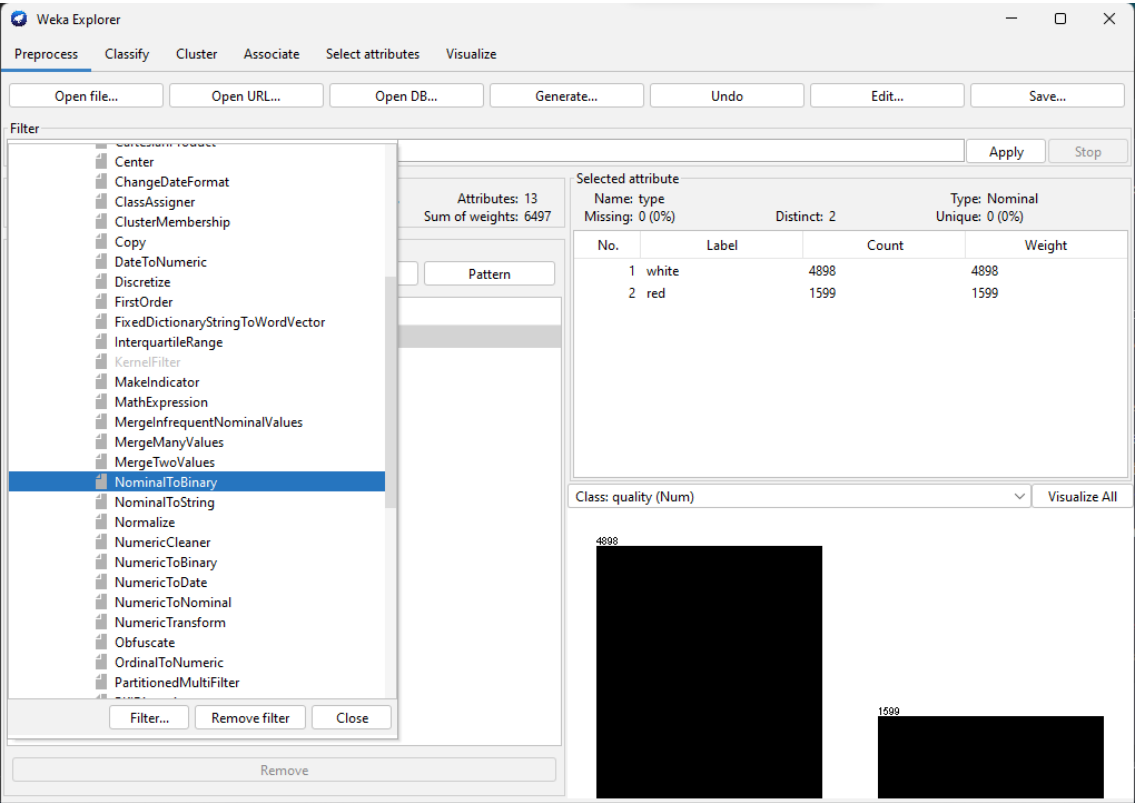
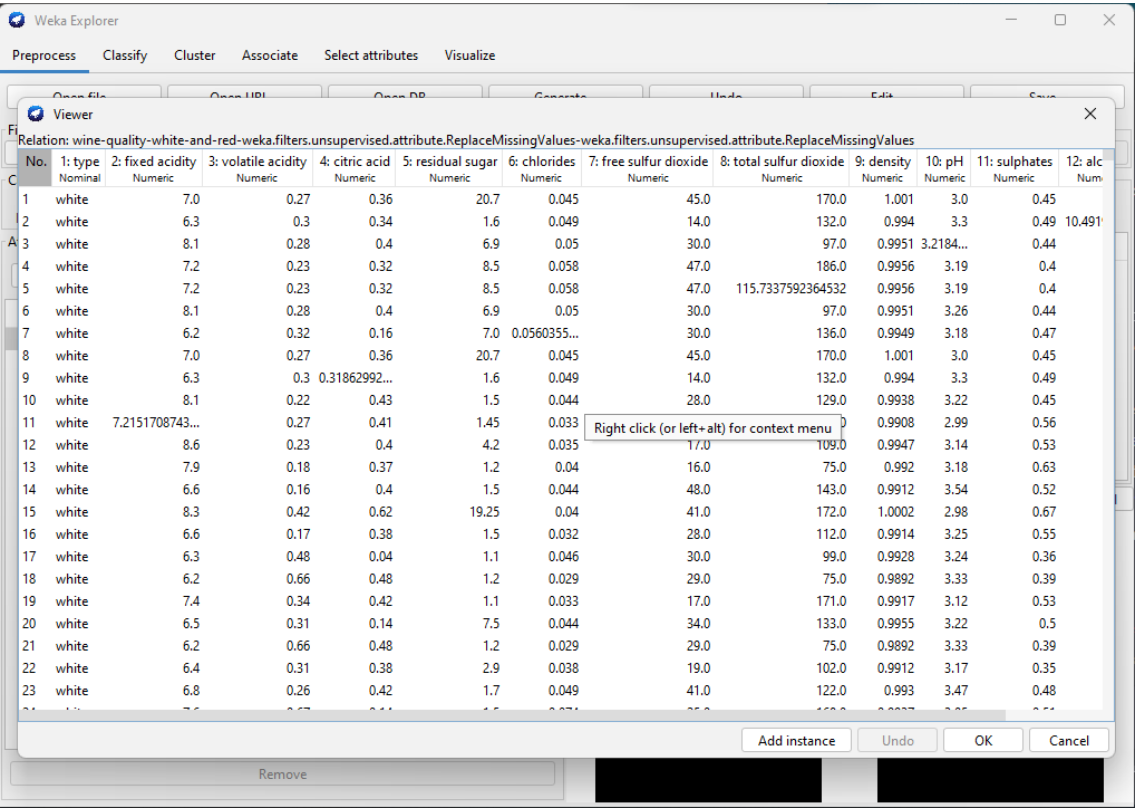
Undo

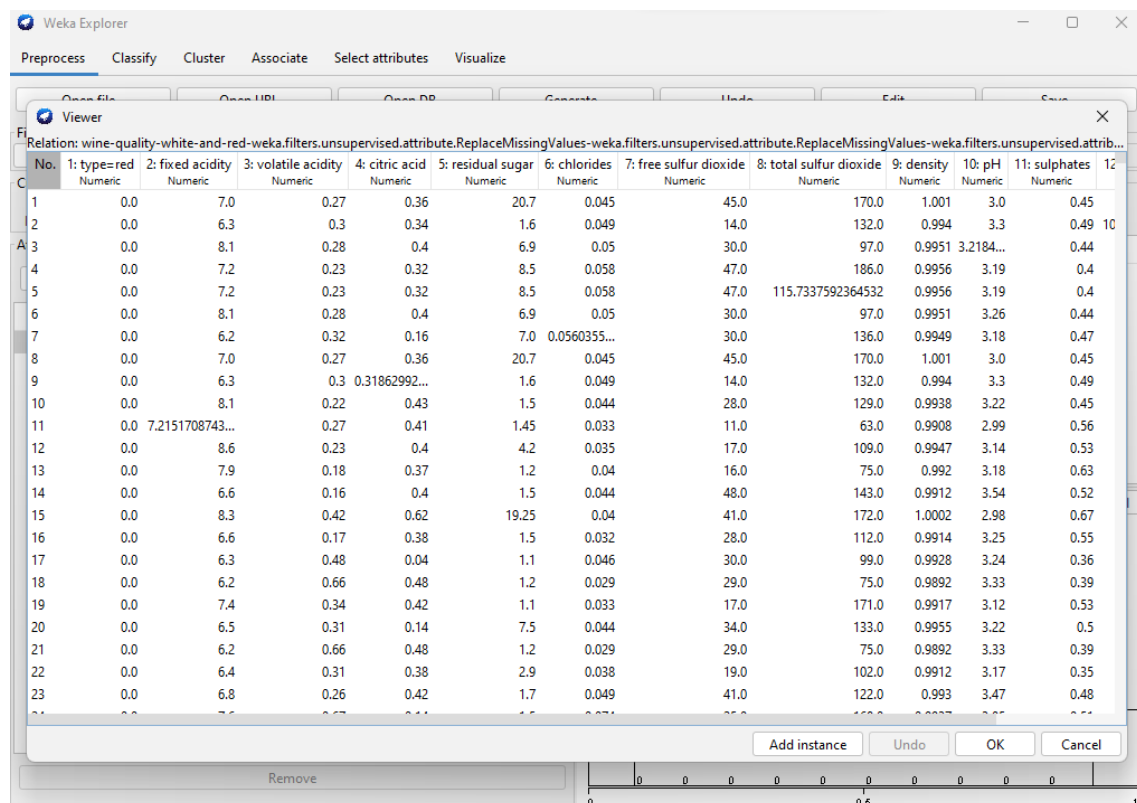
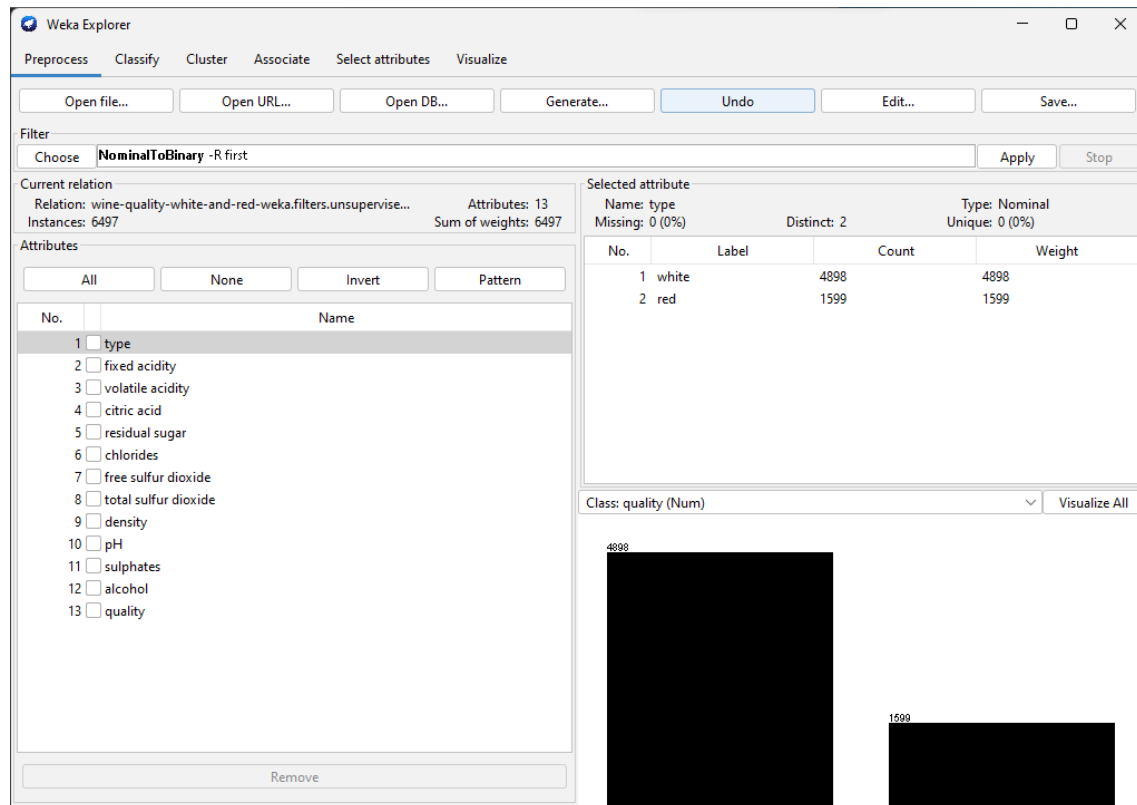
OK

Cancel

2.- Codificación de variables categóricas:

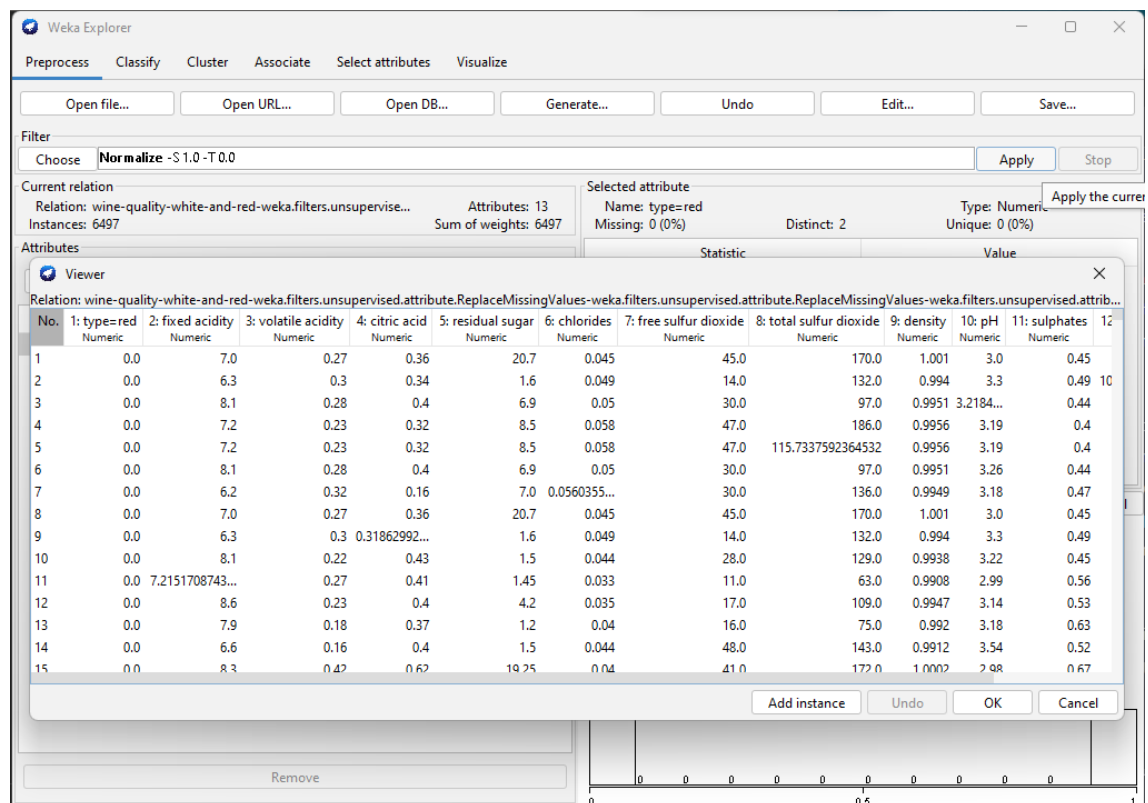
