

# Properati

Analisis exploratorio

Equipo: Gianfranco Quaranta , Nicolas Baranek, Matias Lorenzatti, Vanesa Cespedes, Sebastian Rosido, Joaquin Solis



# Punto de partida

El análisis se hizo en base a un archivo csv con información sobre la oferta de propiedades para venta y alquiler.

El DataFrame original (df) fue leído con Pandas y constaba de 121220 filas y 26 columnas.

Entre las columnas encontramos 13 del tipo float 13, 12 del tipo object y una int64.



# Información de interes para los usuarios de Properatti.

Columnas a trabajar en el data set.

OPERATION (operación):

Tipo de operación inmobiliaria ofrecida, venta, alquiler,permuta,etc.

En este caso se trabaja solo con VENTAS.

PROPERATY\_TYPE:

Tipo de propiedad a adquirir.

PLASE\_NAME:

Barrio de Buenos Aires

COUNTRY\_NAME

Pais-Argentina

STATE\_NAME:

Provincia o region.

LAT:

Dato de la latitud de la posicion geoespacial del inmueble

LON

Dato de la longitud de la posicion geoespacial del inmueble.

PRICE: Precio de la propiedad.

CONCURRENCY: Moneda usada para valuar la propiedad.

SURFACE\_TOTAL\_IN\_m2: Superficie total del inmueble.

FLOOR: pisos en el inmueble.

ROOMS: cantidad de habitaciones.

EXPENSES: expensas del inmueble

# Información data set



EL DATA SET PRESENTA UN ALTO PORCENTAJE DE DATOS NULOS

## Columnas a trabajar

- "state\_name"
- "property\_type"
- "surface\_total\_in\_m2"
- "surface\_covered\_in\_m2"
- "price\_aprox\_usd"
- "lon"
- "lat"
- "rooms"
- "place\_name"
- "geonames\_id"
- "title"
- "description"

Son importantes para el analisis  
predictivo

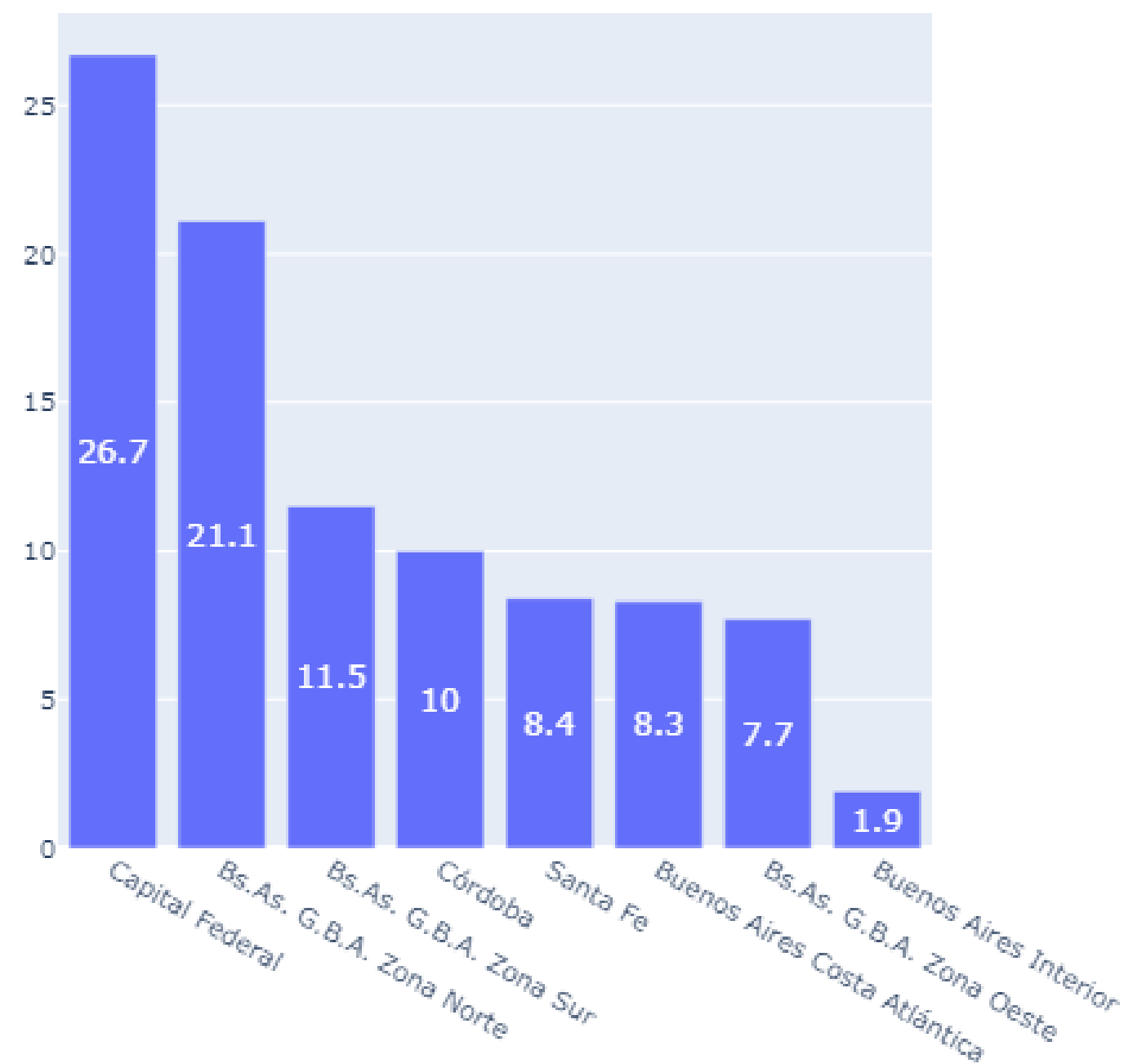
## Columnas eliminadas

- "Unnamed"
- "operation"
- "lat-lon"
- "floor"
- "expenses"
- "currency"
- "price"
- "price\_aprox\_local\_currency"

