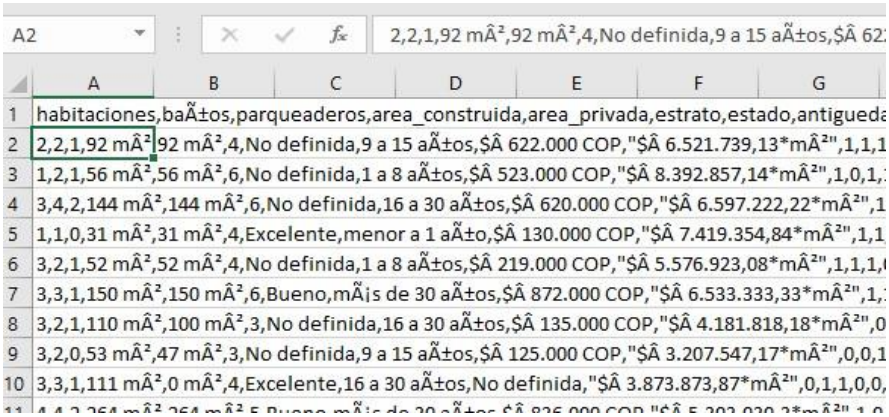
LABORATORIO 3

Gabriel Nicolas Aguilera Niño

1. VERIFICACION DEL ARCHIVO EL CUAL SE VA MANIPULAS

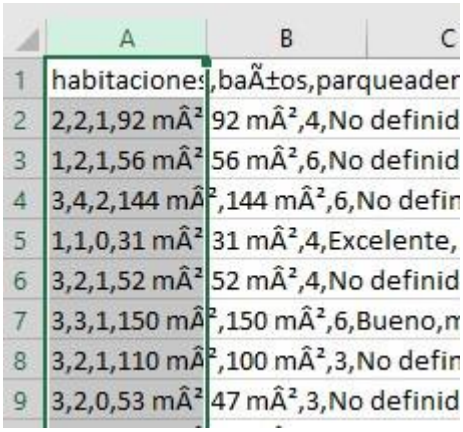
En primer lugar se identifica el formato del archivo el cual se desea limpiar, en este caso el archivo es un CSV llamado
"housing_fincaraiz".

Procedemos abrir el archivo.

