



ORGANIZACIÓN DE DATOS

Trabajo Práctico 1

ANÁLISIS EXPLORATORIO: ZONAPROP MÉXICO

Prof Corrector:

INTEGRANTE	PADRÓN	E-MAIL
Taiel Colavecchia	102510	tcolavecchia@fi.uba.ar
Santiago Klein	102192	sklein@fi.uba.ar
Yuhong Huang	102146	yhuang@fi.uba.ar

GRUPO 10

Índice

Índice	1
Introducción	3
Link al repositorio	3
Procesamiento y análisis inicial	4
Adaptación del DataFrame	4
Cambios del nombre de dato:	4
Porcentajes de datos inválidos	4
Un vistazo a los tipos de propiedades:	6
Cantidad de propiedades de cada tipo	6
Precio de propiedades por tipo	8
Análisis de los posibles 'amenities'	10
Cantidad de propiedades de cada tipo con cada 'amenity'	10
Cantidad de propiedades de cada tipo con cada 'amenity'	11
Precio de cada tipo dado un 'amenity'	14
Proporción de cada tipo por 'amenity'	15
Análisis del precio por metros	16
Análisis de la distribución de precios	16
¿Cómo influye en el precio la cercanía a escuelas?	18
Promedio de precio por metros con escuelas cercanas	18
Promedio de precio/m ² de distintos tipos de propiedades en función de la cercanía a escuelas	19
¿Cómo influye en el precio la cercanía a centros comerciales?	20
Promedio de precio/m ² en inmuebles con centros comerciales cercanos	20
Promedio de precio/m ² de distintos tipos de propiedades en función de la cercanía a centros comerciales	21
¿Cómo varían los precios de los distintos tipos de inmuebles según la antigüedad?	
24	
¿Cómo influye en el precio de los apartamentos según la antigüedad?	24
¿Cómo influye en el precio de las casas según la antigüedad?	25
¿Cómo influye en el precio de las casas en condominio según la antigüedad?	26
¿Cómo es la composición de los inmuebles según su año de construcción?	27
Análisis según cantidad de habitaciones, baños y garages	28
Análisis por provincia	35

Análisis por ciudad	38
Top 10 ciudades más caras del país	38
Top 10 ciudades más baratas del país	39
Precio/m ² de las ciudades de las provincias con mayor publicaciones de ventas	41
Los precios promedios de propiedades de las ciudades de Distrito Federal	41
Promedio de precio/m ² de propiedades de las ciudades del Edo. de México	43
A partir del gráfico, se puede observar que las ciudades más caras no están concentradas en el centro de la ciudad de México, sino en los suburbios de ella.	
44	
Promedio de precio/m ² de propiedades de las ciudades de Jalisco	45
Promedio de precio/m ² de propiedades de las ciudades de Querétaro	48
Promedio de precio/m ² de propiedades de las ciudades de Nuevo León	50
Comparación de precio/m ² entre las casas y los apartamentos de las ciudades	52
Top 10 ciudades con casas más caras del país	52
Top 10 ciudades con casa más barata del país	52
Top 10 ciudades con apartamentos más caro del país	54
Top 10 ciudades con apartamentos más baratos del país	55
Promedio de precio/m ² de casas de las ciudades de Distrito Federal	57
Promedio de precio/m ² de apartamentos de las ciudades del Distrito Federal	58
Promedio de precio/m ² de casas de las ciudades del Edo. de México	60
Promedio de precio/m ² de apartamentos de las ciudades del Edo. de México	63
Promedio de precio/m ² de casas de las ciudades de Jalisco	65
Promedio de precio/m ² de apartamentos de las ciudades de Jalisco	67
Promedio de precio/m ² de casas de las ciudades de Querétaro	69
Promedio de precio/m ² de apartamentos de las ciudades de Querétaro	70
Promedio de precio/m ² de casas de las ciudades de Nuevo León	72
Promedio de precio/m ² de apartamentos de las ciudades de Nuevo León	73
Conclusión respecto al análisis de ciudades	75
Análisis de las descripciones	76
¿Qué palabras fueron las más utilizadas?	76
Conclusiones	77

Introducción

Este trabajo práctico fue realizado para la materia Organización de Datos de la Facultad de Ingeniería de la Universidad de Buenos Aires.

En el presente informe, se detallarán los resultados obtenidos a partir de un análisis exploratorio del registro histórico de publicaciones del sitio <https://www.zonaprop.com.ar>.

El set de datos incluye el registro de todas las publicaciones realizadas entre los años 2012 y 2016 en México, y puede ser obtenido desde:

<metadata.fundacionsadosky.org.ar/media/navent/train.csv>

El objetivo del trabajo práctico es descubrir cuestiones de interés que presente el set de datos.

Link al repositorio

github.com/TaielC/Fiuba_Datos_TP1

Procesamiento y análisis inicial

Adaptación del DataFrame

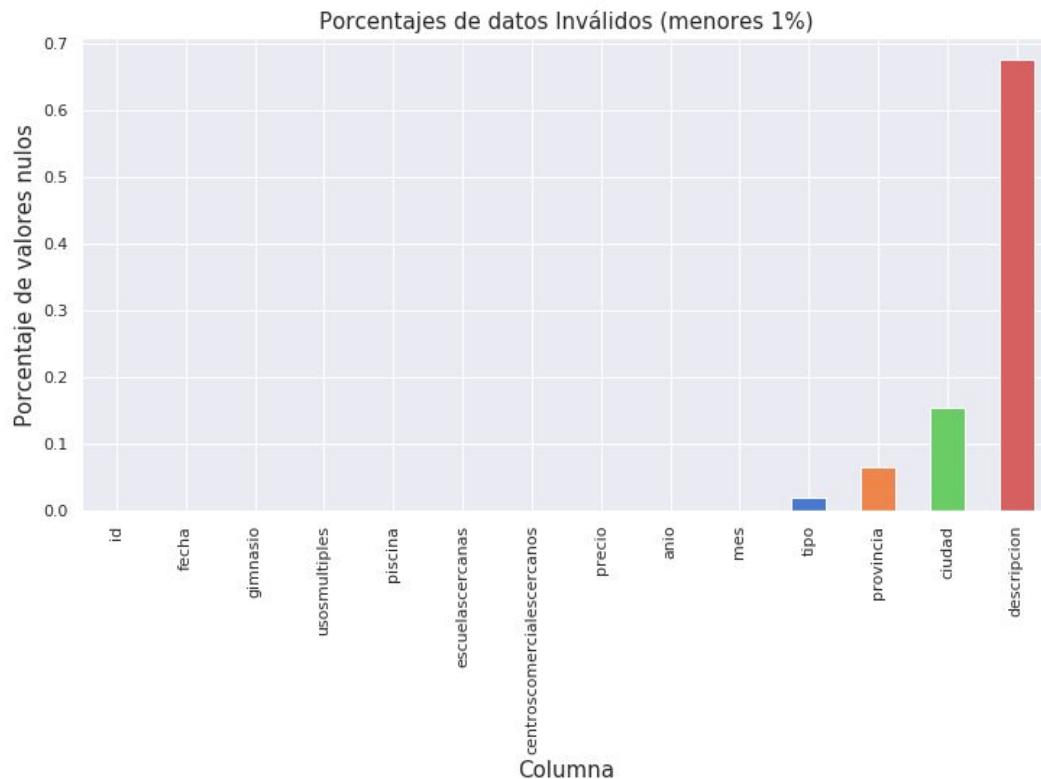
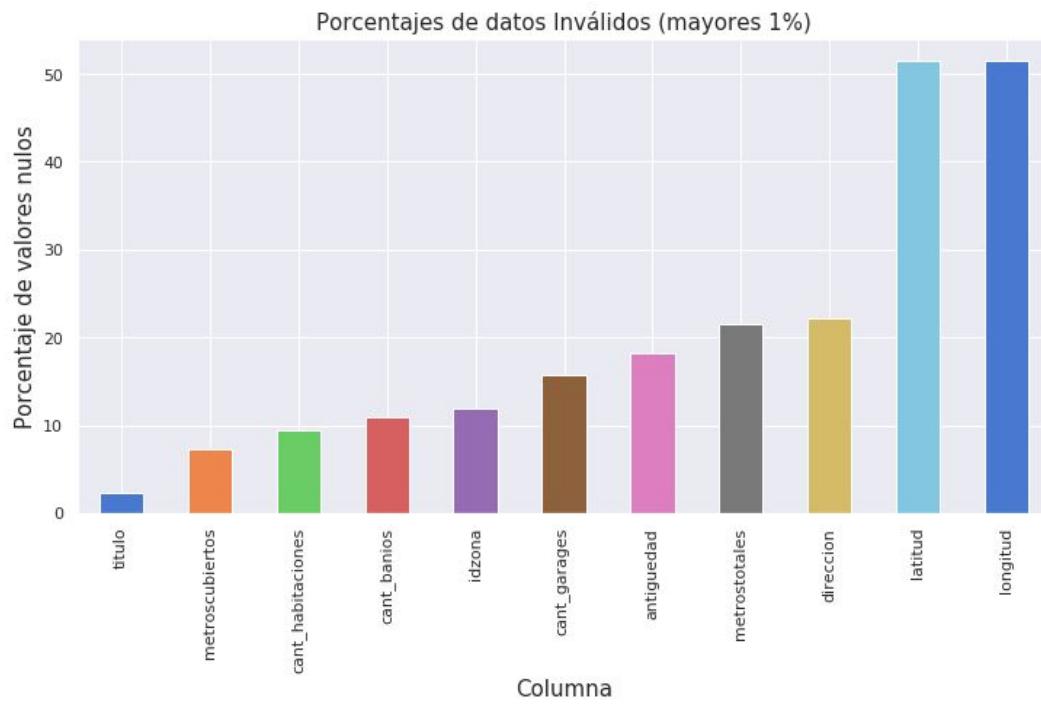
El primer paso fue analizar la forma del dataset, para ver con qué contábamos para efectuar el análisis. Por comodidad, se procedió a realizar cambios del nombre de algunas columnas:

Cambios del nombre de dato:

- “tipodepropiedad” renombrado a: “ tipo”
- “habitaciones” renombrado a: “cant_habiciones”
- “garages” renombrado a: “cant_garages”
- “banos” renombrado a: “cant_banios”
- “lat” renombrado a: “latitud”
- “lng” renombrado a: “longitud”

Porcentajes de datos inválidos

Se realizó un análisis para revisar y graficar el porcentaje de datos nulos de cada columna del set de datos. Fue separado en dos gráficos según el porcentaje obtenido: el primero mostrando aquellas que tienen un porcentaje de datos nulos mayor al 1% y el segundo las que tienen un porcentaje menor, de esta manera podemos notar las que no tienen datos nulos, y así, enfocar nuestro análisis en aquellos datos consistentes, y tener en cuenta a la hora de analizar las respectivas columnas la necesidad de filtrar o no.