Repetitivas respecto a otras columnas  
Muy pocos datos no nulos

<b>floor</b>	<b>% 93.48</b>
<b>expenses</b>	<b>% 88.23</b>
<b>rooms</b>	<b>% 60.91</b>
<b>price_usd_per_m2</b>	<b>% 43.39</b>
<b>lat</b>	<b>% 42.53</b>
<b>lon</b>	<b>% 42.53</b>
<b>surface_total_in_m2</b>	<b>% 32.44</b>
<b>price_per_m2</b>	<b>% 27.69</b>
<b>currency</b>	<b>% 16.84</b>
<b>price</b>	<b>% 16.84</b>
<b>price_aprox_local_currency</b>	<b>% 16.84</b>
<b>surface_covered_in_m2</b>	<b>% 16.42</b>
<b>geonames_id</b>	<b>% 15.44</b>
<b>description</b>	<b>% 0.00</b>

**Propiedades por Provincia**

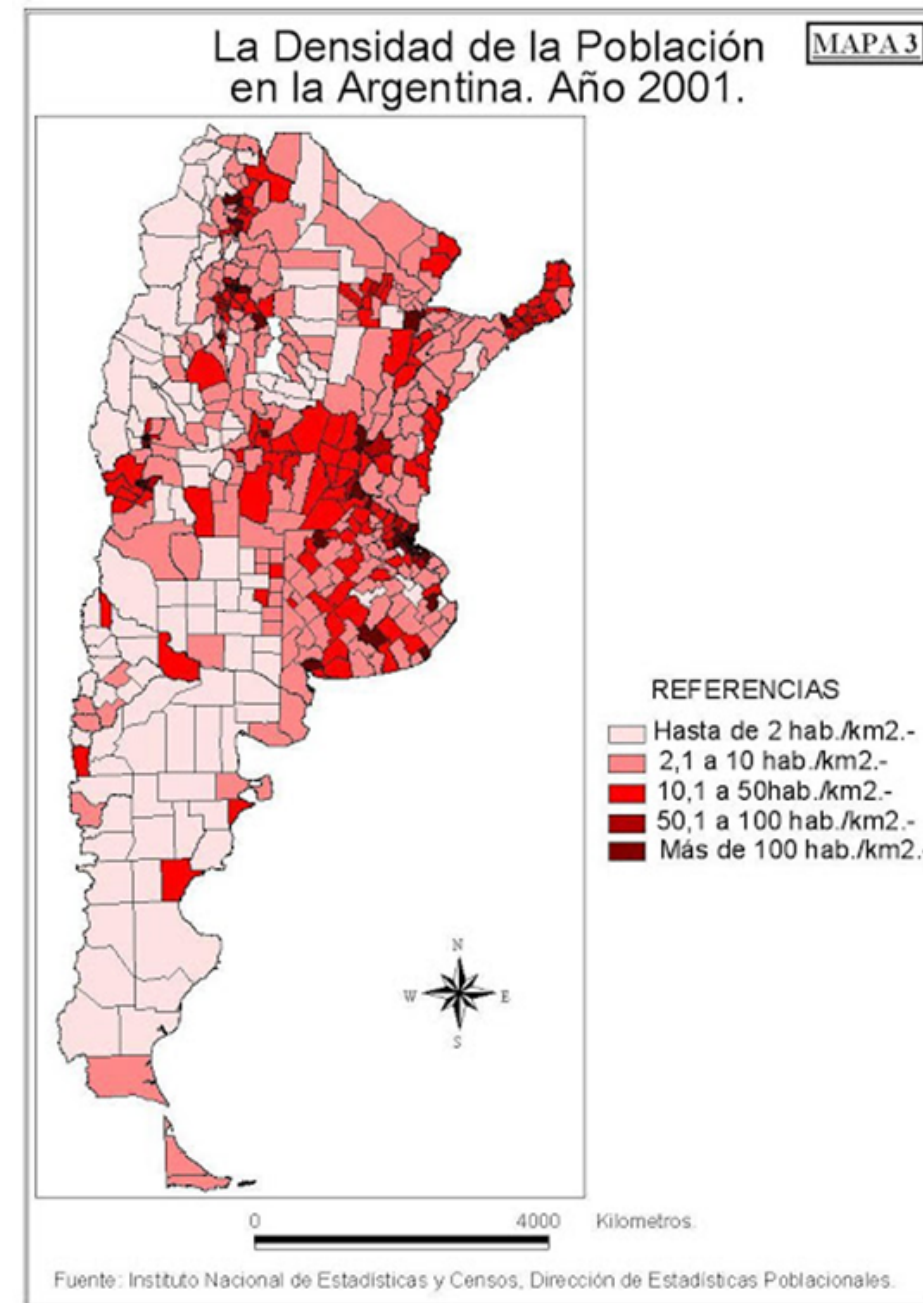
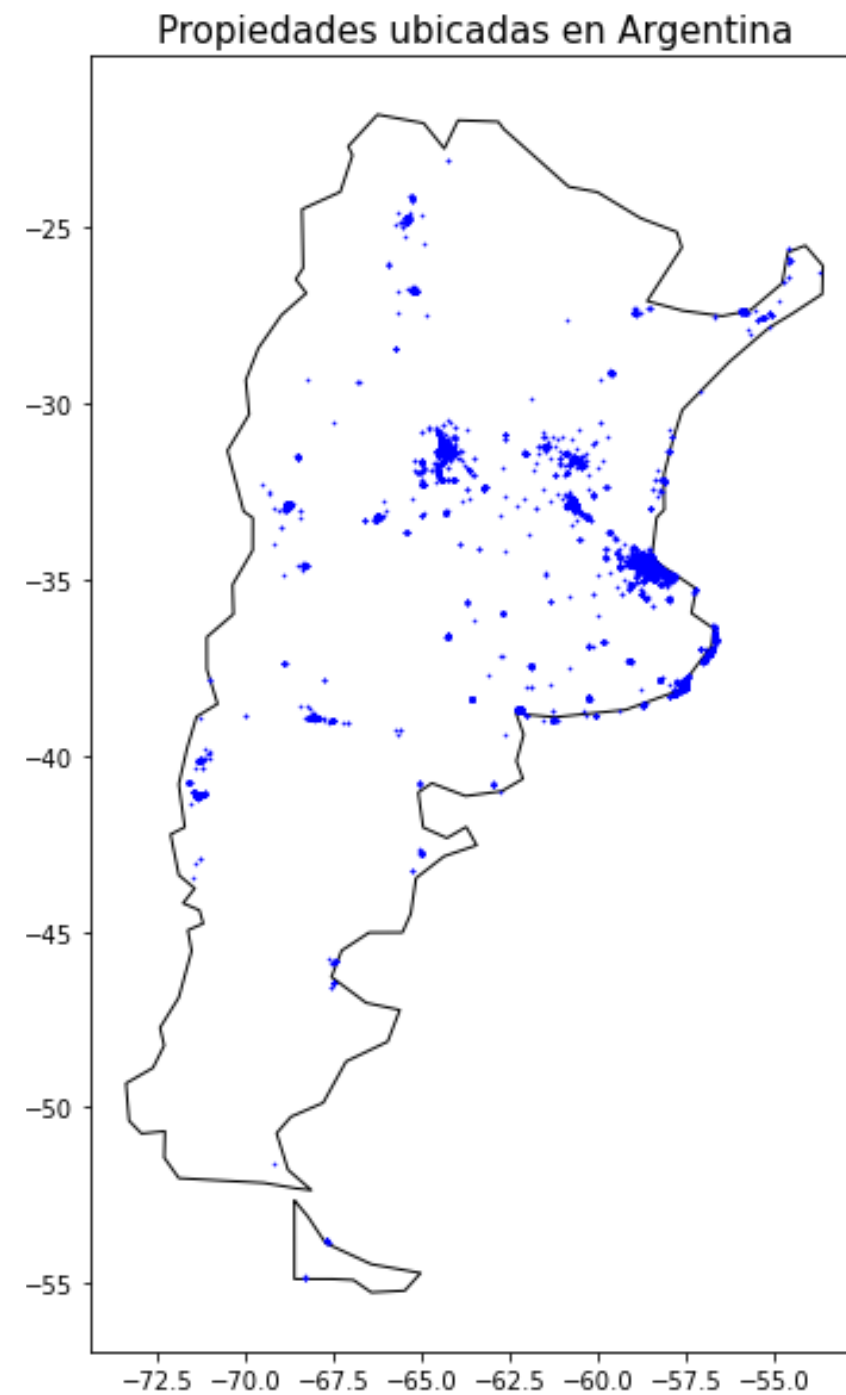