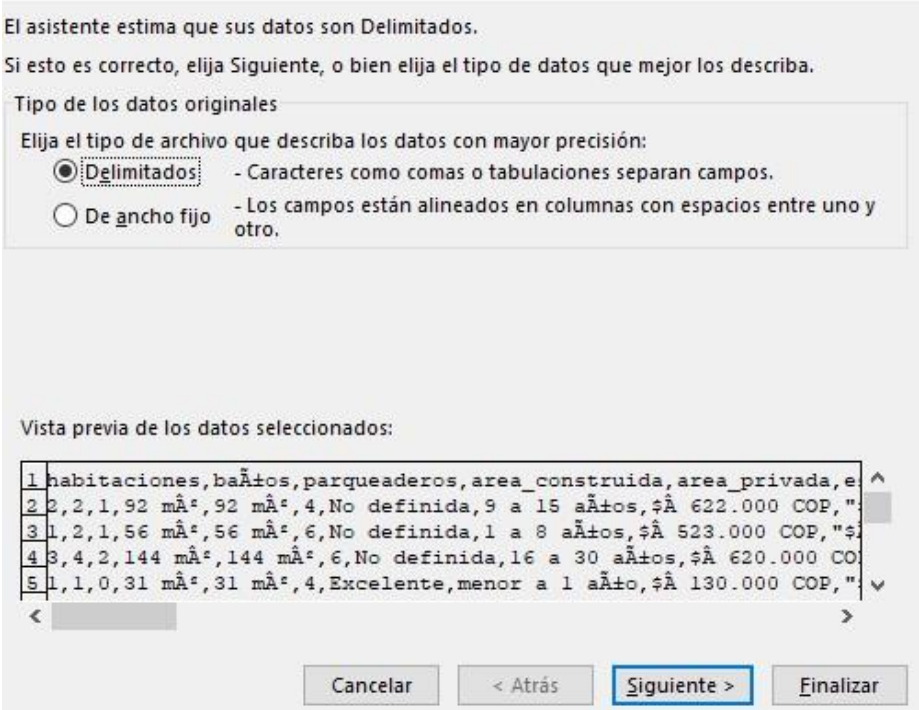


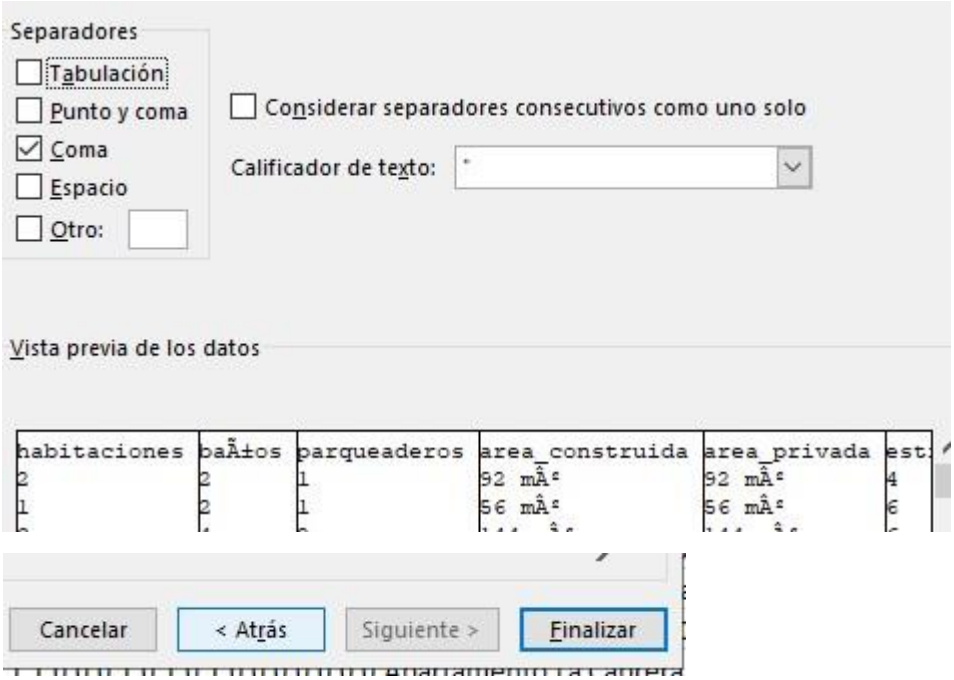
Se aprecia que es un archivo con formato CSV, con longitud amplia de datos, se observa que están delimitados por comas (,) y por el tipo de archivo no reconoce los caracteres especiales.

Procedemos a señalar la columna nuero 1 luego clic en datos y por último texto en columnas



Cuando nos aparece esta pantalla, podemos ver que si están delimitados por comas (,). Damos clic en siguiente y procedemos a seleccionar los separadores o las delimitaciones y damos clic en finalizar.