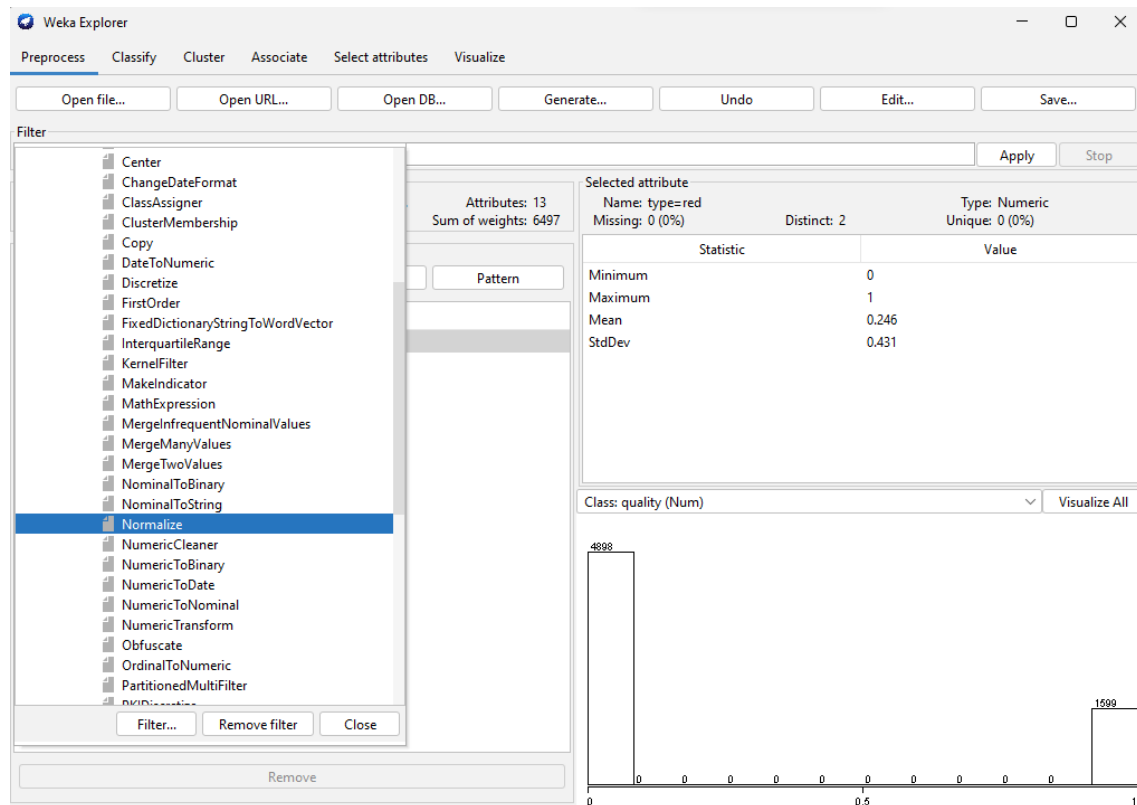
Razón: Si el conjunto de datos contiene variables categóricas (como "wine_type"), es necesario convertirlas en variables numéricas para que los algoritmos de aprendizaje automático puedan trabajar con ellas. Esto generalmente se