## Las primeras 8 regiones acumulan el 95,6% de las publicaciones

De estas 8 regiones el 67% se encuentran en el área metropolitana.

El mayor porcentaje esta representado por la Capital Federal con el 26,7%, que a su vez es la región con más densidad de población.

## El numero de publicaciones por region esta relacionado con la densidad poblacional

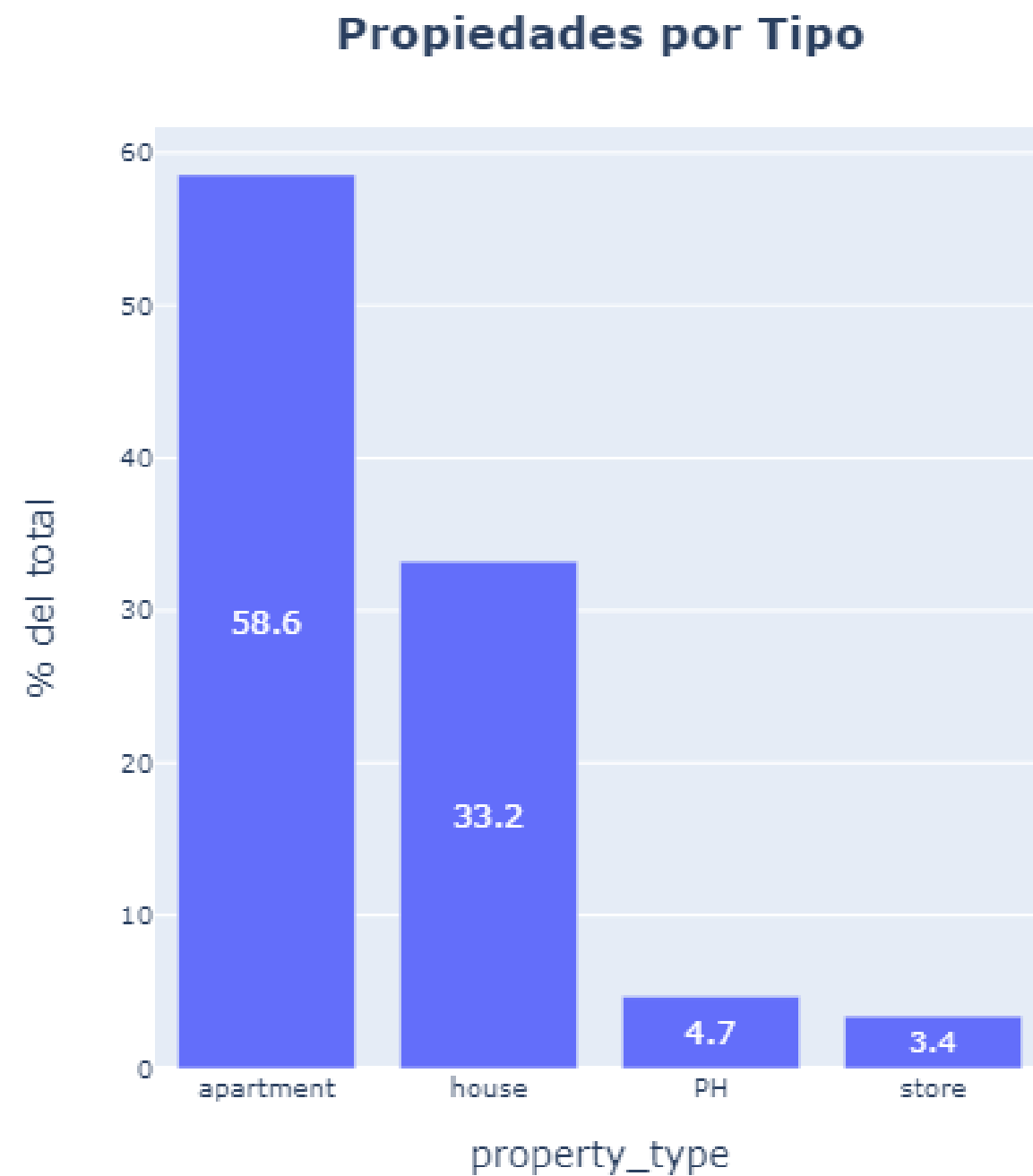


Conocer la dispersión geográfica, será un punto fundamental que nos ayude a caracterizar y agrupar conductas de los precios más adelante

Se puede observar claramente la concentración en las zonas urbanas

Cuando comparamos nuestro gráfico con la densidad de población al último CENSO encontramos una clara relación, pero notamos un sesgo hacia las propiedades en la zona Noreste del país