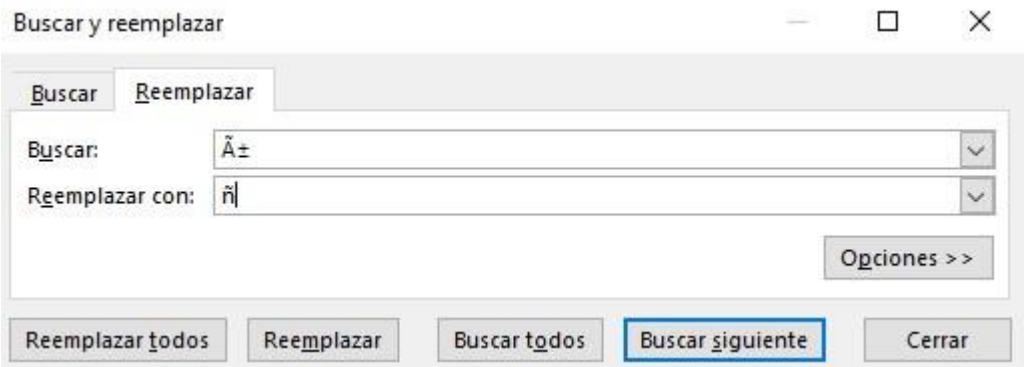


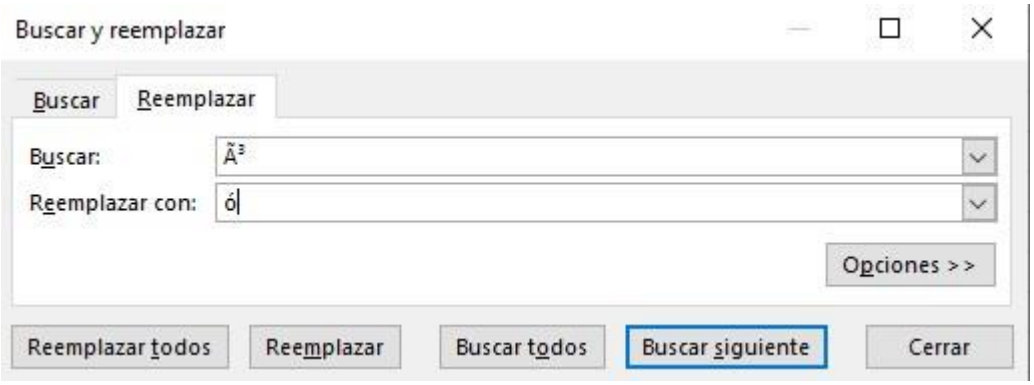


Una vez tengamos el CSV organizado podemos verlo de la siguiente manera.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	habitacio	baÃ±os	parqueaderc	area_constru	area_privada	estrato	estado	antiguedad	administraci	precio_m2	Ascensor	Circuito cerr	Parqueaderc	PorterÃ-a / F
2	2	2	1	92 mÂ²	92 mÂ²	4	No definida	9 a 15 aÃ±os	$Â 622.000 C	$Â 6.521.739	1	1	1	1
3	1	2	1	56 mÂ²	56 mÂ²	6	No definida	1 a 8 aÃ±os	$Â 523.000 C	$Â 8.392.857	1	0	1	1
4	3	4	2	144 mÂ²	144 mÂ²	6	No definida	16 a 30 aÃ±o	$Â 620.000 C	$Â 6.597.222	1	0	0	0
5	1	1	0	31 mÂ²	31 mÂ²	4	Excelente	menor a 1 aÃ	$Â 130.000 C	$Â 7.419.354	1	1	1	0
6	3	2	1	52 mÂ²	52 mÂ²	4	No definida	1 a 8 aÃ±os	$Â 219.000 C	$Â 5.576.923	1	1	1	0
7	3	3	1	150 mÂ²	150 mÂ²	6	Bueno	mÃ¡s de 30 a	$Â 872.000 C	$Â 6.533.333	1	1	1	0
8	3	2	1	110 mÂ²	100 mÂ²	3	No definida	16 a 30 aÃ±o	$Â 135.000 C	$Â 4.181.818	0	1	0	0
9	3	2	0	53 mÂ²	47 mÂ²	3	No definida	9 a 15 aÃ±os	$Â 125.000 C	$Â 3.207.547	0	0	1	1
10	3	3	1	111 mÂ²	0 mÂ²	4	Excelente	16 a 30 aÃ±o	No definida	$Â 3.873.873	0	1	1	0
11	4	4	2	264 mÂ²	264 mÂ²	5	Bueno	mÃ¡s de 30 a	$Â 836.000 C	$Â 5.303.030	1	0	0	0
12	3	2	2	97 mÂ²	0 mÂ²	4	No definida	9 a 15 aÃ±os	$Â 272.000 C	$Â 6.175.257	1	1	0	0
13	2	3	2	87 mÂ²	82 mÂ²	4	No definida	9 a 15 aÃ±os	$Â 350.000 C	$Â 4.712.643	1	1	1	1
14	3	4	2	175 mÂ²	175 mÂ²	4	No definida	16 a 30 aÃ±o	No definida	$Â 4.285.714	0	0	0	0
15	10	5	2	391 mÂ²	0 mÂ²	3	No definida	16 a 30 aÃ±o	No definida	$Â 1.994.884	0	0	0	0