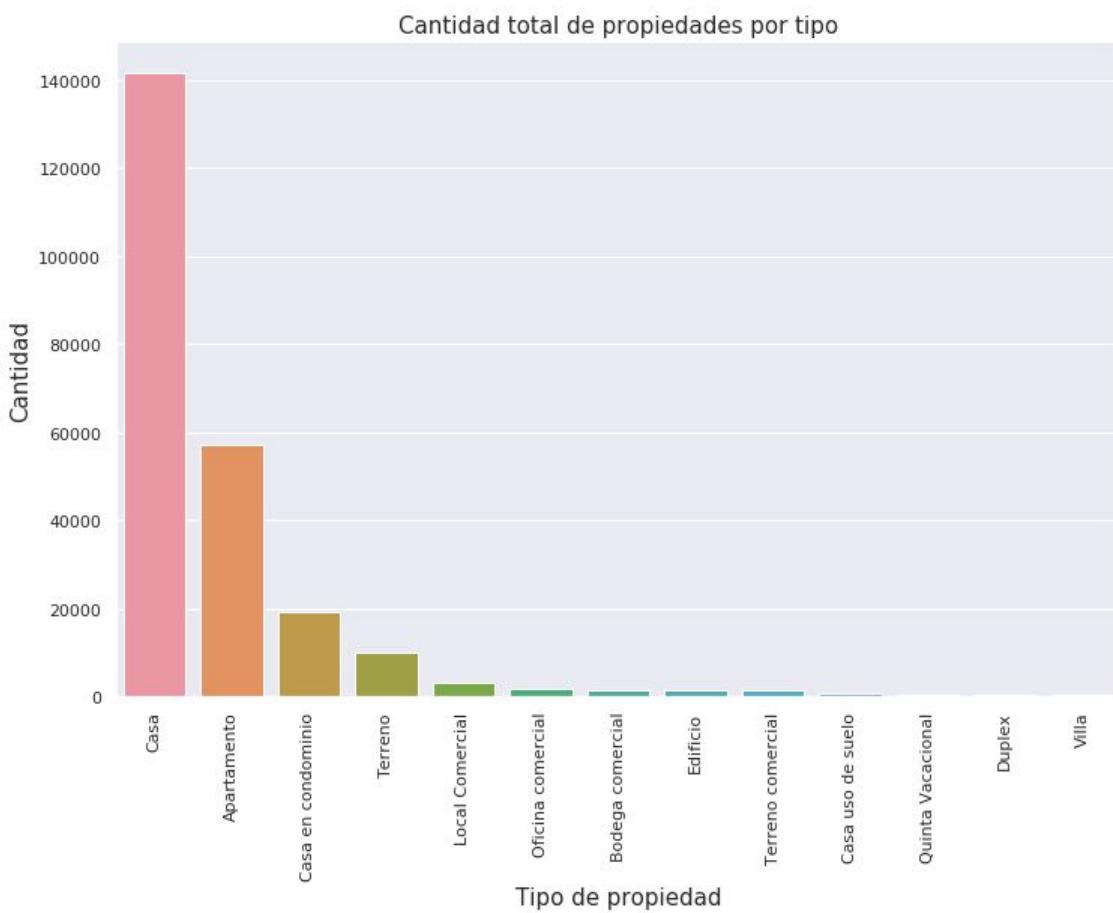
A partir de los gráficos obtenidos, se puede observar que las columnas '**latitud**' y '**longitud**' presentan un porcentaje muy alto de incompletitud (superior al 50%), por lo que se optó por apartarlas del análisis. Aunque presenta un menor porcentaje, se adoptó el mismo criterio para la '**dirección**'.

En los otros casos, se adoptaron diferentes criterios según el análisis en cuestión, preponderando la eliminación de los datos corruptos si se tratara de un porcentaje despreciable ante el volumen de datos total. Un caso particular es el de los '**metros cubiertos**' y '**metros totales**', en los que se optó por intercalar los valores en el caso de que alguna de éstas falte.

Un vistazo a los tipos de propiedades:

Cantidad de propiedades de cada tipo

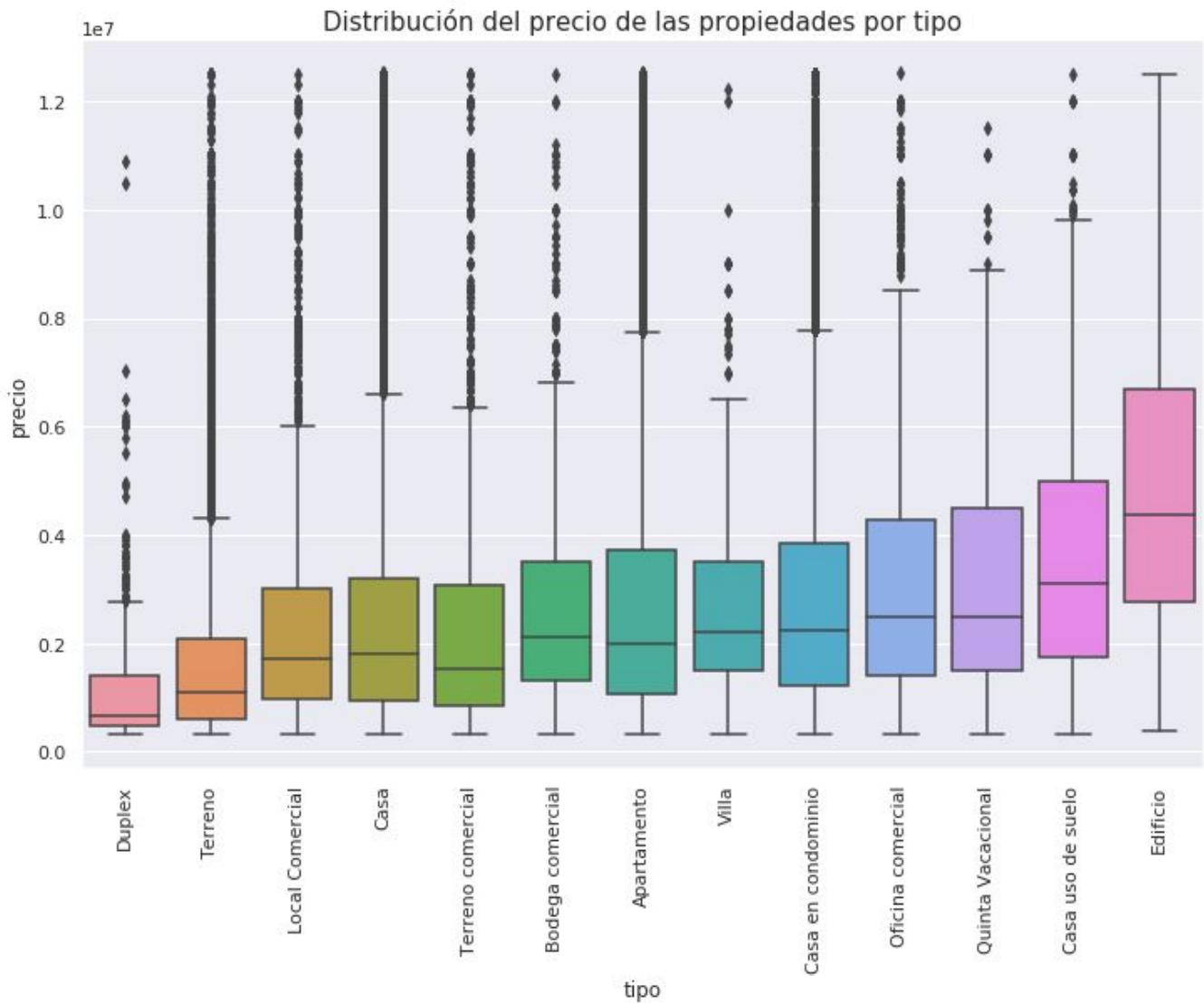
Empezamos por verificar qué tantas propiedades hay de cada tipo y cómo se comportan con el precio de forma general.



En este primer gráfico podemos ver cómo se distribuyen la cantidad de propiedades en los diferentes tipos. Se eliminaron las categorías que se considera que no tienen suficientes datos para ser analizadas y que no aportan al set de datos:

TIPO	CANTIDAD	TIPO	CANTIDAD
Terreno comercial	1326	Departamento Compartido	141
Casa uso de suelo	708	Otros	134
Quinta Vacacional	395	Nave industrial	76
Duplex	343	Terreno industrial	31
Villa	340	Huerta	20
Inmuebles productivos urbanos	200	Lote	5
Rancho	170	Hospedaje	1
Local en centro comercial	165	Garage	1

Así, los datos que pertenecieran a un tipo que tuviera menos de 1000 elementos fueron eliminados.