# El 91,8% de las publicaciones corresponden a casas o departamentos.



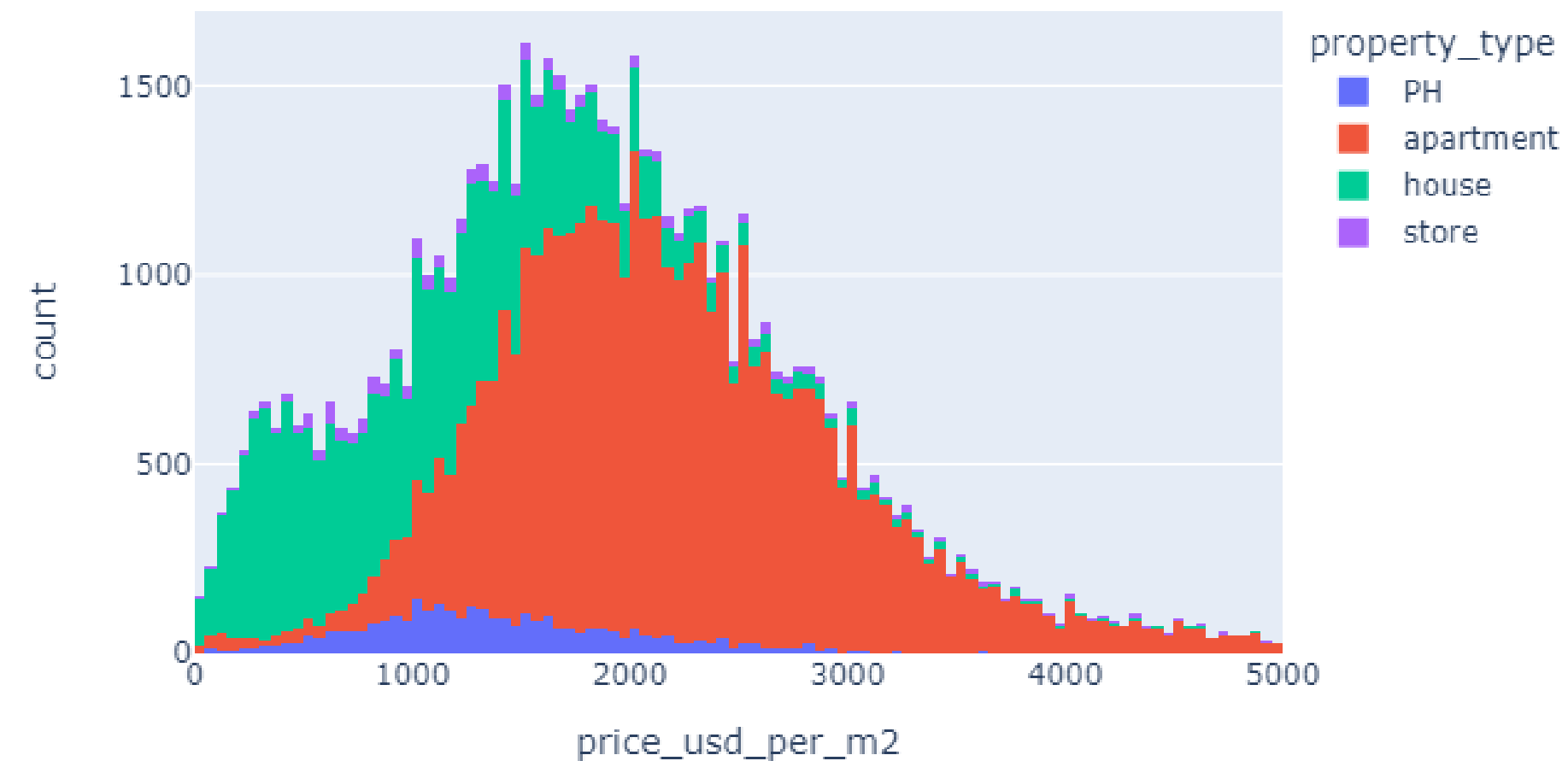
Ya que la columna de TIPO DE PROPIEDAD no tiene valores nulos nos permite tener una mirada general sobre todos los datos

Existen cuatro tipos de propiedades publicadas :  
DEPARTAMENTOS, CASAS, PH y TIENDAS/LOCALES

Los departamentos son el tipo mas común con 71065 registros encontrados.