Buscamos y remplazamos los caracteres especiales



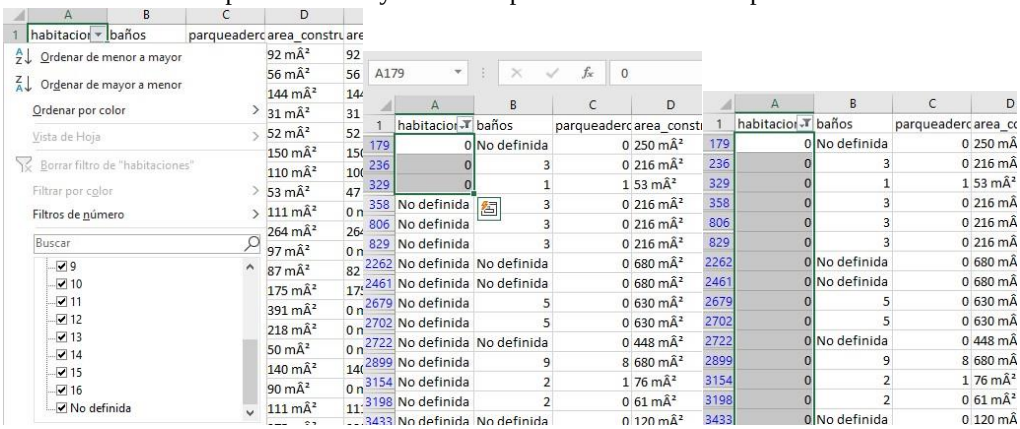


2. DARLE UN FORMATO A LOS DATOS

Después de ordenar el CSV, le damos formato a cada una de nuestras columnas, excepto los títulos.

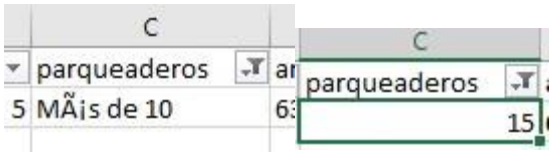
Para el siguiente caso, mi columna de habitaciones es un entero, lo que significa que siempre dice si son 1 habitación, 2 habitación, 3 habitación o 4 habitaciones. Sin embargo, la palabra "No definido" no es un número, por lo que la cambiaré por 0 para indicar que no está definido.

Aplico un filtro y cambio la palabra "No definida" por 0.

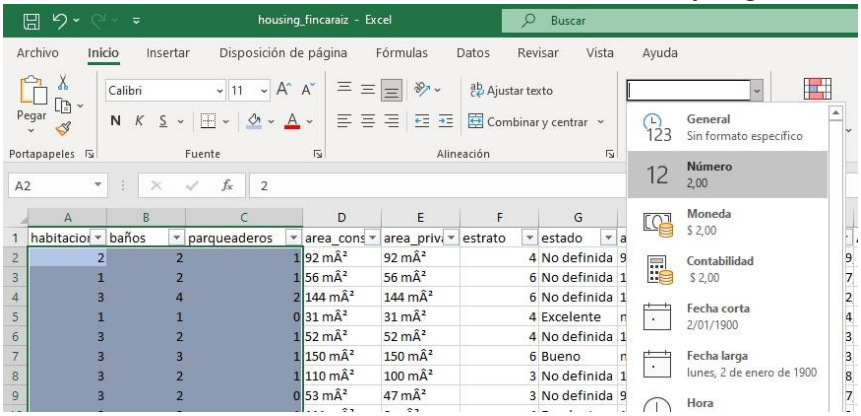


Para las columnas de tipo número entero, hacemos lo mismo. En este caso, voy a pintar las columnas de azul para ayudarme a mostrar cuales ya les di formato.

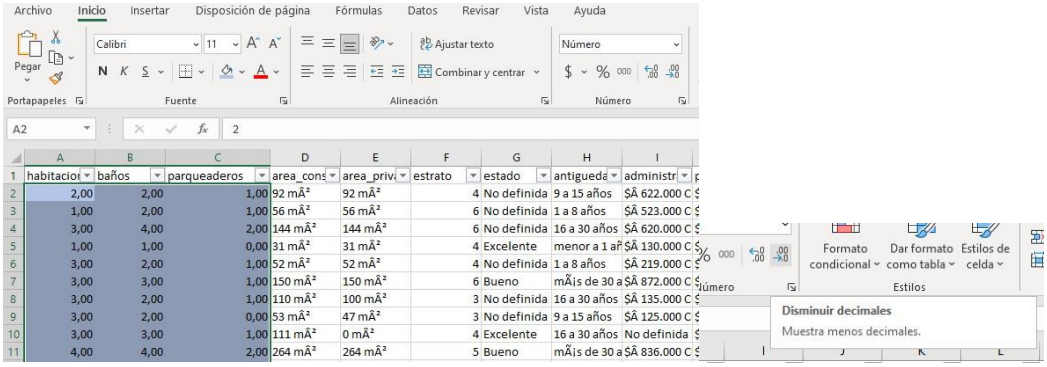
Para este caso de estacionamientos, establecí un número que fuera superior a 10, y para este caso, establecí el 15.



