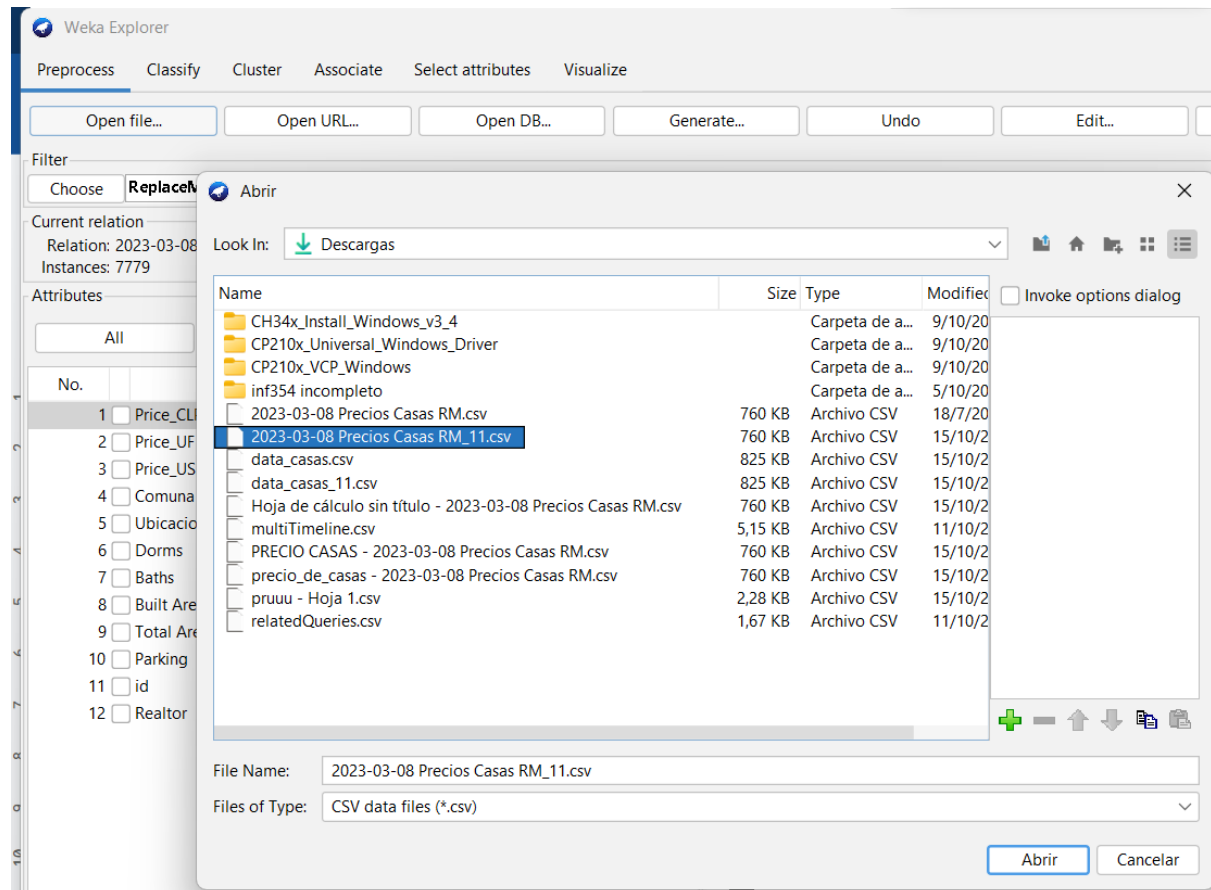


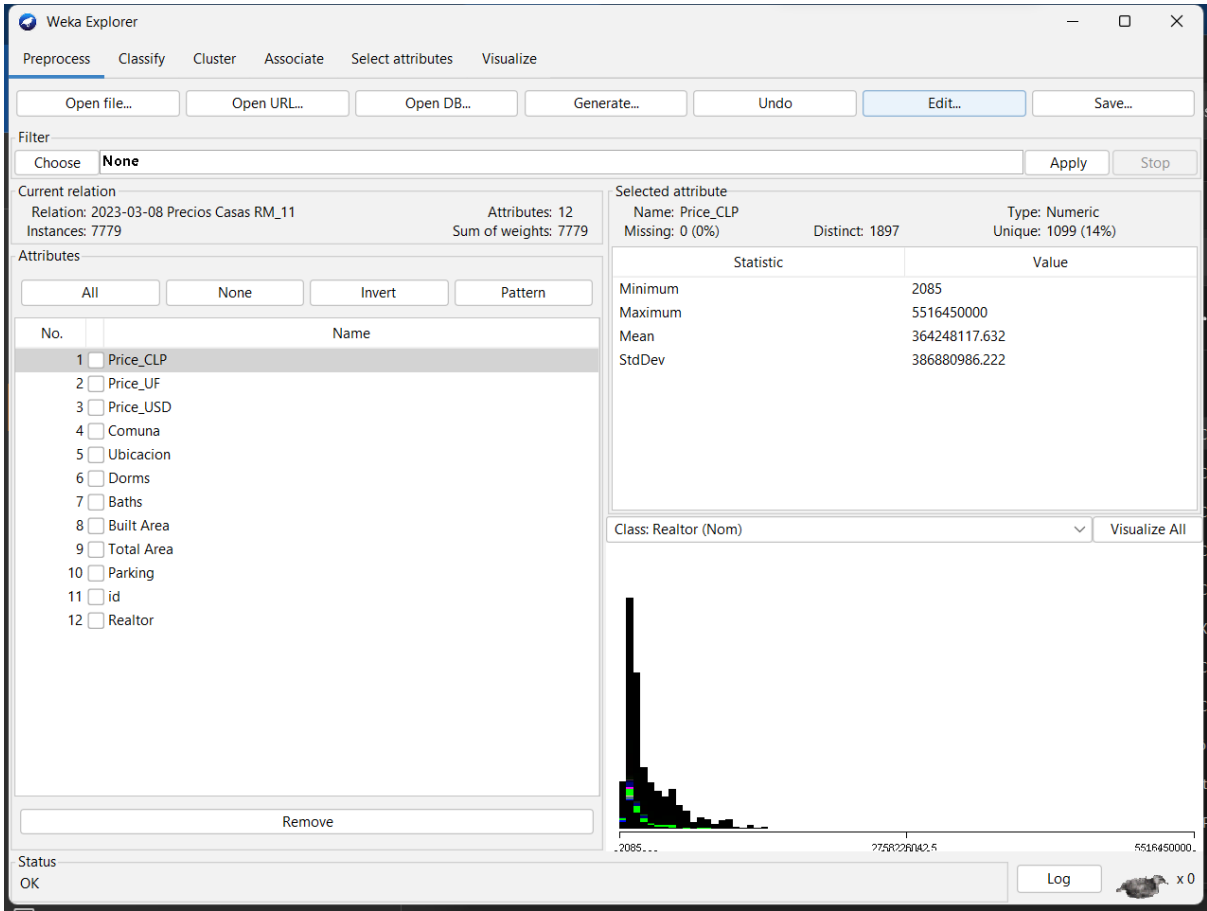
Pregunta 5

Del dataset elegido, migre el mismo a WEKA y utilice cuatro técnicas de preprocesamiento (realice la captura de pantallas de estos por fases). Explique la razón de aplicar estas técnicas.

paso 1. cargamos el dataset de “precios de casas” e importamos



paso2. vemos que se ha importado con éxito.



paso3. pulsamos el boton “**editar**” y observamos que el dataset tiene datos faltantes.