# Precios

Frecuencia acumulada de precio por m2 en USD por tipo de propiedad



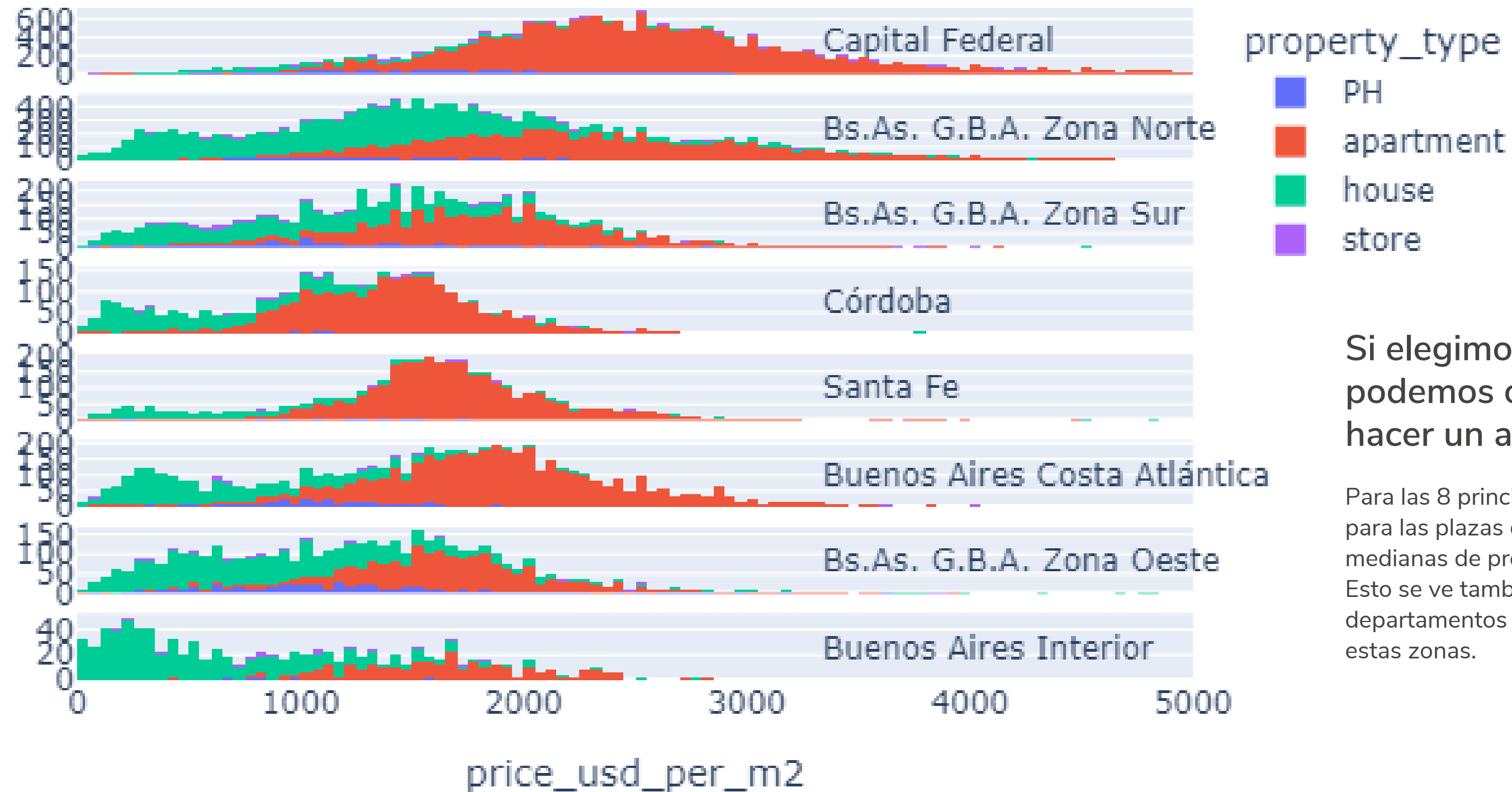
Si elegimos cortar los precios extremos, podemos observar mejor la distribución y hacer un análisis más exhaustivo

Para las 8 principales regiones, se observa la mayor dispersión para las plazas de Capital Federal y GBA Norte; ambas con medianas de precio superiores.

Esto se ve también sumado a una mayor cantidad de departamentos sobre el total de propiedades publicadas en estas zonas.



# Precios por zona



Si elegimos cortar los precios extremos, podemos observar mejor la distribución y hacer un análisis más exhaustivo

Para las 8 principales regiones, se observa la mayor dispersión para las plazas de Capital Federal y GBA Norte; ambas con medianas de precio superiores. Esto se ve también sumado a una mayor cantidad de departamentos sobre el total de propiedades publicadas en estas zonas.

# Precios:

		Media	Desviación Estándar	Precio Minimo	Precio Maximo
Tipo de propiedad	Provincia				
apartment	Bs.As. G.B.A. Zona Norte	2539.198037	2382.482184	31.246621	48296.222222
	Bs.As. G.B.A. Zona Oeste	1718.724709	871.935979	22.107656	24500.000000
	Bs.As. G.B.A. Zona Sur	1888.912977	1239.061350	91.135875	23140.764706
	Buenos Aires Costa Atlántica	1961.070956	815.550387	26.262626	21897.058824
	Buenos Aires Interior	1643.059367	506.580200	107.853077	3657.142857
	Capital Federal	3325.388251	3895.143204	49.000000	206333.333333
	Córdoba	1428.786003	532.797780	1.184571	14960.000000
	Santa Fe	2392.245874	4331.372338	0.600000	44425.675676
TOTALES		2693.050228	3208.031692	0.600000	206333.333333

		Media	Desviación Estándar	Precio Minimo	Precio Maximo
Tipo de propiedad	Provincia				
house	Bs.As. G.B.A. Zona Norte	1341.578924	1188.927989	4.166667	38000.000000
	Bs.As. G.B.A. Zona Oeste	1053.891080	1956.782600	7.000000	80000.000000
	Bs.As. G.B.A. Zona Sur	1161.544603	761.060416	19.314341	9411.764706
	Buenos Aires Costa Atlántica	749.953688	534.328605	3.000000	5698.005128
	Buenos Aires Interior	647.793745	558.987525	2.072029	4166.666667
	Capital Federal	1546.437925	1658.398601	190.000000	23064.516129
	Córdoba	714.335569	777.596681	12.254902	16666.666667
	Santa Fe	992.574612	2107.895565	37.916667	36549.707237
TOTALES		1176.186030	1286.090049	2.072029	80000.000000

		Media	Desviación Estándar	Precio Minimo	Precio Maximo
Tipo de propiedad	Provincia				
PH	Bs.As. G.B.A. Zona Norte	1451.712566	938.917931	49.485588	11333.333333
	Bs.As. G.B.A. Zona Oeste	1224.018825	1159.268926	77.893944	16000.000000
	Bs.As. G.B.A. Zona Sur	1114.012013	490.746480	50.475300	2800.000000
	Buenos Aires Costa Atlántica	1156.454370	554.608810	55.000000	5957.446809
	Buenos Aires Interior	1019.888682	512.080247	45.824847	1867.647059
	Capital Federal	1695.517729	887.360632	120.964706	25000.000000
	Córdoba	1544.048178	4070.615915	36.601307	30000.000000
	Santa Fe	1207.188748	406.055353	218.103167	1724.573000
TOTALES		1444.379348	1048.326069	36.601307	30000.000000

		Media	Desviación Estándar	Precio Minimo	Precio Maximo
Tipo de propiedad	Provincia				
store	Bs.As. G.B.A. Zona Norte	1862.009755	1819.141377	158.730159	20833.333333
	Bs.As. G.B.A. Zona Oeste	1556.947316	1627.518711	28.041833	11818.181818
	Bs.As. G.B.A. Zona Sur	1541.024184	1524.366776	22.222222	13000.000000
	Buenos Aires Costa Atlántica	1401.045662	953.937115	82.540833	4740.259740
	Buenos Aires Interior	987.670901	610.157332	133.928571	2333.333333
	Capital Federal	2899.667204	2222.673859	13.797467	22000.000000
	Córdoba	1458.775597	844.103398	31.406853	5521.472393
	Santa Fe	3002.901146	7963.224291	416.666667	57382.080000
TOTALES		2234.083533	2353.550340	13.797467	57382.080000