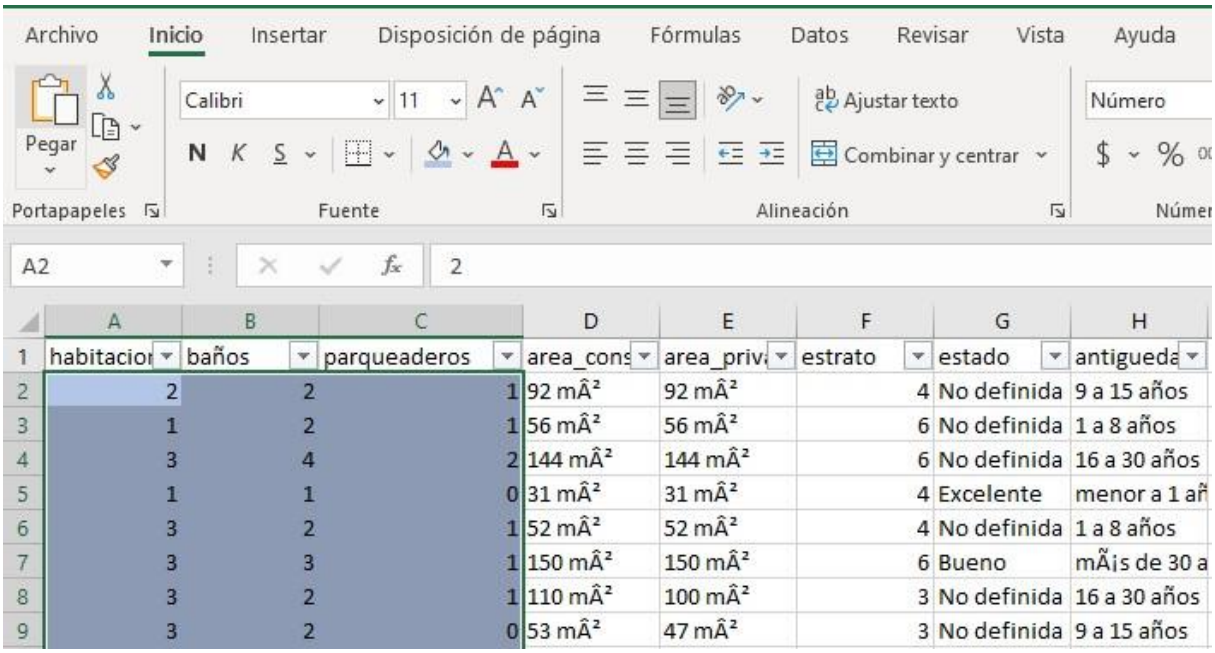
Para darle formato a las columnas, seleccione todo y luego seleccione el formato número.



Los números quedarán como sigue: Para eliminar esos ceros, haga clic en disminuir decimales hasta que se eliminen todos los ceros.

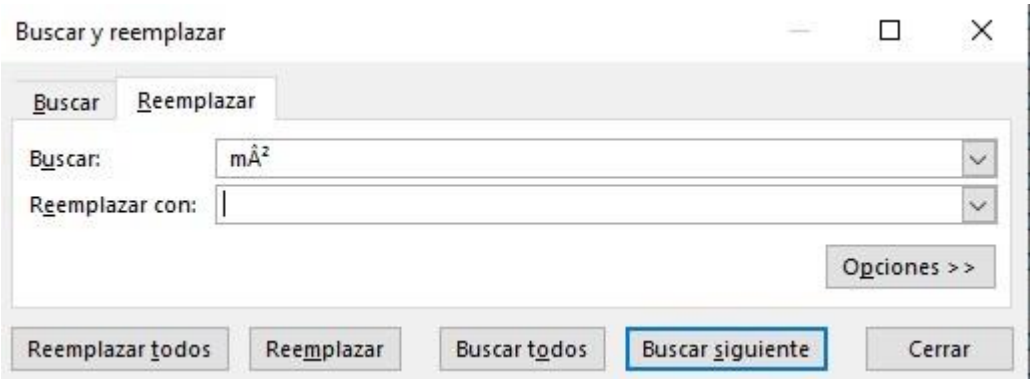


Luego veremos nuestros numero nuevamente de la misma manera, pero en formato numérico.