hace mediante técnicas de codificación como la codificación one-hot o la codificación ordinal.





3.- Normalización o estandarización de características

Razón: Algunos algoritmos de aprendizaje automático son sensibles a la escala de las características. La normalización (escalamiento a un rango específico) o la estandarización (escalamiento con media cero y desviación estándar uno) de las características pueden ayudar a los modelos a converger más rápido y obtener un mejor rendimiento.



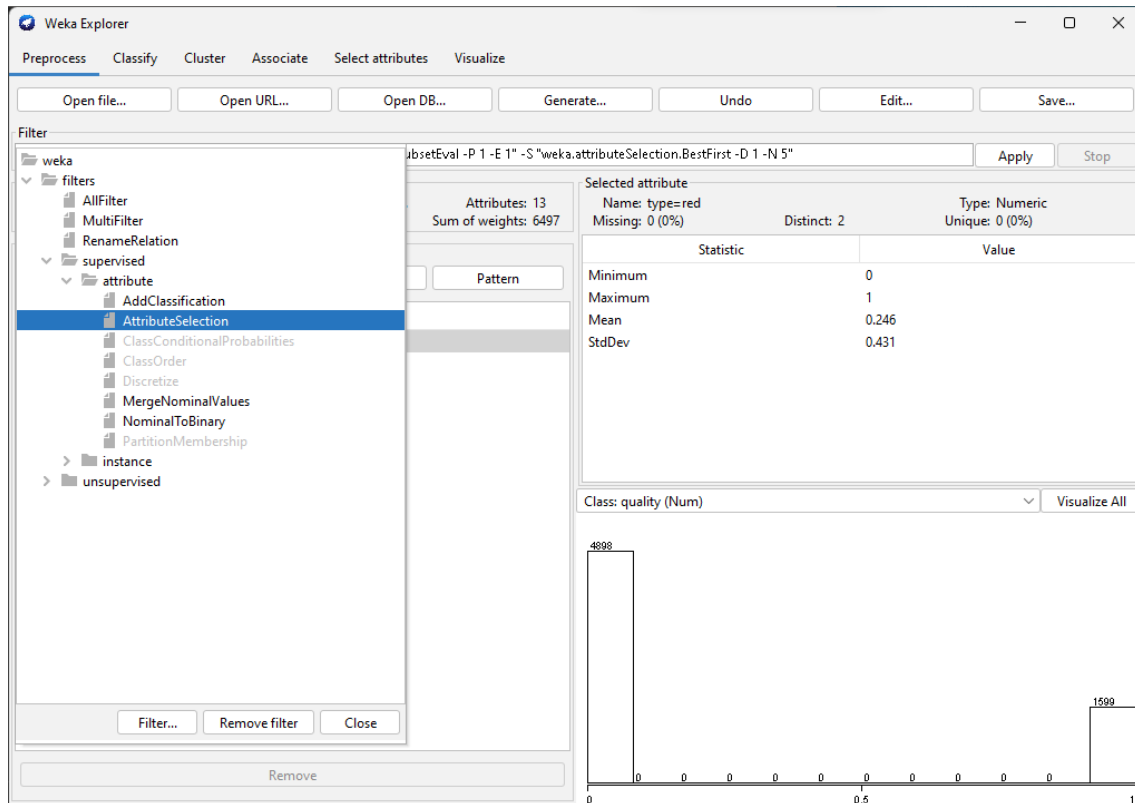
- S:** Este valor representa el valor de escala superior. En otras palabras, es el valor máximo al que se normalizarán las características. Por ejemplo, si estableces **-S 1.0**, significa que las características se normalizarán de modo que su valor máximo sea 1.0.