# Correcciones | apartment |:

## Estadísticas ANTES de la corrección

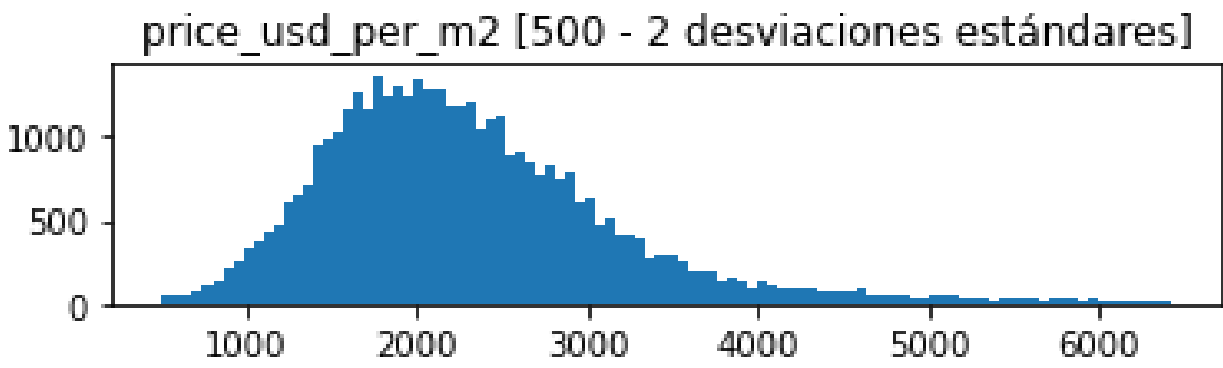
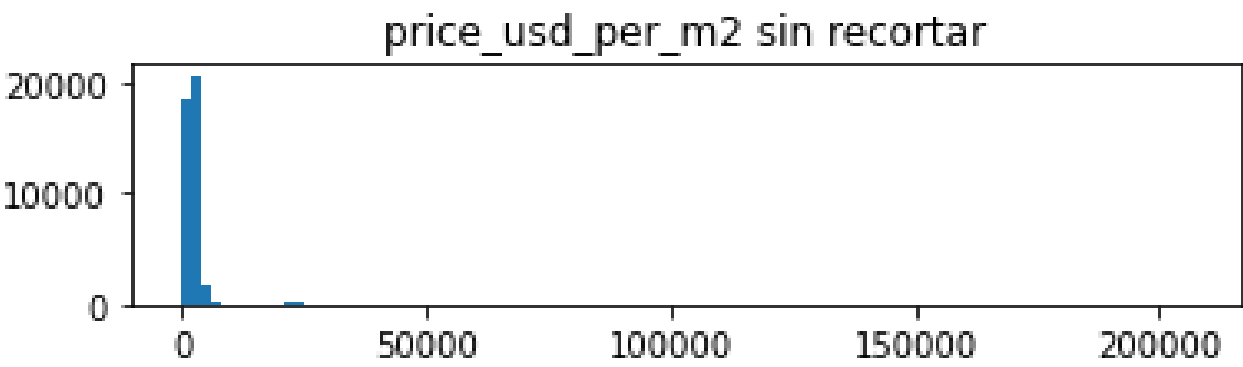
		Media	Desviación Estándar	Precio Minimo	Precio Maximo
Tipo de propiedad	Provincia				
apartment	Bs.As. G.B.A. Zona Norte	2539.198037	2382.482184	31.246621	48296.222222
	Bs.As. G.B.A. Zona Oeste	1718.724709	871.935979	22.107656	24500.000000
	Bs.As. G.B.A. Zona Sur	1888.912977	1239.061350	91.135875	23140.764706
	Buenos Aires Costa Atlántica	1961.070956	815.550387	26.262626	21897.058824
	Buenos Aires Interior	1643.059367	506.580200	107.853077	3657.142857
	Capital Federal	3325.388251	3895.143204	49.000000	206333.333333
	Córdoba	1428.786003	532.797780	1.184571	14960.000000
	Santa Fe	2392.245874	4331.372338	0.600000	44425.675676
TOTALES		2693.050228	3208.031692	0.600000	206333.333333

## Estadísticas DESPUES de la corrección

		Media	Desviación Estándar	Precio Minimo	Precio Maximo
Tipo de propiedad	Provincia				
apartment	Bs.As. G.B.A. Zona Norte	2403.372171	833.286992	500.746786	6290.247423
	Bs.As. G.B.A. Zona Oeste	1703.961463	465.171209	523.809524	3900.588235
	Bs.As. G.B.A. Zona Sur	1844.914339	543.091074	504.753000	5542.168675
	Buenos Aires Costa Atlántica	1964.500122	565.428031	522.388060	4651.162791
	Buenos Aires Interior	1670.199894	473.743170	615.384615	3657.142857
	Capital Federal	2678.141516	917.158167	516.304348	6413.043478
	Córdoba	1440.726098	411.233323	504.753000	5694.650192
	Santa Fe	1710.199541	456.932555	516.560132	5990.756364
TOTALES		2302.393373	887.229021	500.746786	6413.043478

El porcentaje de departamentos cuyo 'price\_usd\_per\_m2' es menor a \$500 es de: 0.56 %. En total hay 276 por debajo de los \$500, de los 49484 departamentos

Elegimos como corte departamentos cuyo price\_usd\_per\_m2 sea mayor a \$500 y este dentro de las 2 desviaciones estándares del precio.