	A	B	C	D	E	F	G	H
1	habitacior	baños	parqueaderos	area_cons	area_priv	estrato	estado	antigueda
2	2	2	1	92 mÂ²	92 mÂ²	4	No definida	9 a 15 años
3	1	2	1	56 mÂ²	56 mÂ²	6	No definida	1 a 8 años
4	3	4	2	144 mÂ²	144 mÂ²	6	No definida	16 a 30 años
5	1	1	0	31 mÂ²	31 mÂ²	4	Excelente	menor a 1 añ
6	3	2	1	52 mÂ²	52 mÂ²	4	No definida	1 a 8 años
7	3	3	1	150 mÂ²	150 mÂ²	6	Bueno	mÃ¡s de 30 a
8	3	2	1	110 mÂ²	100 mÂ²	3	No definida	16 a 30 años
9	3	2	0	53 mÂ²	47 mÂ²	3	No definida	9 a 15 años

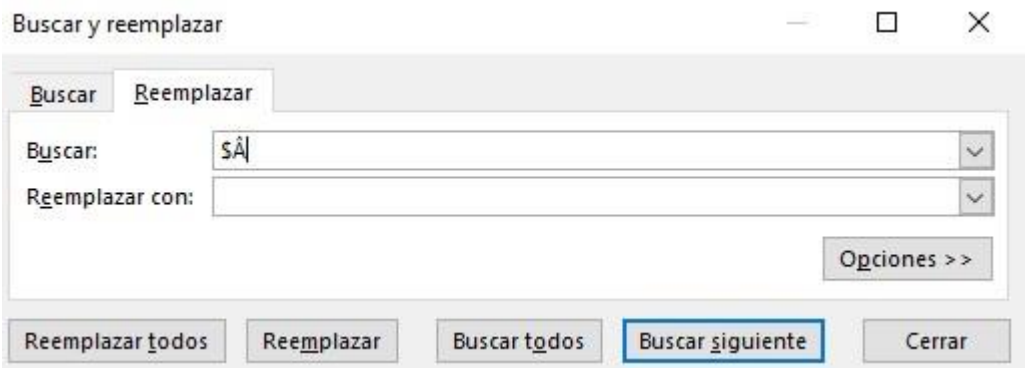
Para las dos columnas posterioes, la unidad de medida es el metro cuadrado. Como este campo puede ser decimal, seguimos el mismo procedimiento, eliminamos el metro cuadrado y ponemos en formato numérico, aunque dejamos al menos un decimal.

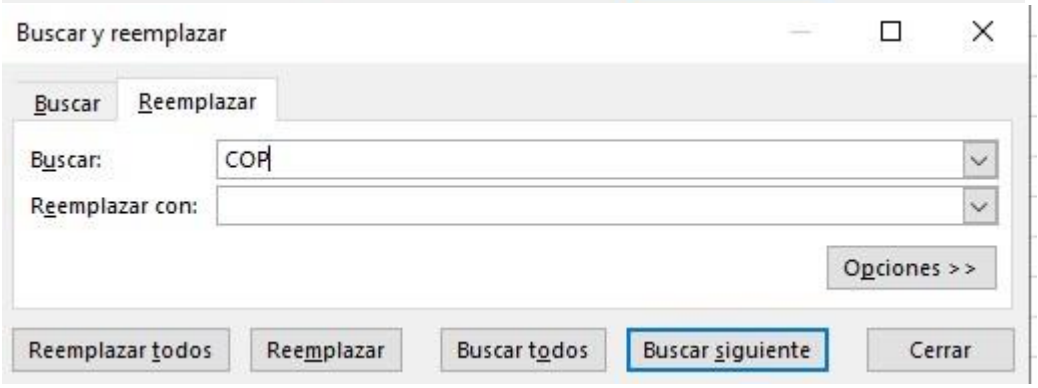


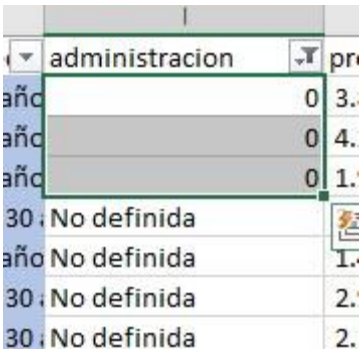
area_construida	area_privada
92,0	92,0
56,0	56,0
144,0	144,0
31,0	31,0
52,0	52,0
150,0	150,0
110,0	100,0
53,0	47,0
111,0	0,0
264,0	264,0
97,0	0,0
87,0	82,0
175,0	175,0