Relation: 2023-03-08 Precios Casas RM_11

No.	1: Price_CLP Numeric	2: Price_UF Numeric	3: Price_USD Numeric	4: Comuna Nominal	5: Ubicacion Nominal	6: Dorms Numeric	7: Baths Numeric	8: Built Area Numeric	9: Total Area Numeric	10: Parking Numeric	11: id Numeric	12: Realtor Nominal
1	4.09285E8	11500.0	509695.0	QuintaNor...	Hoevel4548...	7.0	4.0	384.0	732.0	3.0	1.1700...	
2	1.05E8	2950.0	130760.0	PedroAgu...	Rucalhue	2.0	1.0	112.0	145.0	1.0	1.0894...	Legales y P...
3	1.28124E8	3600.0	159557.0	EstaciÃ³n...	AvenidaLas...	3.0	1.0	59.0	243.0	2.0	1.0257...	Propiedad...
4	7.5E7	2107.0	93400.0	Colina	PasajeGonz...	3.0	1.0	103.0	73.0	1.0	92320...	Patricio Ga...
5	5.3E7	1489.0	66002.0	Colina	HernAnDÃ...	2.0	1.0	57.0	67.0	1.0	70853...	Patricio Ga...
6	9.4E7	2641.0	117061.0	EstaciÃ³n...	Avenida5de...	2.0	1.0	72.0	131.0		75863...	Patricio Ga...
7	7.5E7	2107.0	93400.0	Colina	GabrielaMis...	3.0	1.0	70.0	80.0	2.0	1.1688...	Patricio Ga...
8	1.05E8	2950.0	130760.0	LaFlorida	SanCristÃ³b...	2.0	1.0	50.0	120.0	1.0	1.2229...	Gesinprop ...
9	7.6E7	2135.0	94645.0	MaipÃº	PasajeJosÃ...	3.0	1.0	72.0	72.0	1.0	1.2223...	Gesinprop ...
10	7.6E7	2135.0	94645.0	MaipÃº	MatÃ-as	3.0	1.0	46.0	90.0	1.0	1.2221...	Gesinprop ...
11	9.0E7	2529.0	112080.0	MaipÃº	LasTinas&...	3.0	1.0	100.0	134.0	3.0	1.2188...	Gesinprop ...
12	7.5E7	2107.0	93400.0	SanBernar...	ValledelAco...	3.0	2.0	50.0	88.0	1.0	1.2187...	Gesinprop ...
13	1.53037E8	4300.0	190582.0	MaipÃº	Merlin	3.0	2.0	98.0	130.0	2.0	1.2187...	Gesinprop ...
14	1.32E8	3709.0	164384.0	QuintaNor...	SalvadorGut...	4.0	1.0	69.0	200.0	2.0	84930...	Paula Viva...
15	1.99E8	5591.0	247821.0	Santiago	Cuevas	7.0	3.0	162.0	185.0		1.2155...	Todo Propi...
16	7.9E7	2220.0	98381.0	PuenteAlto	PasajeLago...	2.0	1.0			2.0	1.2149...	Fe Propied...
17	3.73695E8	10500.0	465374.0	LaFlorida	JosÃ©Migu...	7.0	3.0	173.0	850.0	3.0	78509...	
18	4.8E8	13487.0	597758.0	LaFlorida	Altomacul	5.0	6.0	240.0	600.0	2.0	64748...	Inmobiliari...
19	1.42E8	3990.0	176837.0	Santiago	Sandaniel	3.0	2.0	100.0	163.0	2.0	1.2065...	
20	2.95397E8	8300.0	367867.0	LaFlorida	Caminodela...	3.0	3.0	105.0	280.0	2.0	1.0765...	Ecorredore...
21	3.9149E8	11000.0	487534.0	LaFlorida	Cam.LasCu...	5.0	3.0	150.0	400.0	2.0	1.0105...	Ecorredore...
22	1.21006E9	34000.0	1506924.0	LasCondes	Laforesta	5.0	5.0	360.0	1150.0	4.0	1.0080...	
23	1.60155E8	4500.0	199446.0	Lampa	AvenidaLa...	3.0	3.0	80.0	156.0	2.0	86770...	Behouse
24	1.299035E8	3650.0	161773.0	Lampa	AvenidaLaH...	3.0	2.0	68.0	221.0	1.0	78506...	Behouse

paso 4. elegimos la funcion “**ReplaceMissingValues**” para completar datos.

Filter: ReplaceMissingValues

Attributes: 12
Sum of weights: 7779

Selected attribute
Name: Price_CLP
Missing: 0 (0%)
Distinct: 1897
Type: Numeric
Unique: 1099 (14%)

Statistic	Value
Minimum	2085
Maximum	5516450000
Mean	364248117.632
StdDev	386880986.222

Replaces all missing values for nominal and numeric attributes in a dataset with the modes and means from the training data

The class attribute is skipped by default.

CAPABILITIES
Class -- Binary class, Date class, Empty nominal class, Missing class values, No class, Nominal class, Numeric class, Relational class, String class, Unary class

Attributes -- Binary attributes, Date attributes, Empty nominal attributes, Missing values, Nominal attributes, Numeric attributes, Relational attributes, String attributes, Unary attributes

Interfaces -- Sourcable, UnsupervisedFilter, WeightedAttributesHandler, WeightedInstancesHandler

Additional
Minimum number of instances: 0

paso 5. vemos que no se han completado los datos con valores de tipo object o String.