# Correcciones | house |:

## Estadísticas ANTES de la corrección

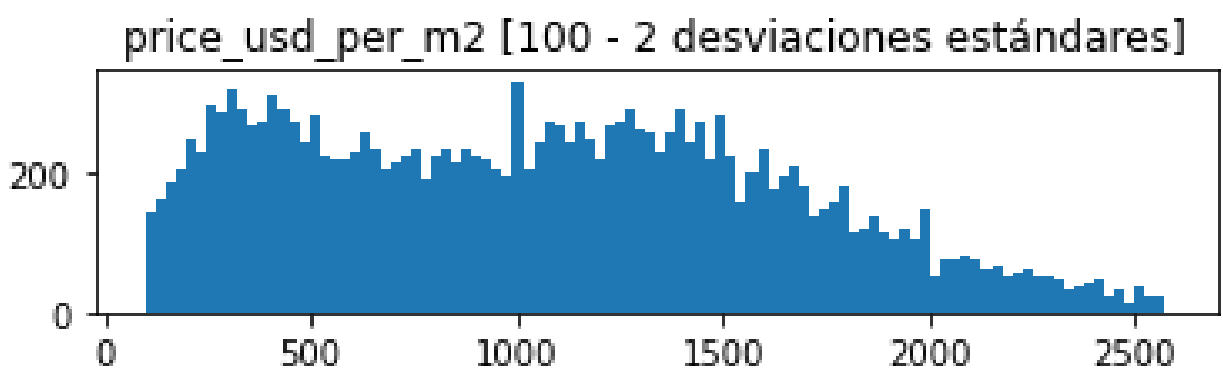
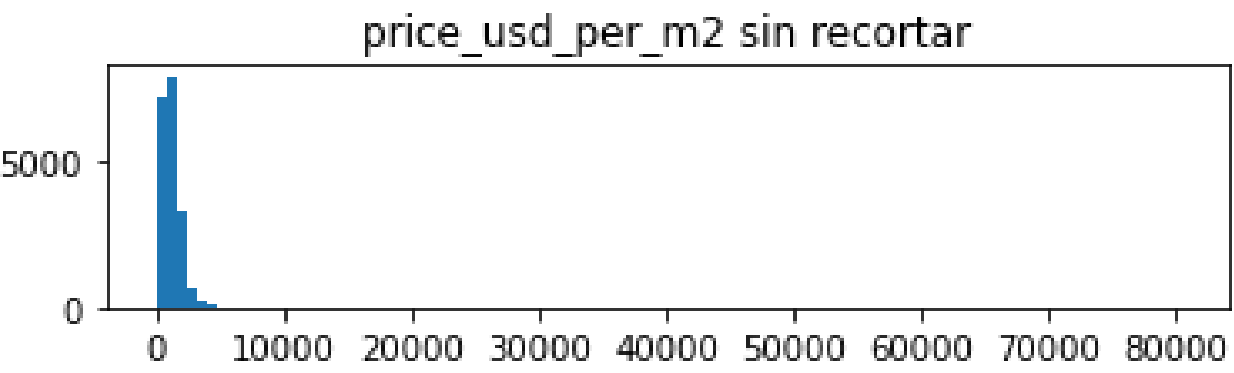
		Media	Desviación Estándar	Precio Minimo	Precio Maximo
Tipo de propiedad	Provincia				
house	Bs.As. G.B.A. Zona Norte	1341.578924	1188.927989	4.166667	38000.000000
	Bs.As. G.B.A. Zona Oeste	1053.891080	1956.782600	7.000000	80000.000000
	Bs.As. G.B.A. Zona Sur	1161.544603	761.060416	19.314341	9411.764706
	Buenos Aires Costa Atlántica	749.953688	534.328605	3.000000	5698.005128
	Buenos Aires Interior	647.793745	558.987525	2.072029	4166.666667
	Capital Federal	1546.437925	1658.398601	190.000000	23064.516129
	Córdoba	714.335569	777.596681	12.254902	16666.666667
	Santa Fe	992.574612	2107.895565	37.916667	36549.707237
TOTALES		1176.186030	1286.090049	2.072029	80000.000000

## Estadísticas DESPUES de la corrección

		Media	Desviación Estándar	Precio Minimo	Precio Maximo
Tipo de propiedad	Provincia				
house	Bs.As. G.B.A. Zona Norte	1208.231810	577.439813	100.950618	2571.428571
	Bs.As. G.B.A. Zona Oeste	918.238873	519.898963	101.267427	2549.019608
	Bs.As. G.B.A. Zona Sur	1109.268234	568.544206	100.950533	2555.555556
	Buenos Aires Costa Atlántica	748.313865	478.035027	102.857143	2542.372881
	Buenos Aires Interior	689.148025	506.608839	102.500000	2500.000000
	Capital Federal	1235.301316	548.759605	190.000000	2564.102564
	Córdoba	681.498916	452.993973	100.463679	2416.666667
	Santa Fe	784.967876	516.336001	110.184267	2543.859649
TOTALES		1060.563800	582.778768	100.463679	2571.428571


El porcentaje de casas cuyo 'price\_usd\_per\_m2' es menor a \$100 es de: 0.83 %. En total hay 304 por debajo de los \$100, de las 36654 casas

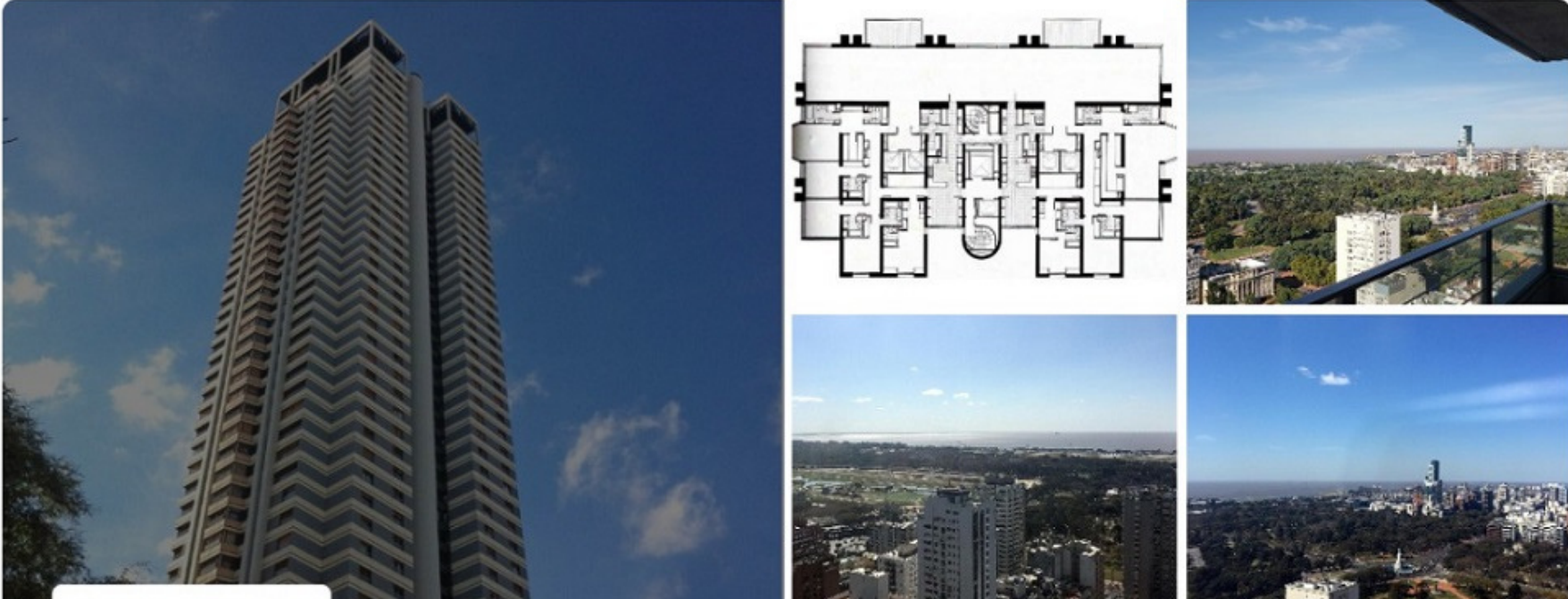
Elegimos como corte casas cuyo price\_usd\_per\_m2 sea mayor a \$100 y este dentro de las 2 desviaciones estándares del precio.