Para las colu

administracion	precio_m2
$Â 622.000 COP	$Â 6.521.739,13*mÂ²
$Â 523.000 COP	$Â 8.392.857,14*mÂ²
$Â 620.000 COP	$Â 6.597.222,22*mÂ²
$Â 130.000 COP	$Â 7.419.354,84*mÂ²
$Â 219.000 COP	$Â 5.576.923,08*mÂ²
$Â 872.000 COP	$Â 6.533.333,33*mÂ²



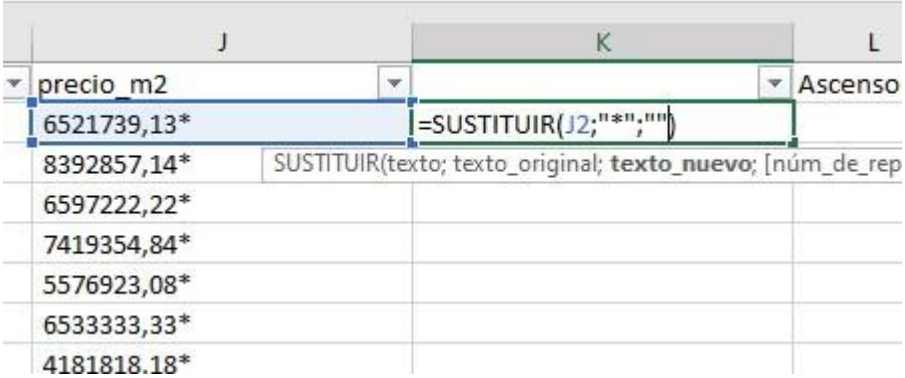




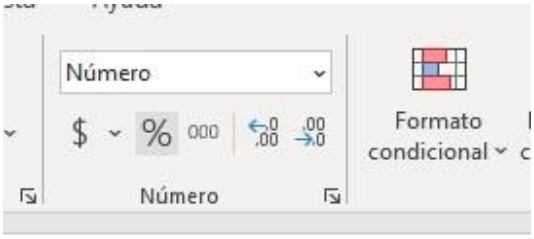
Antes de obtener solo un número, conviértelo a formato entero.

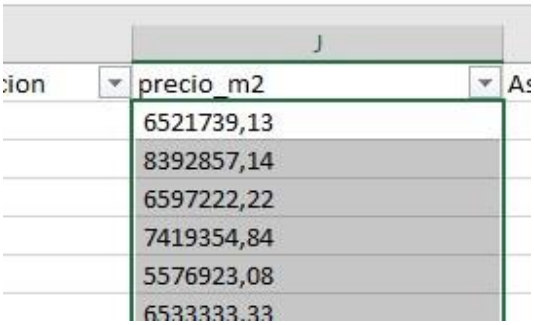
Para la columna Precio por metro cuadrado, utiliza Reemplazar para eliminar los datos, pero debes hacerlo de esta manera para eliminar los asteriscos.

Después de reemplazar, debe copiar y pegar el valor sobre el valor con asterisco en lugar de crearlo.