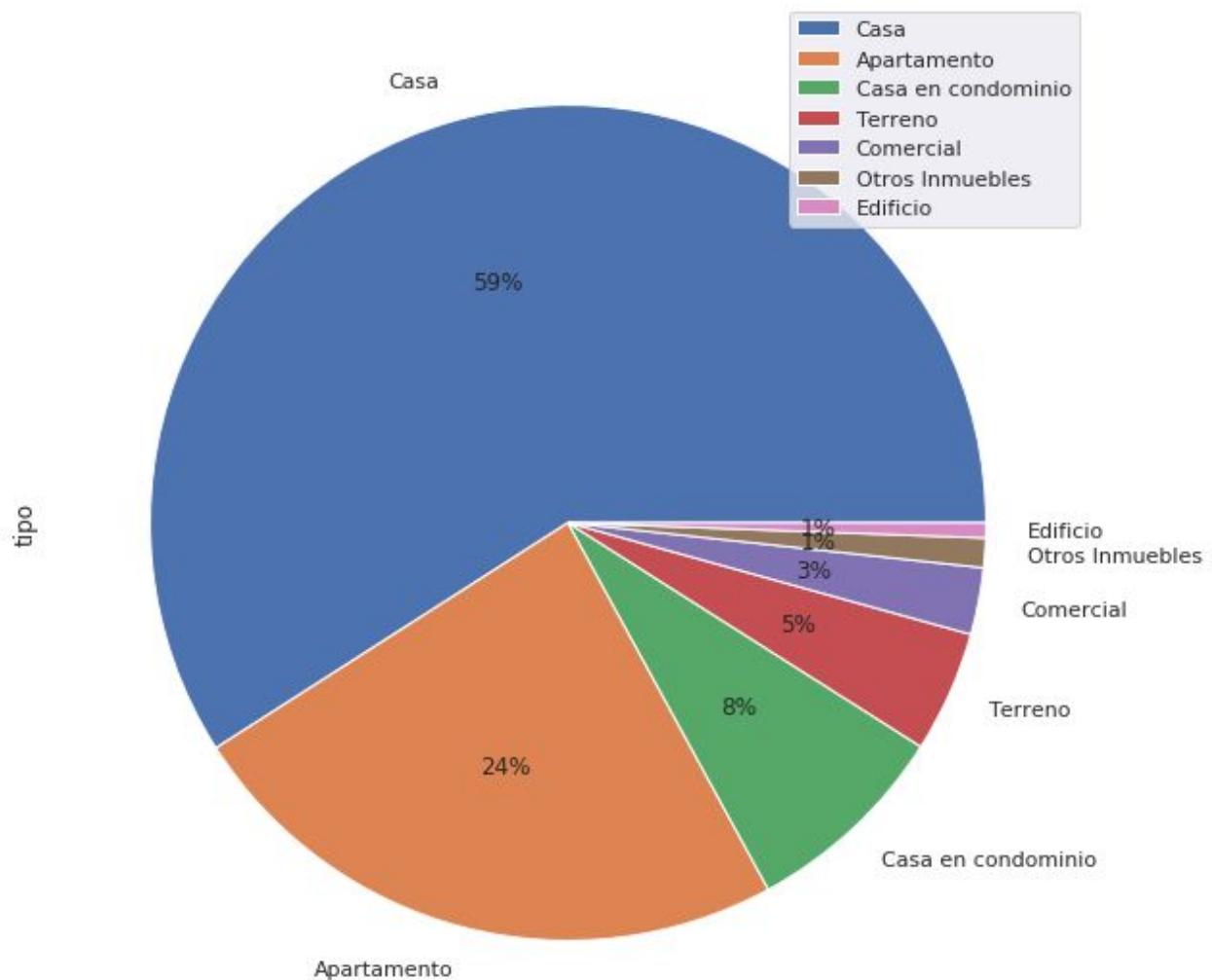
Precio de propiedades por tipo

A partir de este gráfico en combinación con el anterior, se decidió agrupar los tipos que tuvieran cantidades de datos menos relevantes según su precio. De esta manera, quedaron individualmente las 3 categorías que tenían la mayor cantidad de datos: '**Casa**', '**Apartamento**', y '**Casa en condominio**', y la que tenía precios generalmente distinguidamente más altos que el resto '**Edificio**'. Y el resto fueron agrupados de la siguiente manera:

- 'Comercial' contiene a: 'Local Comercial', 'Bodega comercial', 'Oficina comercial'
- 'Terreno' contiene a: 'Terreno', 'Terreno comercial'

De esta manera se obtuvo una mejor distribución de los tipos de propiedades para ser analizadas por separado.

Distribucion de los tipos de propiedades en porcentaje

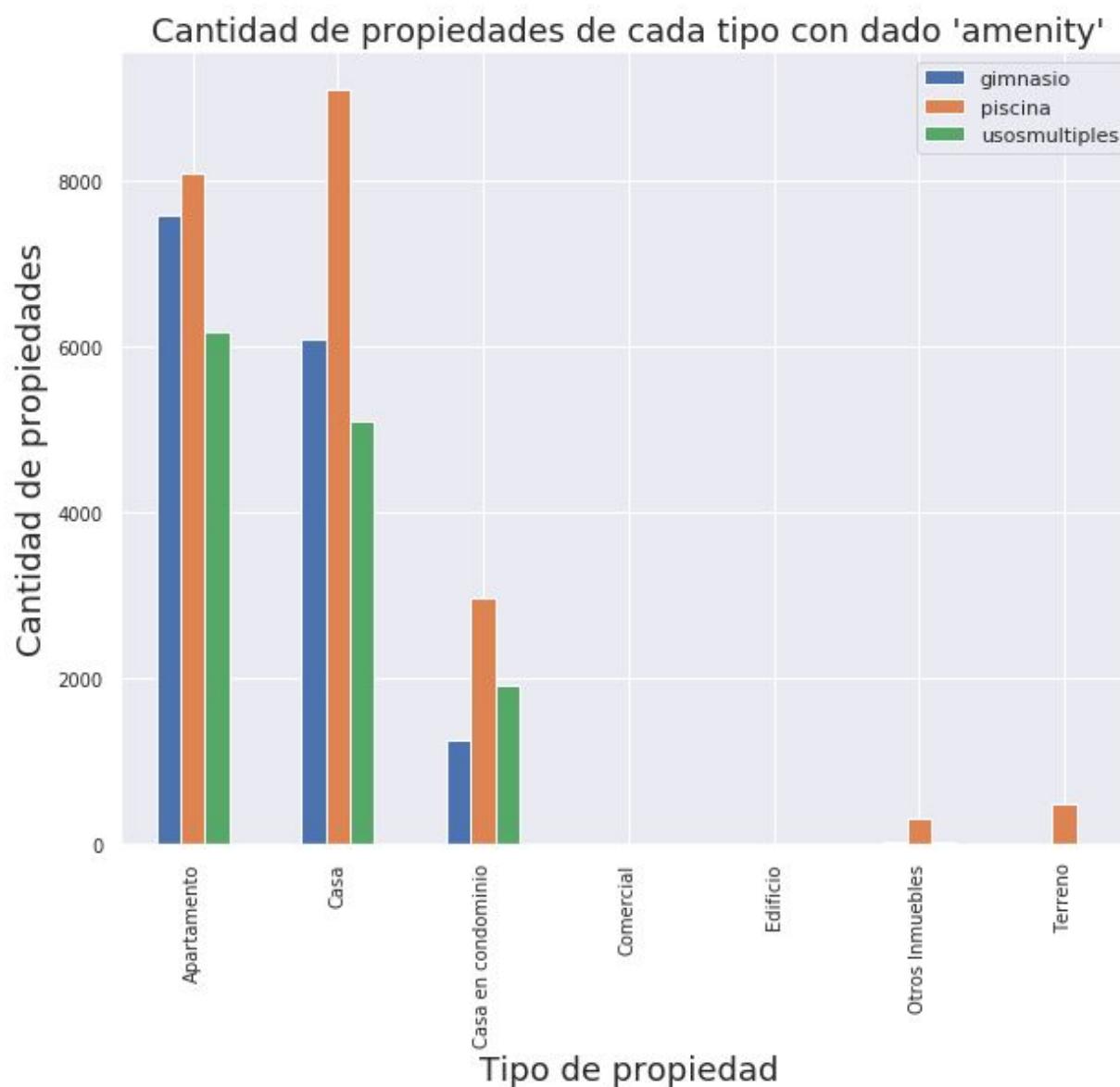


Análisis de los posibles 'amenities'

Llamamos 'amenity' a el agregado de 'piscina', 'gimnasio' o salón de 'usos múltiples'.

Cantidad de propiedades de cada tipo con cada 'amenity'

Se graficó para cada tipo de propiedad la cantidad que tienen cada 'amenity'.



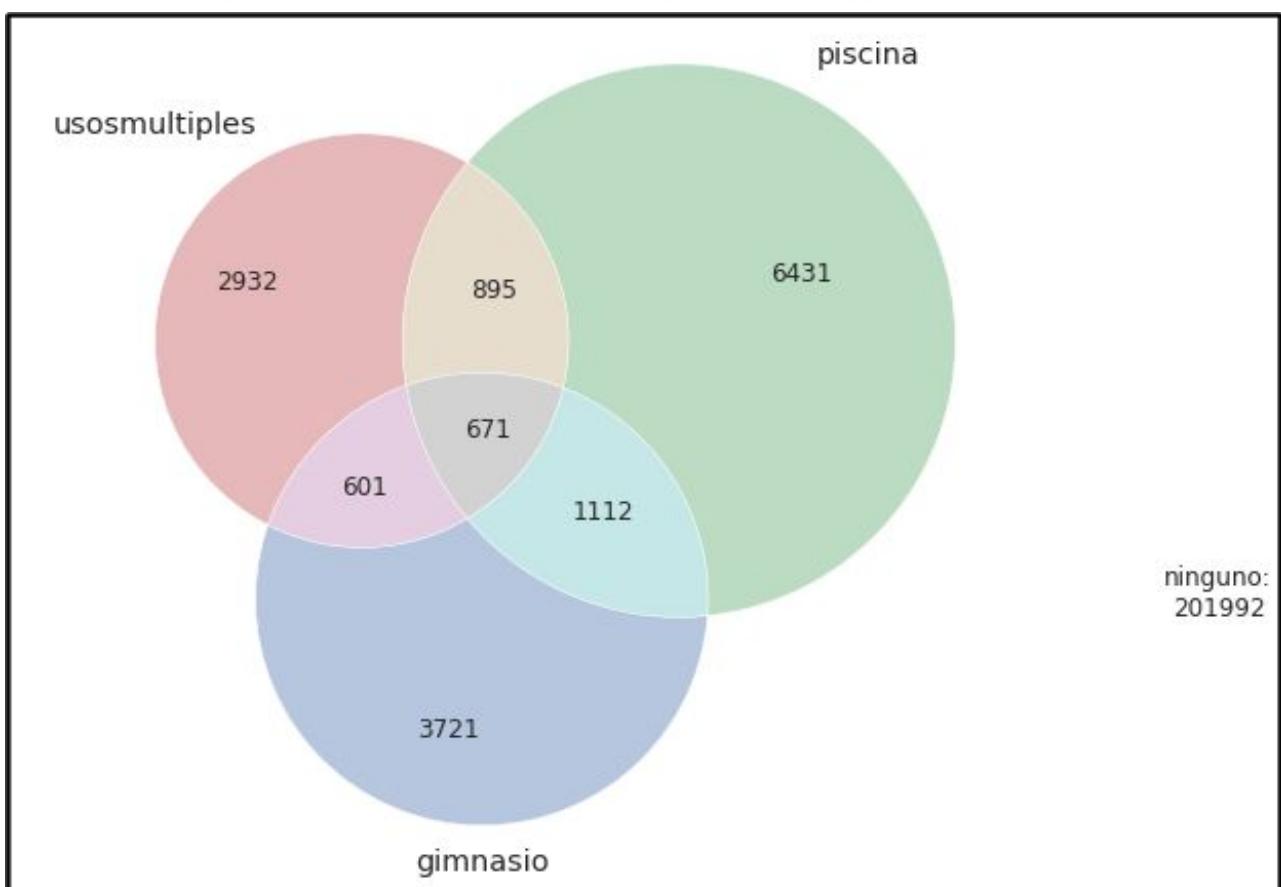
Para los siguientes análisis se eliminaron aquellos tipos que tuvieran menos de 500 propiedades con todos los 'amenities'. Esto deja únicamente para el análisis los 3 tipos de propiedades más abundantes, es decir: '**Apartamento**', '**Casa**' y '**Casa en condominio**'.