# Se encontraron problemas con los datos de los portales de búsqueda de inmuebles

zonaprop Comprar ▾ Alquilar ▾ Temporal ▾ Emprendimientos Servicios ▾ Publicar 



 Fotos 




**USD 1.111.111**  
\$ 140.000 Expensas




**Migueletes al 1100**  
Las Cañitas, Palermo


🏠 370 m<sup>2</sup> 📏 370 m<sup>2</sup> 🛏 5 amb. 🛎 4 dorm. 🚿 3 baños 🚗 4 coch.

**Departamento - Piso - Las Cañitas - Torre - Seguridad - 4 Coch - 370 m - Amenities - Vista**

Venta piso alto en torre migueletes plazaSuper luminoso. Imponente vista a la Ciudad y al Río de la Plat...



   **Contactar**




**USD 1.111.111**  
\$ 75.000 Expensas




**Av. Libertador al 5100**  
Belgrano, Capital Federal

🏠 257 m<sup>2</sup> 📏 257 m<sup>2</sup> 🛏 5 amb. 🛎 4 dorm. 🚿 3 baños 🚗 2 coch.

**Hermoso Sempiso - Style Libertador - 5 Amb - 260 m2 - 2 Coch - Seg 24 Hs - Full Amen.**

Contacto: Ver datoscihermosisimo sempiso en style libertador - vistas al verdePalier privado. Comodo ...



   **Contactar**

- Muchos datos de distintos tipos
- Datos faltantes
- Datos inusuales y llamativos.
- Datos repetidos
- Datos mal cargados
- Distintos tipos de monedas

# Investigacion - Definición de umbrales

Hablamos con especialistas e investigamos en internet

## Casas

- Tamaño maximo: 2000 m2
- Tamaño minimo: 30 m2

## Departamento

- Tamaño maximo: 800 m2
- Tamaño minimo: 16 m2

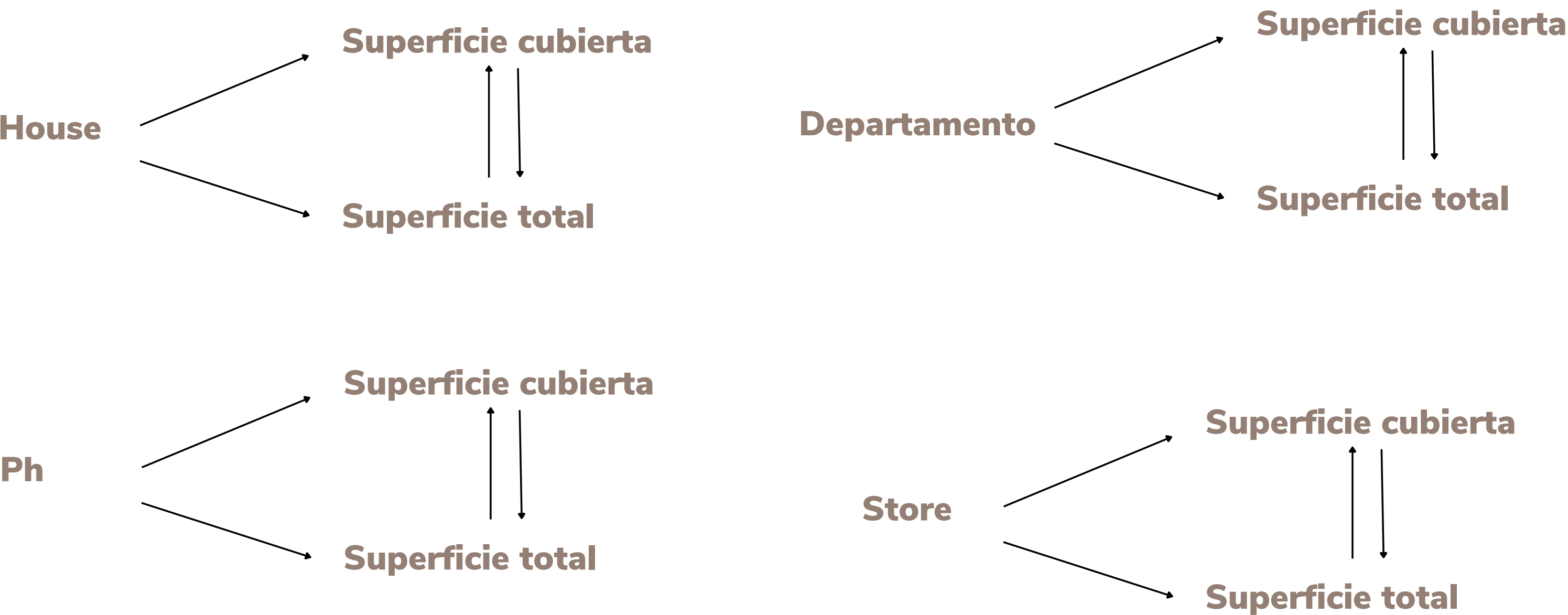
## Ph

- Tamaño maximo: pendiente...
- Tamaño minimo: pendiente...

## Store

- Tamaño maximo: pendiente...
- Tamaño minimo: pendiente...

# Limpieza e imputacion



# Tareas pendientes

- Variables Dummies:
  - State\_name.
  - Property\_type
- Regex description → rooms.
- Mas graficos