Por último, debemos damos formato a la columna y listo

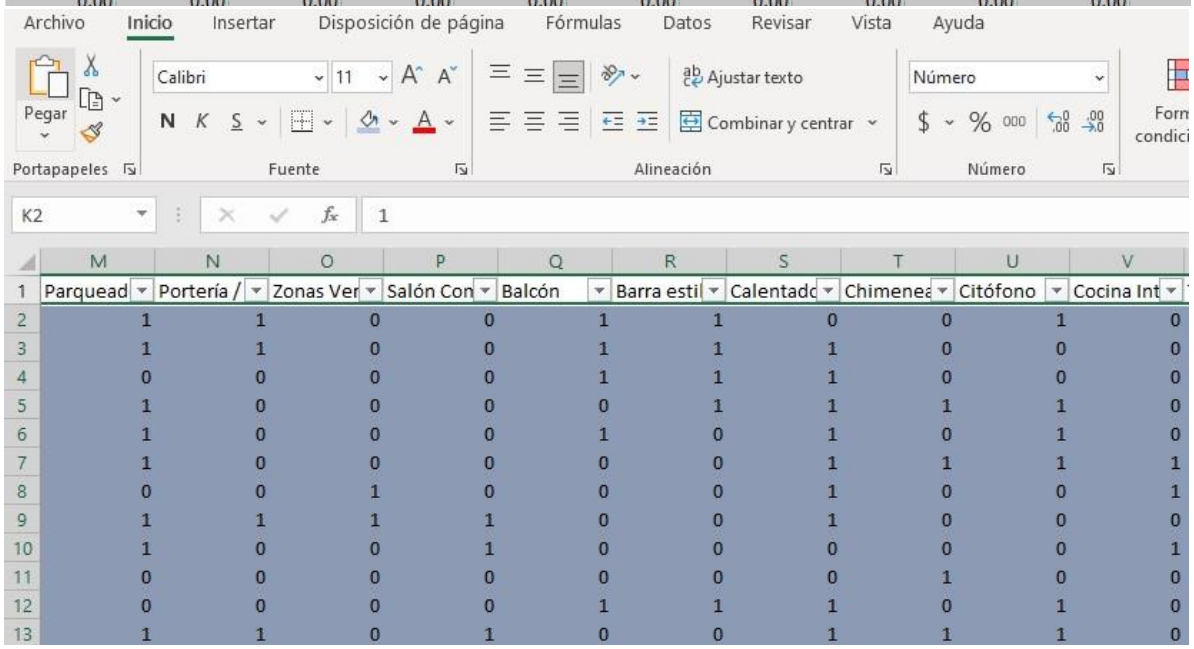




Hacemos los mismos pasos con las siguientes columnas y procedemos a dejarlas tipo numéricas.

Ascensor	Circuito c	Parquead	Portería /	Zonas Ver	Salón Con	Balcón	Barra esti	Calentado	Chimenea	Citófono	Cocina Int	Ter
1	1	1	1	0	0	1	1	0	0	1	0	
1	0	1	1	0	0	1	1	1	0	0	0	
1	0	0	0	0	0	1	1	1	0	0	0	
1	1	1	0	0	0	0	1	1	1	1	0	
1	1	1	0	0	0	1	0	1	0	1	0	
1	1	1	0	0	0	0	0	1	1	1	1	
0	1	0	0	1	0	0	0	1	0	0	1	
0	0	1	1	1	1	0	0	1	0	0	0	
0	1	1	0	0	1	0	0	0	0	0	1	
1	0	0	0	0	0	0	0	0	1	0	0	
1	1	0	0	0	0	1	1	1	0	1	0	
1	1	1	1	0	1	0	0	1	1	1	0	
0	0	0	0	1	0	0	0	0	1	0	1	

Ascensor	Circuito c	Parquead	Portería /	Zonas Ver	Salón Con	Balcón	Barra esti	Calentado	Chimenea	Citófon
1,00	1,00	1,00	1,00	0,00	0,00	1,00	1,00	0,00	0,00	
1,00	0,00	1,00	1,00	0,00	0,00	1,00	1,00	1,00	0,00	
1,00	0,00	0,00	0,00	0,00	0,00	1,00	1,00	1,00	0,00	
1,00	1,00	1,00	0,00	0,00	0,00	0,00	1,00	1,00	1,00	
1,00	1,00	1,00	0,00	0,00	0,00	1,00	0,00	1,00	0,00	
1,00	1,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	1,00	
0,00	1,00	0,00	0,00	1,00	0,00	0,00	0,00	1,00	0,00	
0,00	0,00	1,00	1,00	1,00	1,00	0,00	0,00	1,00	0,00	
0,00	1,00	1,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	
1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	
1,00	1,00	0,00	0,00	0,00	0,00	1,00	1,00	1,00	0,00	
1,00	1,00	1,00	1,00	0,00	1,00	0,00	0,00	1,00	1,00	
0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00	1,00	
0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	



	M	N	O	P	Q	R	S	T	U	V
1	Parquead	Portería /	Zonas Ver	Salón Con	Balcón	Barra esti	Calentado	Chimenea	Citófono	Cocina Int
2	1	1	0	0	1	1	0	0	1	0
3	1	1	0	0	1	1	1	0	0	0
4	0	0	0	0	1	1	1	0	0	0
5	1	0	0	0	0	1	1	1	1	0
6	1	0	0	0	1	0	1	0	1	0
7	1	0	0	0	0	0	1	1	1	1
8	0	0	1	0	0	0	1	0	0	1
9	1	1	1	1	0	0	1	0	0	0
10	1	0	0	1	0	0	0	0	0	1
11	0	0	0	0	0	0	0	1	0	0
12	0	0	0	0	1	1	1	0	1	0
13	1	1	0	1	0	0	1	1	1	0

Una vez tengamos limpios los datos procedemos a guardar un nuevo CSV.