Cantidad de propiedades de cada tipo con cada 'amenity'

Se realizó un diagrama de conjuntos (Venn) para cada tipo principal que muestra la cantidad de propiedades que tienen cada amenity. En estos se puede notar, para cada tipo de propiedad, la cantidad de propiedades que tienen cada amenity.

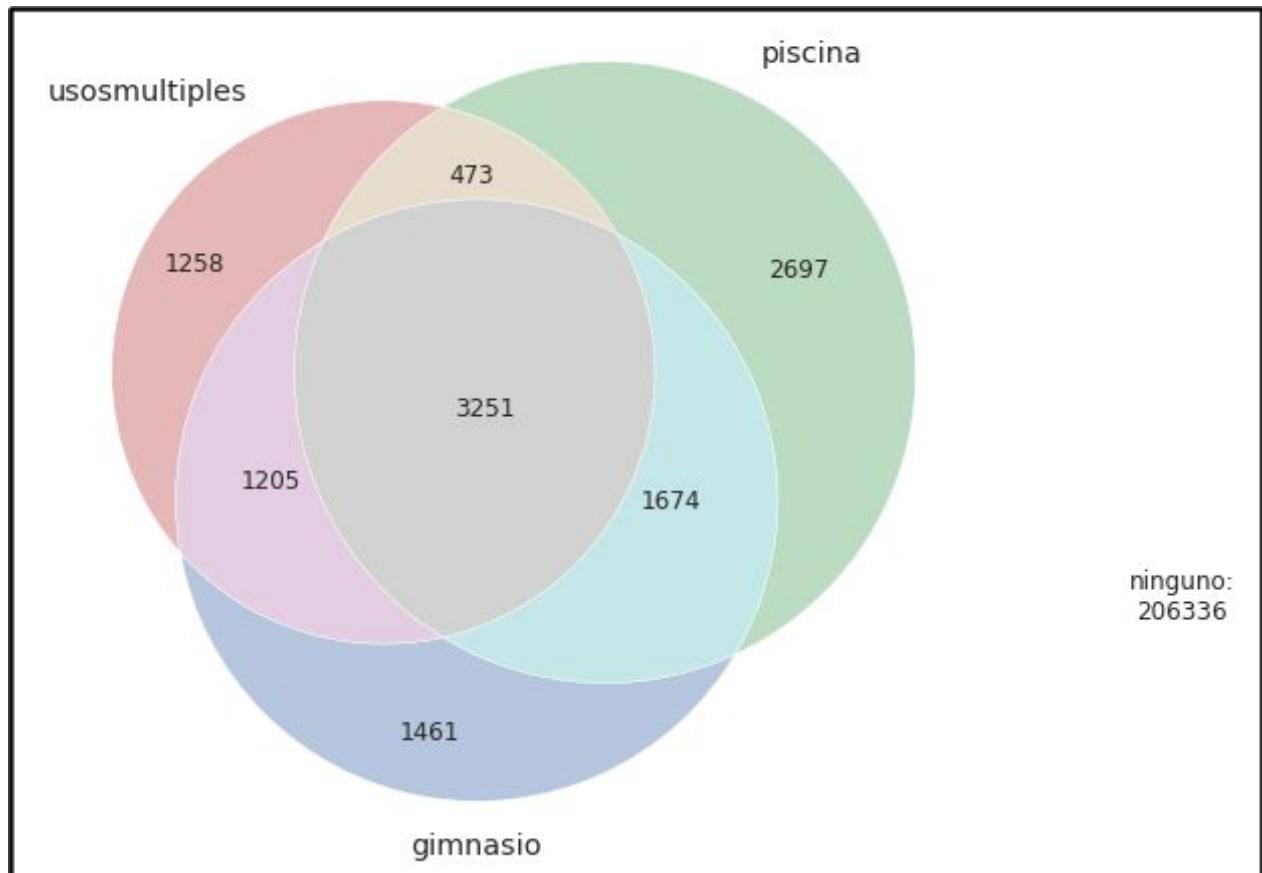
- Total con amenities de '**Casa**': 16363

Diagrama de Venn de Casa con cantidades de cada 'amenity'



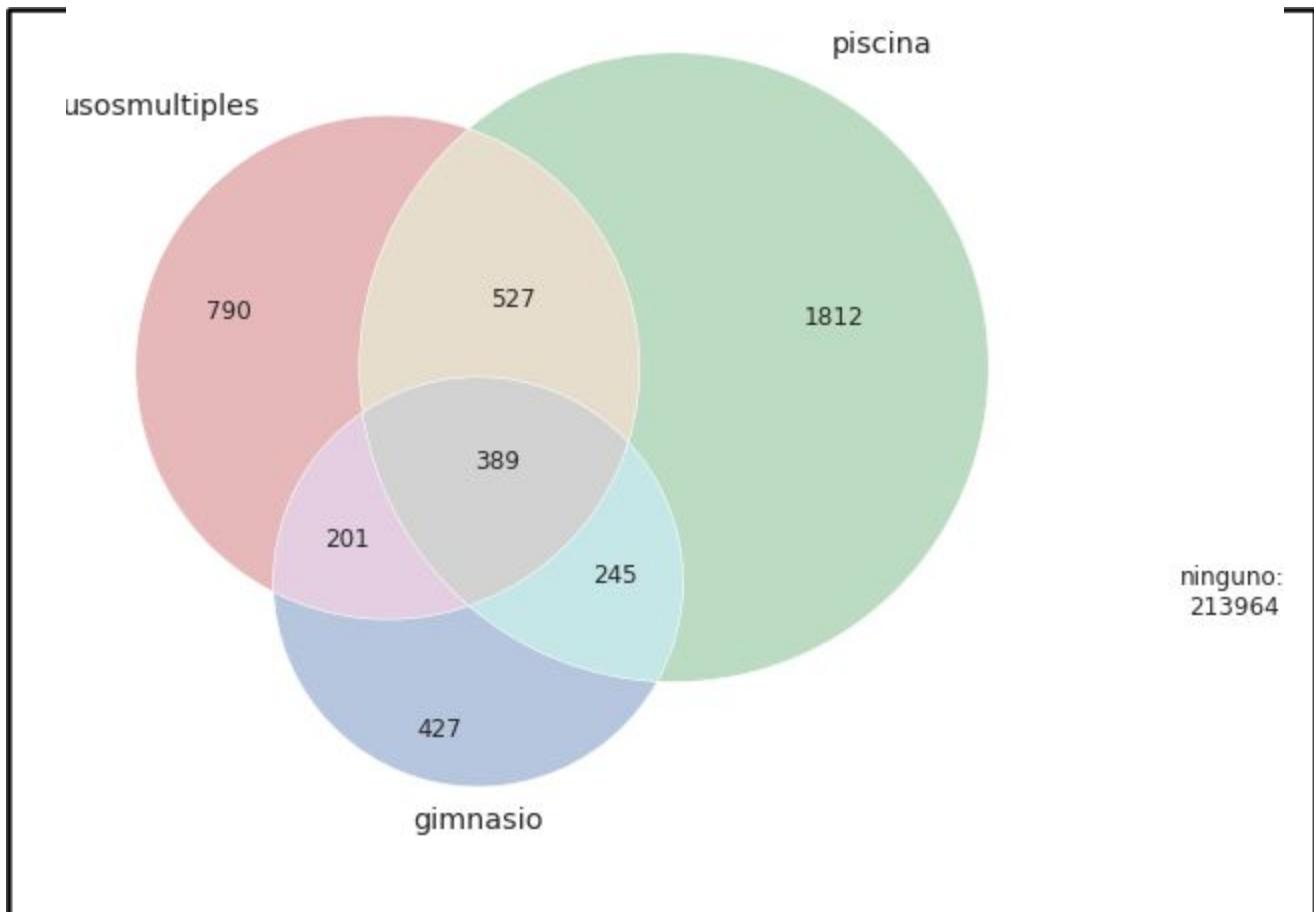
- Total con amenities de 'Apartamento': 12019

Diagrama de Venn de Apartamento con cantidades de cada 'amenity'