No.	1: Price_CLP	2: Price_UF	3: Price_USD	4: Comuna	5: Ubicacion	6: Dorms	7: Baths	8: Built Area	9: Total Area	10: Parking	11: id	12: Realtor
1	4.09285E8	11500.0	509695.0	QuintaNor...	Hoevel4548...	7.0	4.0	384.0	732.0	3.0	1.1700...	
2	1.05E8	2950.0	130760.0	PedroAgui...	Rucalhue	2.0	1.0	112.0	145.0	1.0	1.0894...	Legales y P...
3	1.28124E8	3600.0	159557.0	EstaciÃ³nC...	AvenidaLas...	3.0	1.0	59.0	243.0	2.0	1.0257...	Propiedad...
4	7.5E7	2107.0	93400.0	Colina	PasajeGonz...	3.0	1.0	103.0	73.0	1.0	92320...	Patricio Ga...
5	5.3E7	1489.0	66002.0	Colina	HernAnDÃ-...	2.0	1.0	57.0	67.0	1.0	70853...	Patricio Ga...
6	9.4E7	2641.0	117061.0	EstaciÃ³nC...	Avenida5de...	2.0	1.0	72.0	131.0	2.98050646...	75863...	Patricio Ga...
7	7.5E7	2107.0	93400.0	Colina	GabrielaMis...	3.0	1.0	70.0	80.0	2.0	1.1688...	Patricio Ga...
8	1.05E8	2950.0	130760.0	LaFlorida	SanCristÃ³b...	2.0	1.0	50.0	120.0	1.0	1.2229...	Gesinprop ...
9	7.6E7	2135.0	94645.0	MaipÃº	PasajeJosÃ...	3.0	1.0	72.0	72.0	1.0	1.2223...	Gesinprop ...
10	7.6E7	2135.0	94645.0	MaipÃº	MatÃ-as	3.0	1.0	46.0	90.0	1.0	1.2221...	Gesinprop ...
11	9.0E7	2529.0	112080.0	MaipÃº	LasTinajas&...	3.0	1.0	100.0	134.0	3.0	1.2188...	Gesinprop ...
12	7.5E7	2107.0	93400.0	SanBernar...	ValledelAco...	3.0	2.0	50.0	88.0	1.0	1.2187...	Gesinprop ...
13	1.53037E8	4300.0	190582.0	MaipÃº	Merlin	3.0	2.0	98.0	130.0	2.0	1.2187...	Gesinprop ...
14	1.32E8	3709.0	164384.0	QuintaNor...	SalvadorGut...	4.0	1.0	69.0	200.0	2.0	84930...	Paula Viva...
15	1.99E8	5591.0	247821.0	Santiago	Cuevas	7.0	3.0	162.0	185.0	2.98050646...	1.2155...	Todo Propi...
16	7.9E7	2220.0	98381.0	PuenteAlto	PasajeLago...	2.0	1.0	229.923669...	807.9198256...	2.0	1.2149...	Fe Propied...
17	3.73695E8	10500.0	465374.0	LaFlorida	JosÃ©Migu...	7.0	3.0	173.0	850.0	3.0	78509...	
18	4.8E8	13487.0	597758.0	LaFlorida	Altomacul	5.0	6.0	240.0	600.0	2.0	64748...	Inmobiliari...
19	1.42E8	3990.0	176837.0	Santiago	Sandaniel	3.0	2.0	100.0	163.0	2.0	1.2065...	
20	2.95397E8	8300.0	367867.0	LaFlorida	Caminodela...	3.0	3.0	105.0	280.0	2.0	1.0765...	Ecorredore...
21	3.9149E8	11000.0	487534.0	LaFlorida	Cam.LasCu...	5.0	3.0	150.0	400.0	2.0	1.0105...	Ecorredore...
22	1.21006E9	34000.0	1506924.0	LasCondes	Laforesta	5.0	5.0	360.0	1150.0	4.0	1.0080...	
23	1.60155E8	4500.0	199446.0	Lampa	AvenidaLa...	3.0	3.0	80.0	156.0	2.0	86770...	Behouse
24	1.299035E8	3650.0	161773.0	Lampa	AvenidaLaH...	3.0	2.0	68.0	221.0	1.0	78506...	Behouse

paso 6. hacemos click en la barra de la funcion a aplicar, y elegimos la opcion *ignoreClass* *True*.

Filter: **ReplaceMissingValues** -unset-class-temporarily

Current relation: 2023-03-08 Precios Casas RM_11-weka.filters.unsuperv...
Instances: 7779

Attributes: 12
Sum of weights: 7779

Selected attribute: Name: Price_CLP
Missing: 0 (0%)
Distinct: 1897
Type: Numeric
Unique: 1099 (14%)

weka.gui.GenericObjectEditor

About
Replaces all missing values for nominal and numeric attributes in a dataset with the modes and means from the training data.

debug: False
doNotCheckCapabilities: False
ignoreClass: **True**

Open... False

2085
5516450000
364248117.632
386880986.222

Visualize All

paso 7. vemos el resultado de aplicar las función “ReplaceMissingValues”

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Open file Open URL Open DB Generate Undo Edit Save