- **-T:** Este valor representa el valor de escala inferior. Es el valor mínimo al que se normalizarán las características. Si estableces **-T 0.0**, significa que las características se normalizarán de modo que su valor mínimo sea 0.0.

No.	1: type=red	2: fixed acidity	3: volatile acidity	4: citric acid	5: residual sugar	6: chlorides	7: free sulfur dioxide	8: total sulfur dioxide	9: density	10: pH	11: sulphates	12
1	0.0	0.2644628099...	0.126666666666...	0.21686746...	0.308282208588...	0.0598006...	0.1527777777777778	0.3778801843317972	0.267784...	0.2170...	0.129213483...	0.
2	0.0	0.2066115702...	0.146666666666...	0.20481927...	0.015337423312...	0.0664451...	0.0451388888888889	0.2903225806451613	0.132832...	0.4496...	0.151685393...	0.
3	0.0	0.3553719008...	0.133333333333...	0.24096385...	0.096625766871...	0.0681063...	0.1006944444444445	0.20967741935483872	0.154038...	0.3864...	0.123595505...	0.
4	0.0	0.2809917355...	0.100000000000...	0.19277108...	0.121165644171...	0.0813953...	0.1597222222222222	0.4147465437788018	0.163678...	0.3643...	0.101123595...	0.
5	0.0	0.2809917355...	0.100000000000...	0.19277108...	0.121165644171...	0.0813953...	0.1597222222222222	0.2528427632176341	0.163678...	0.3643...	0.101123595...	0.
6	0.0	0.3553719008...	0.133333333333...	0.24096385...	0.096625766871...	0.0681063...	0.1006944444444445	0.20967741935483872	0.154038...	0.4186...	0.123595505...	0.
7	0.0	0.1983471074...	0.16	0.09638554...	0.098159509202...	0.0781321...	0.1006944444444445	0.2995391705069124	0.150183...	0.3565...	0.140449438...	0.
8	0.0	0.2644628099...	0.126666666666...	0.21686746...	0.308282208588...	0.0598006...	0.1527777777777778	0.3778801843317972	0.267784...	0.2170...	0.129213483...	0.
9	0.0	0.2066115702...	0.146666666666...	0.19194573...	0.015337423312...	0.0664451...	0.0451388888888889	0.2903225806451613	0.132832...	0.4496...	0.151685393...	0.
10	0.0	0.3553719008...	0.093333333333...	0.25903614...	0.013803680981...	0.0581395...	0.09375	0.2834101382488479	0.128976...	0.3875...	0.129213483...	0.
11	0.0	0.2822455268...	0.126666666666...	0.24698795...	0.013036809815...	0.0398671...	0.0347222222222222	0.1313364055299539	0.071139...	0.2093...	0.191011235...	0.
12	0.0	0.3966942148...	0.100000000000...	0.24096385...	0.055214723926...	0.0431893...	0.0555555555555555	0.23732718894009217	0.146327...	0.3255...	0.174157303...	0.