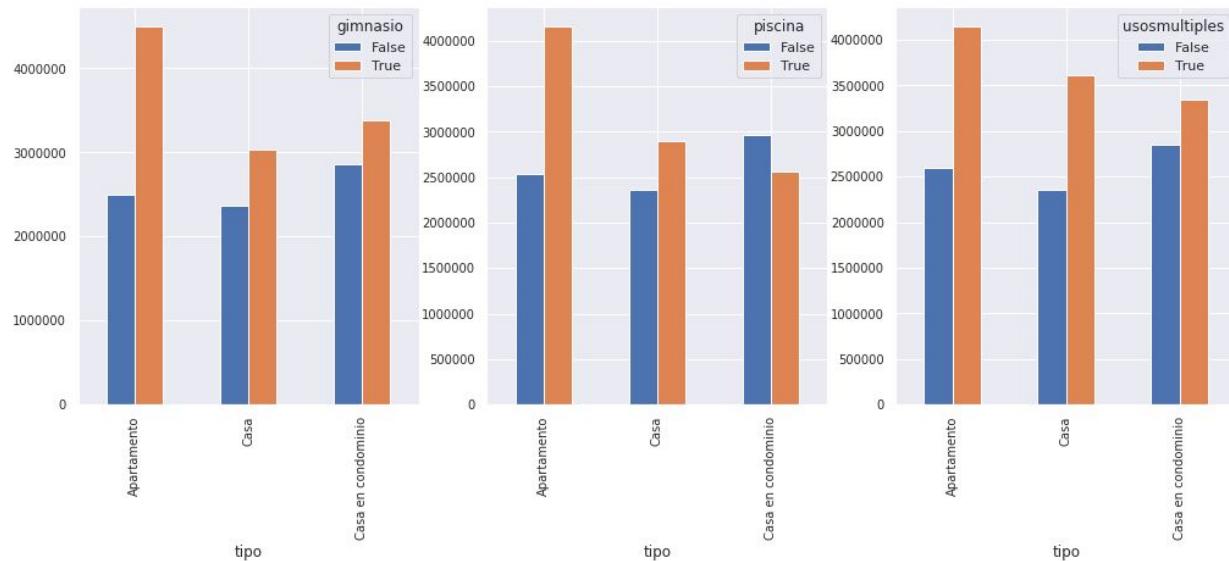
- Total con amenities de 'Casa en condominio': 4391

Diagrama de Venn de Casa en condominio con cantidades de cada 'amenity'



Precio de cada tipo dado un 'amenity'

Se graficó para cada 'amenity' la diferencia de precio promedio para cada tipo (de los principales filtrados anteriormente) dado si la propiedad tiene el 'amenity' o no.



Para la mayor parte de los casos es notable que para cualquier tipo de propiedad el precio promedio es mayor cuando se tiene cualquier 'amenity'.

Proporción de cada tipo por 'amenity'

Se realizó un gráfico con secciones proporcionales a cada tipo que tienen un dato 'amenity'. En este se puede notar qué tipo de propiedades

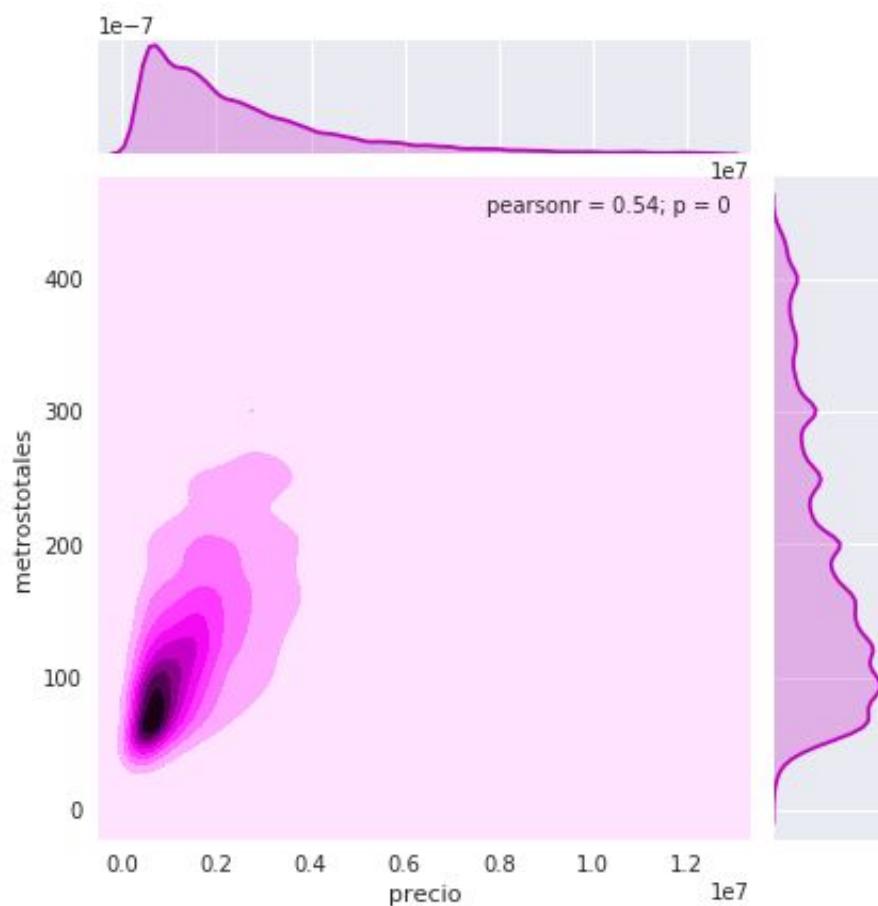
Proporción de cada tipo de propiedad por cada extra



Análisis del precio por metros

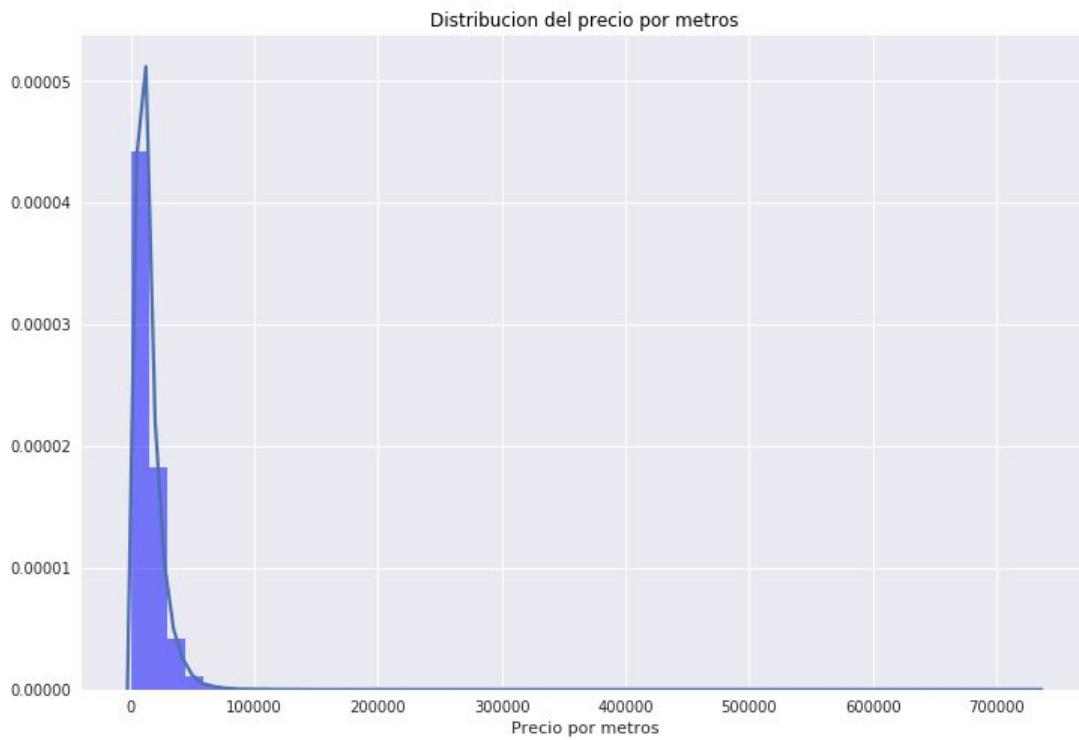
Análisis de la distribución de precios

En principio, se analiza la distribución de precio total y metros totales:

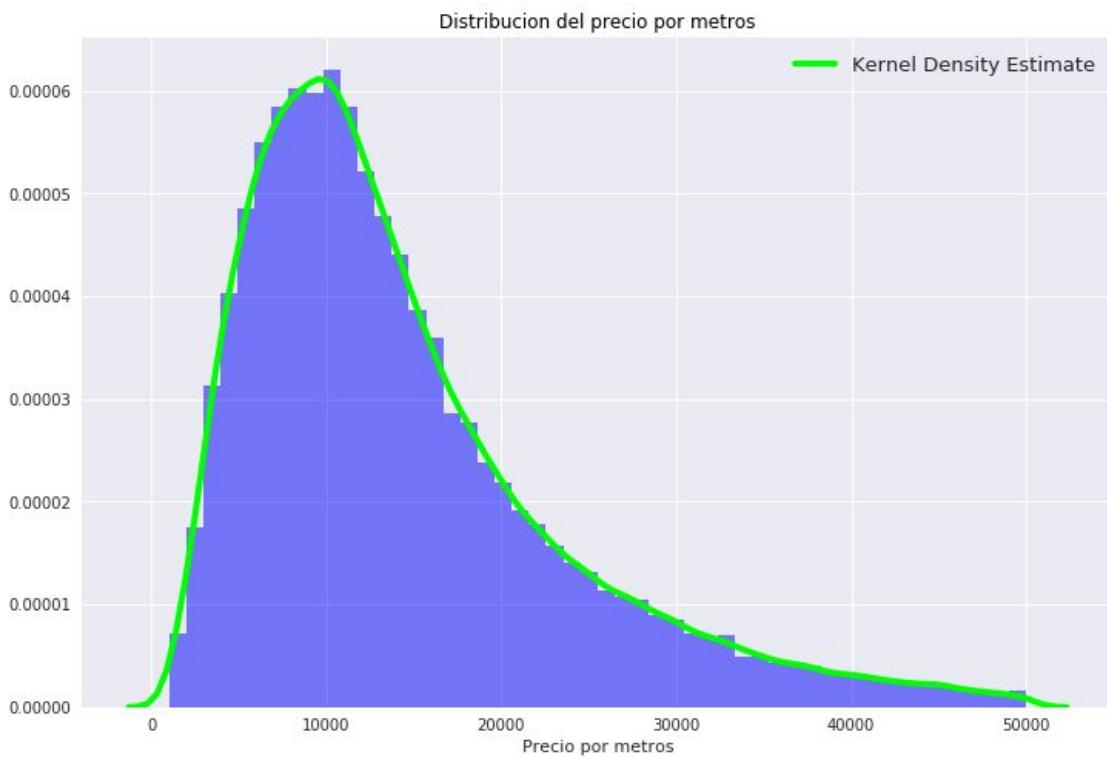


En el gráfico se puede observar el rango de la mayor distribución de `metrostotales` es entre 50m² y 150m², cuyos precios están concentrados entre 1000000 \$MX y 3500000 \$MX.