Viewer

Relation: 2023-03-08 Precios Casas RM_11-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.ReplaceMissingValues-unset-class-temporarily

No.	1: Price_CLP Numeric	2: Price_UF Numeric	3: Price_USD Numeric	4: Comuna Nominal	5: Ubicacion Nominal	6: Dorms Numeric	7: Baths Numeric	8: Built Area Numeric	9: Total Area Numeric	10: Parking Numeric	11: id Numeric	12: Realtor Nominal
1	4.09285E8	11500.0	509695.0	QuintaNor...	Hoevel4548...	7.0	4.0	384.0	732.0	3.0	1.1700...	Unne
2	1.05E8	2950.0	130760.0	PedroAgui...	Rucalhue	2.0	1.0	112.0	145.0	1.0	1.0894...	Legales y P...
3	1.28124E8	3600.0	159557.0	EstaciÃ³nC...	AvenidaLas...	3.0	1.0	59.0	243.0	2.0	1.0257...	Propiedad...
4	7.5E7	2107.0	93400.0	Colina	PasajeGonz...	3.0	1.0	103.0	73.0	1.0	92320...	Patricio Ga...
5	5.3E7	1489.0	66002.0	Colina	HernAndÃ...	2.0	1.0	57.0	67.0	1.0	70853...	Patricio Ga...
6	9.4E7	2641.0	117061.0	EstaciÃ³nC...	Avenida5de...	2.0	1.0	72.0	131.0	2.98050646...	75863...	Patricio Ga...
7	7.5E7	2107.0	93400.0	Colina	GabrielaMis...	3.0	1.0	70.0	80.0	2.0	1.1688...	Patricio Ga...
8	1.05E8	2950.0	130760.0	LaFlorida	SanCristÃ³b...	2.0	1.0	50.0	120.0	1.0	1.2229...	Gesinprop ...
9	7.6E7	2135.0	94645.0	MaipÃº	PasajeJosÃ...	3.0	1.0	72.0	72.0	1.0	1.2223...	Gesinprop ...
10	7.6E7	2135.0	94645.0	MaipÃº	MatÃ-as	3.0	1.0	46.0	90.0	1.0	1.2221...	Gesinprop ...
11	9.0E7	2529.0	112080.0	MaipÃº	LasTinajas&...	3.0	1.0	100.0	134.0	3.0	1.2188...	Gesinprop ...
12	7.5E7	2107.0	93400.0	SanBernar...	ValledelAco...	3.0	2.0	50.0	88.0	1.0	1.2187...	Gesinprop ...
13	1.53037E8	4300.0	190582.0	MaipÃº	Merlin	3.0	2.0	98.0	130.0	2.0	1.2187...	Gesinprop ...
14	1.32E8	3709.0	164384.0	QuintaNor...	SalvadorGut...	4.0	1.0	69.0	200.0	2.0	84930...	Paula Viva...
15	1.99E8	5591.0	247821.0	Santiago	Cuevas	7.0	3.0	162.0	185.0	2.98050646...	1.2155...	Todo Propi...
16	7.9E7	2220.0	98381.0	PuenteAlto	PasajeLago...	2.0	1.0	229.923669...	807.9198256...	2.0	1.2149...	Fe Propied...
17	3.73695E8	10500.0	465374.0	LaFlorida	JosÃ©Migu...	7.0	3.0	173.0	850.0	3.0	78509...	Unne
18	4.8E8	13487.0	597758.0	LaFlorida	Altomacul	5.0	6.0	240.0	600.0	2.0	64748...	Inmobiliari...
19	1.42E8	3990.0	176837.0	Santiago	Sandaniel	3.0	2.0	100.0	163.0	2.0	1.2065...	Unne
20	2.95397E8	8300.0	367867.0	LaFlorida	Caminodela...	3.0	3.0	105.0	280.0	2.0	1.0765...	Ecorredore...
21	3.9149E8	11000.0	487534.0	LaFlorida	Cam.LasCu...	5.0	3.0	150.0	400.0	2.0	1.0105...	Ecorredore...
22	1.21006E9	34000.0	1506924.0	LasCondes	Laforesta	5.0	5.0	360.0	1150.0	4.0	1.0080...	Unne
23	1.60155E8	4500.0	199446.0	Lampa	AvenidaLa...	3.0	3.0	80.0	156.0	2.0	86770...	Behouse
24	1.299035E8	3650.0	161773.0	Lampa	AvenidaLaH...	3.0	2.0	68.0	221.0	1.0	78506...	Behouse

Add instance Undo OK Cancel

Remove

Status OK

Log x 0

eso fue el preprocesamiento en weka