13	0.0	0.3388429752...	0.066666666666...	0.22289156...	0.009202453987...	0.0514950...	0.0520833333333333	0.15898617511520738	0.094274...	0.3565...	0.230337078...	0.
14	0.0	0.2314049586...	0.053333333333...	0.24096385...	0.013803680981...	0.0581395...	0.1631944444444445	0.315668202764977	0.078850...	0.6356...	0.168539325...	0.
15	0.0	0.3719008264...	0.226666666666...	0.37349397...	0.286042944785...	0.0514950...	0.1388888888888889	0.3824884792626728	0.252361...	0.2015...	0.252808988...	0.
16	0.0	0.2314049586...	0.060000000000...	0.22891566...	0.013803680981...	0.0382059...	0.09375	0.24423963133640553	0.082706...	0.4108...	0.185393258...	0.
17	0.0	0.2066115702...	0.266666666666...	0.02409638...	0.007668711656...	0.0614617...	0.1006944444444445	0.21428571428571427	0.109697...	0.4031...	0.078651685...	0.
18	0.0	0.1983471074...	0.386666666666...	0.28915662...	0.009202453987...	0.0332225...	0.0972222222222222	0.15898617511520738	0.040293...	0.4728...	0.095505617...	0.
19	0.0	0.2975206611...	0.173333333333...	0.25301204...	0.007668711656...	0.0398671...	0.0555555555555555	0.38018433179723504	0.088490...	0.3100...	0.174157303...	0.
20	0.0	0.2231404958...	0.153333333333...	0.08433734...	0.105828220858...	0.0581395...	0.1145833333333333	0.2926267281105991	0.161750...	0.3875...	0.157303370...	0.
21	0.0	0.1983471074...	0.386666666666...	0.28915662...	0.009202453987...	0.0332225...	0.0972222222222222	0.15898617511520738	0.040293...	0.4728...	0.095505617...	0.
22	0.0	0.2148760330...	0.153333333333...	0.22891566...	0.035276073619...	0.0481727...	0.0625	0.22119815668202766	0.078850...	0.3488...	0.073033707...	0.
23	0.0	0.2479338842...	0.12	0.25301204...	0.016871165644...	0.0664451...	0.1388888888888889	0.2672811059907834	0.113553...	0.5813...	0.146067415...	0.
24	0.0	0.3140495867...	0.393333333333...	0.08433734...	0.013803680981...	0.1079734...	0.0833333333333333	0.37327188940092165	0.127048...	0.2558...	0.162921348...	0.
25	0.0	0.2314049586...	0.126666666666...	0.24698795...	0.010736196319...	0.0714285...	0.0520833333333333	0.31336405529953915	0.154038...	0.5426...	0.140449438...	0.

4.- Selección de características

Razón: Puede haber características en el conjunto de datos que no aporten información significativa para el problema de clasificación. La selección de características es una técnica que ayuda a identificar y eliminar características irrelevantes o redundantes, lo que puede mejorar la eficiencia y el rendimiento de los modelos.



- **-E "weka.attributeSelection.CfsSubsetEval -P 1 -E 1"**: Esta parte del comando especifica el evaluador de características que se utilizará. En este caso, se utiliza **weka.attributeSelection.CfsSubsetEval**, que es un evaluador que implementa el algoritmo "Correlation-based Feature Selection" (Selección de características basada en correlación). Los argumentos **-P 1 -E 1** son opciones específicas para este evaluador y controlan su comportamiento. Por ejemplo, **-P 1** indica que se utilizará la búsqueda de características con punto flotante y **-E 1** indica que se utilizará el método de evaluación de bondad de características por defecto.
- **-S "weka.attributeSelection.BestFirst -D 1 -N 5"**: Esta parte del comando especifica el método de búsqueda de características que se utilizará. En este caso, se utiliza **weka.attributeSelection.BestFirst**, que es un método de búsqueda que busca el mejor subconjunto de características mediante una estrategia de búsqueda en el espacio de características. Los argumentos **-D 1 -N 5** son opciones específicas para este método de búsqueda y controlan su comportamiento. Por ejemplo, **-D 1** indica que se realizará una búsqueda hacia delante (forward) y **-N 5** indica que se seleccionarán como máximo 5 características.