Una vez filtrado y completado el DataFrame con los precios y los metrostotales, se generó una columna de precio por m², y se puede observar el resultado a partir del gráfico obtenido:



Intervalo de filtrado: [1000,50000] que representa la mayor parte de la distribución.



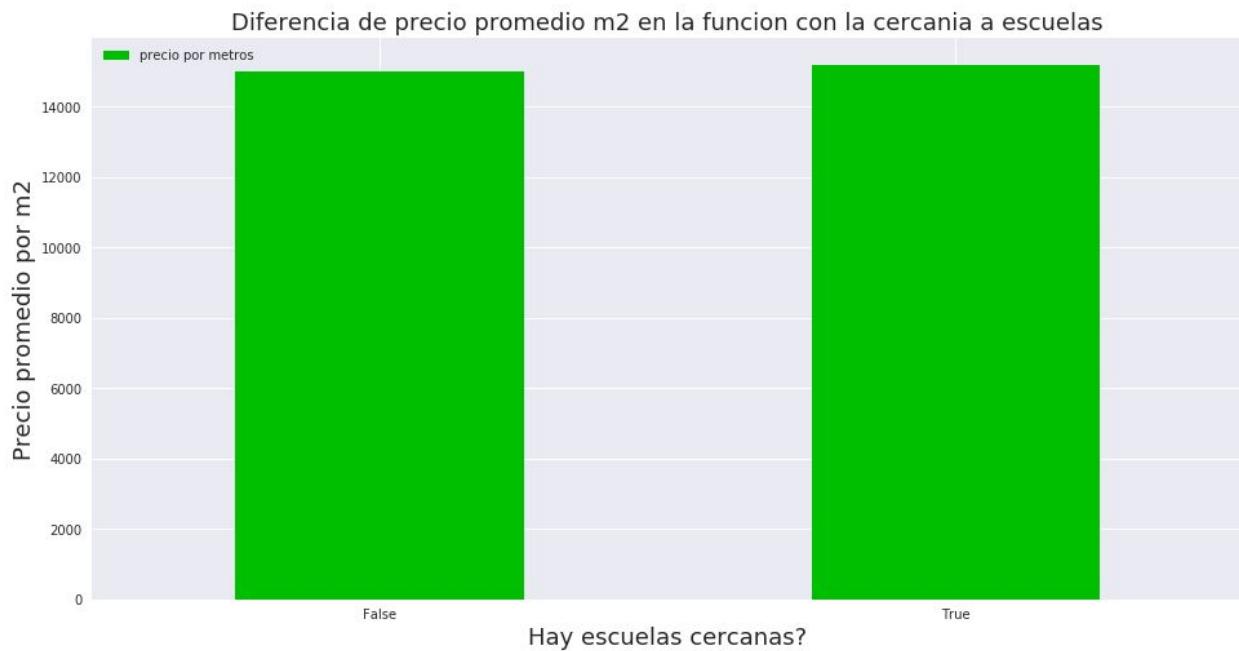
¿Cómo influye en el precio la cercanía a escuelas?

Promedio de precio por metros con escuelas cercanas

Para comenzar, se analizó el precio/m² promedio de todas las publicaciones en función de la cercanía con escuelas, para tener una primera impresión.

Se obtuvo el siguiente resultado:

	Precio/m ² promedio
Con escuelas cercanas	14998.22 \$MX/m ²
Sin escuelas cercanas	15200.69 \$MX/m ²



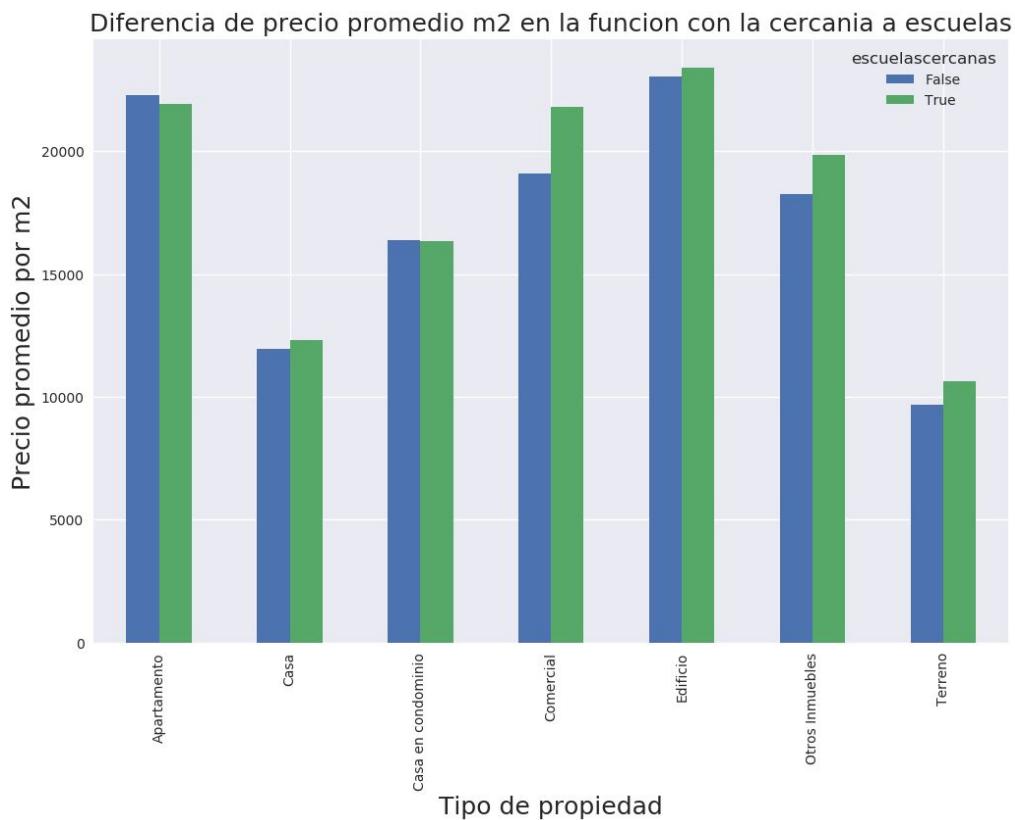
Se puede observar que la diferencia entre con escuelas cercanas y sin escuela cercanas es muy pequeña, teniendo en cuenta todos los tipos de propiedades. A raíz de esto, se analizará en los distintos tipos de interés.

Promedio de precio/m² de distintos tipos de propiedades en función de la cercanía a escuelas

A continuación, se analizó, según el tipo de propiedad, si existe alguna relación entre el precio del inmueble y la cercanía a escuelas.

Se obtuvo el siguiente resultado:

Tipo de propiedad	Precio/m ² (Sin escuelas cercanas)	Precio/m ² (Con escuelas cercanas)
Apartamento	22271.54 \$MX/m ²	21921.81 \$MX/m ²
Casa	11970.51 \$MX/m ²	12314.61 \$MX/m ²
Casa en condominio	16388.73 \$MX/m ²	16338.34 \$MX/m ²
Comercial	19095.22 \$MX/m ²	21816.92 \$MX/m ²
Edificio	23047.13 \$MX/m ²	23396.08 \$MX/m ²
Otros Inmuebles	18248.57 \$MX/m ²	19835.80 \$MX/m ²
Terreno	9694.75 \$MX/m ²	10653.30 \$MX/m ²



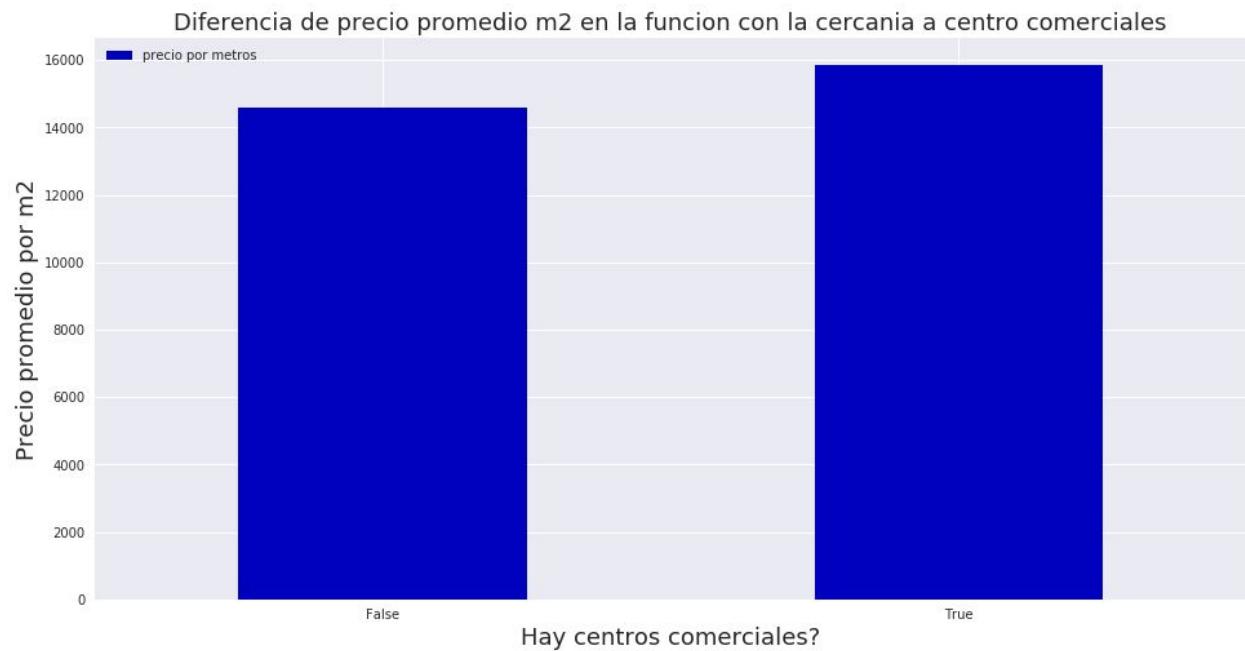
Se puede notar que salvo en los tipos Apartamento y Casa en condominio, hay una clara tendencia creciente en el precio por metros con escuelas cercanas. Se considera que es apreciable que en la mayoría de los tipos haya un aumento del precio si la propiedad tiene escuelas cercanas.

¿Cómo influye en el precio la cercanía a centros comerciales?

Promedio de precio/m² en inmuebles con centros comerciales cercanos

Al principio, se analizó el precio por metros genérico , si existe alguna relación con escuelas cercanas.se obtiene el siguiente resultado:

Centros comerciales cercanos	Precio Por Metros
True	14576.73 \$MX/m ²
False	15866.45 \$MX/m ²



Dichos datos no aportan información útil, con lo cual, a continuación, se analizará según los distintos tipos.

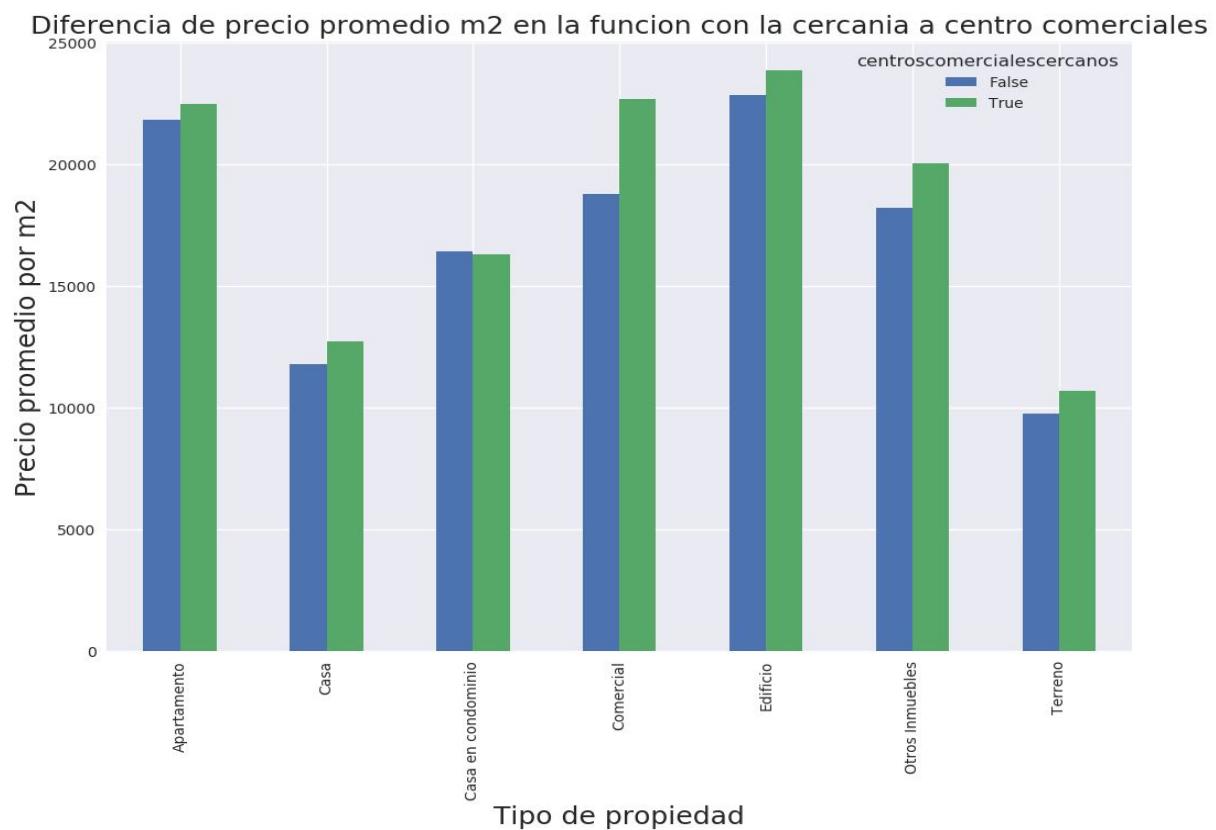
Promedio de precio/m² de distintos tipos de propiedades en función de la cercanía a centros comerciales

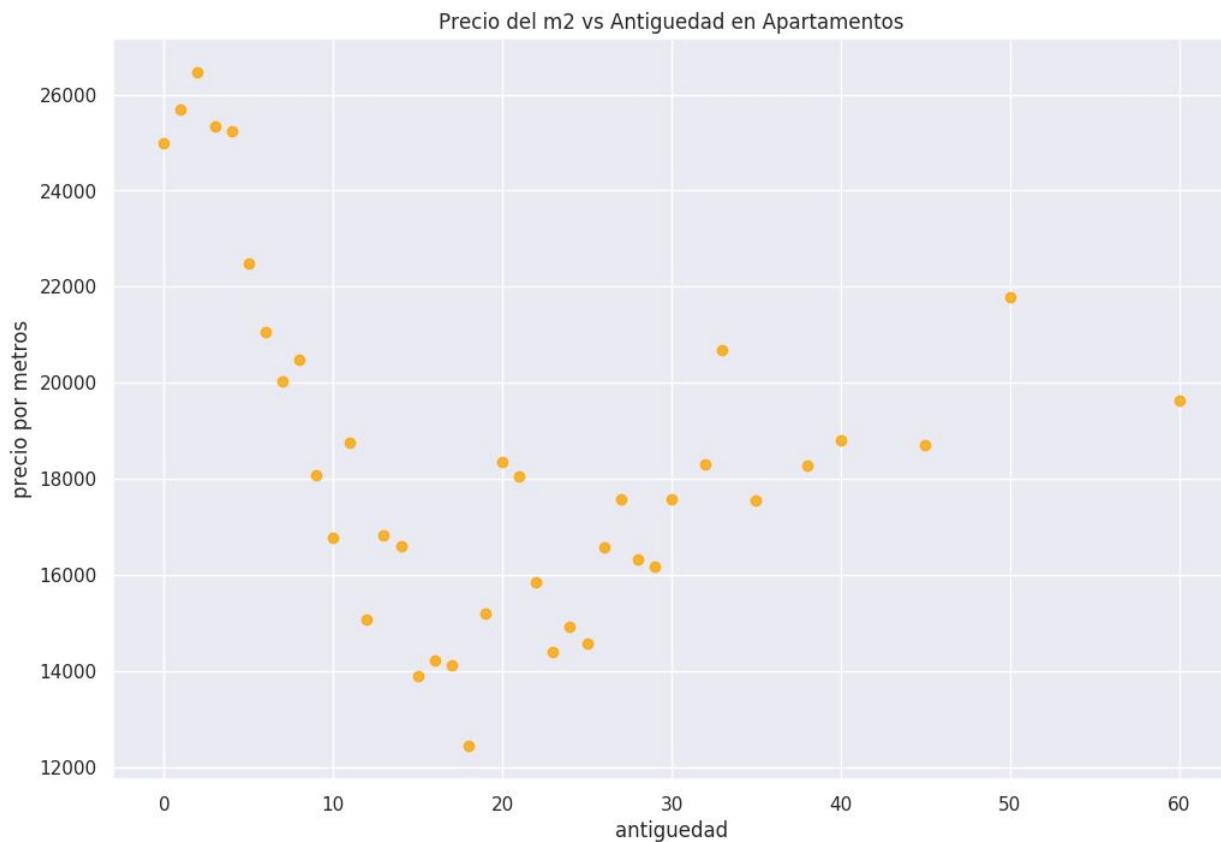
A continuación, se analizó, según el tipo de propiedad, si existe alguna relación entre el precio del inmueble y la cercanía a centros comerciales.

Se obtuvo el siguiente resultado:

Tipo de propiedad	Precio/m ² (Sin escuelas cercanas)	Precio/m ² (Con escuelas cercanas)
Apartamento	21832.81 \$MX/m ²	22449.95 \$MX/m ²
Casa	11770.65 \$MX/m ²	12695.52 \$MX/m ²
Casa en condominio	16416.36 \$MX/m ²	16301.51 \$MX/m ²
Comercial	22841.66 \$MX/m ²	22661.91 \$MX/m ²
Edificio	22841.66 \$MX/m ²	20034.41 \$MX/m ²
Otros Inmuebles	18198.87 \$MX/m ²	20034.41 \$MX/m ²
Terreno	9735.32 \$MX/m ²	10681.24 \$MX/m ²

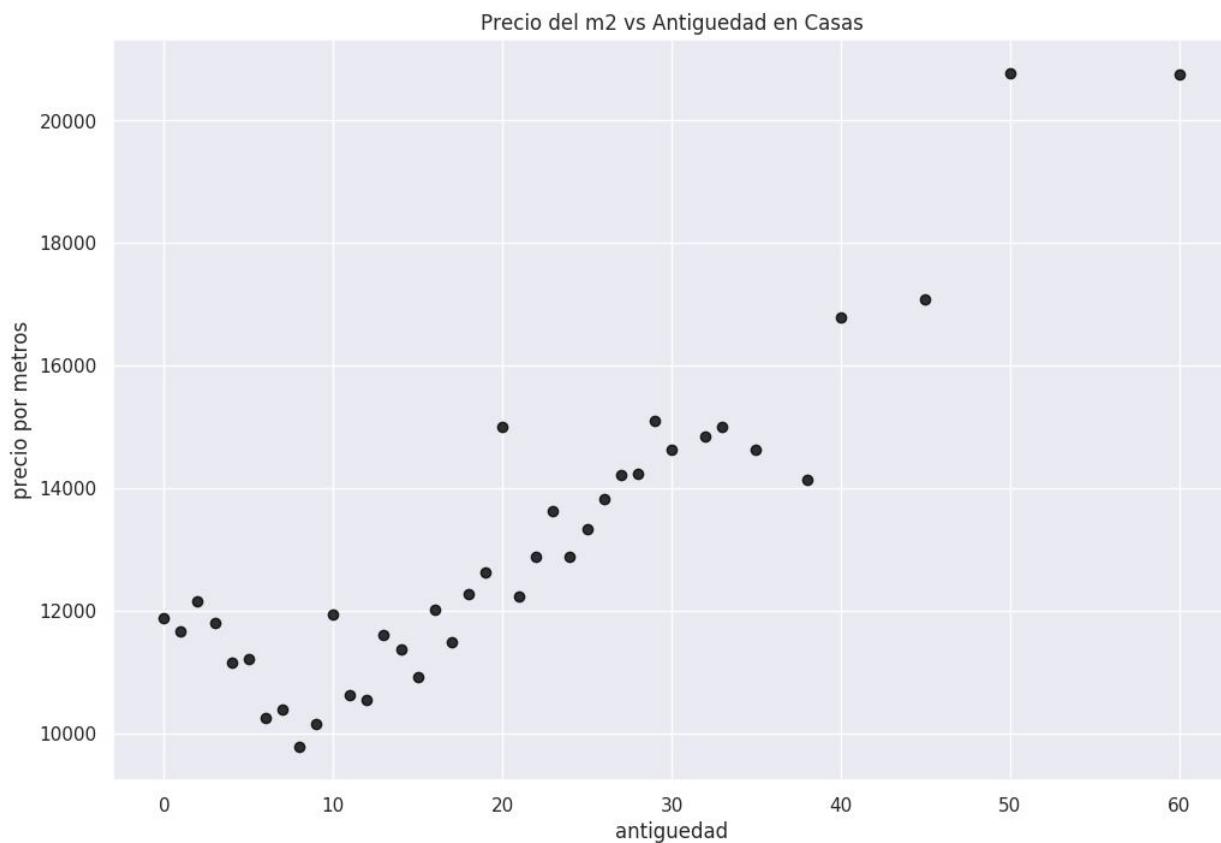
Se puede observar en la mayoría de los tipos que aumenta el precio/m² en aquellos inmuebles cercanos a centros comerciales.



¿Cómo varían los precios de los distintos tipos de inmuebles según la antigüedad?**¿Cómo influye en el precio de los apartamentos según la antigüedad?**

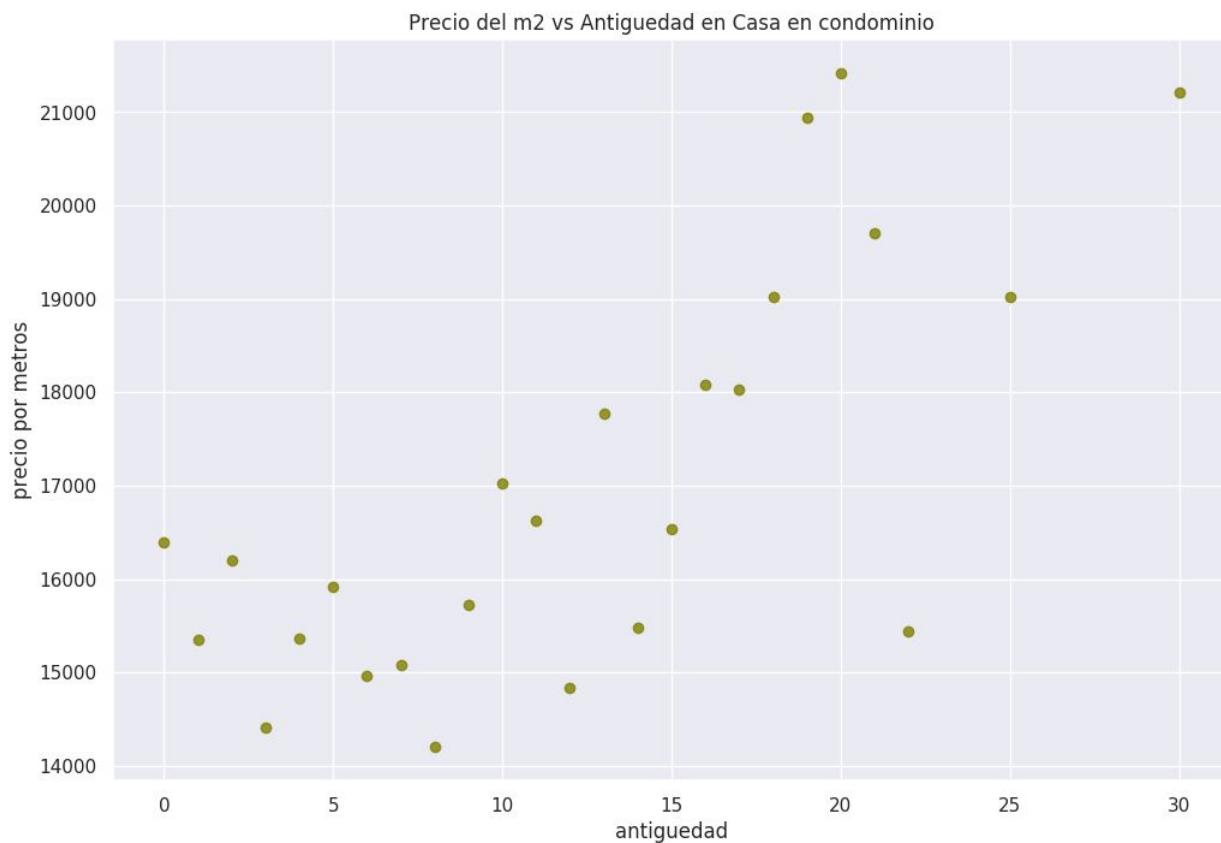
Como se puede observar, el valor promedio del m² de los apartamentos decrece con una tendencia muy clara hasta los 20 años de antigüedad, para luego presentar una distribución menos uniforme y más dispersa. Dicho comportamiento inicial de decrecimiento del precio era el esperado, pues los departamentos añejos tienden a perder valor en el mercado inmobiliario. Sin embargo, la distribución presentada a partir de los 20 años de antigüedad es llamativa, y puede ser atribuida al valor histórico de las mismas.

¿Cómo influye en el precio de las casas según la antigüedad?



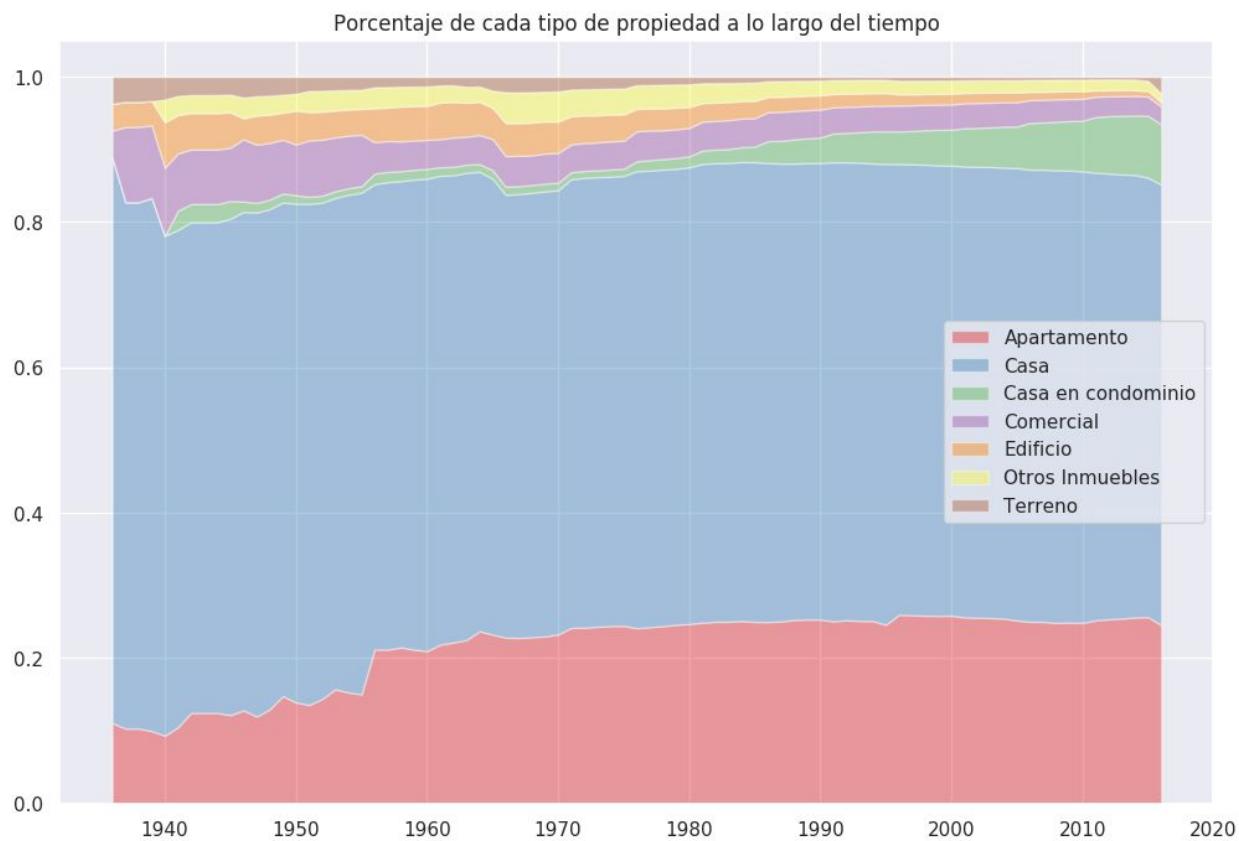
Con respecto a las casas, la evolución del precio por m² en función de la antigüedad es decreciente hasta los 10 años de antigüedad, para luego tornarse creciente. Es curioso nuevamente resaltar cómo cambia la tendencia en cierto punto y se revalorizan los inmuebles, llegando incluso a valores muy altos en aquellos que tienen muchos años de antigüedad (50 y 60), lo cual se ve claramente reflejado en el gráfico. Dicho comportamiento puede ser atribuido nuevamente al valor histórico de las casas, pues es llamativamente superior el precio de las mismas.

¿Cómo influye en el precio de las casas en condominio según la antigüedad?



En el caso de las casas en condominio, el comportamiento es poco claro, y lo único que se puede resaltar del gráfico es el valor de aquellas históricas (con muchos años de antigüedad), que es significativamente superior al resto.

¿Cómo es la composición de los inmuebles según su año de construcción?

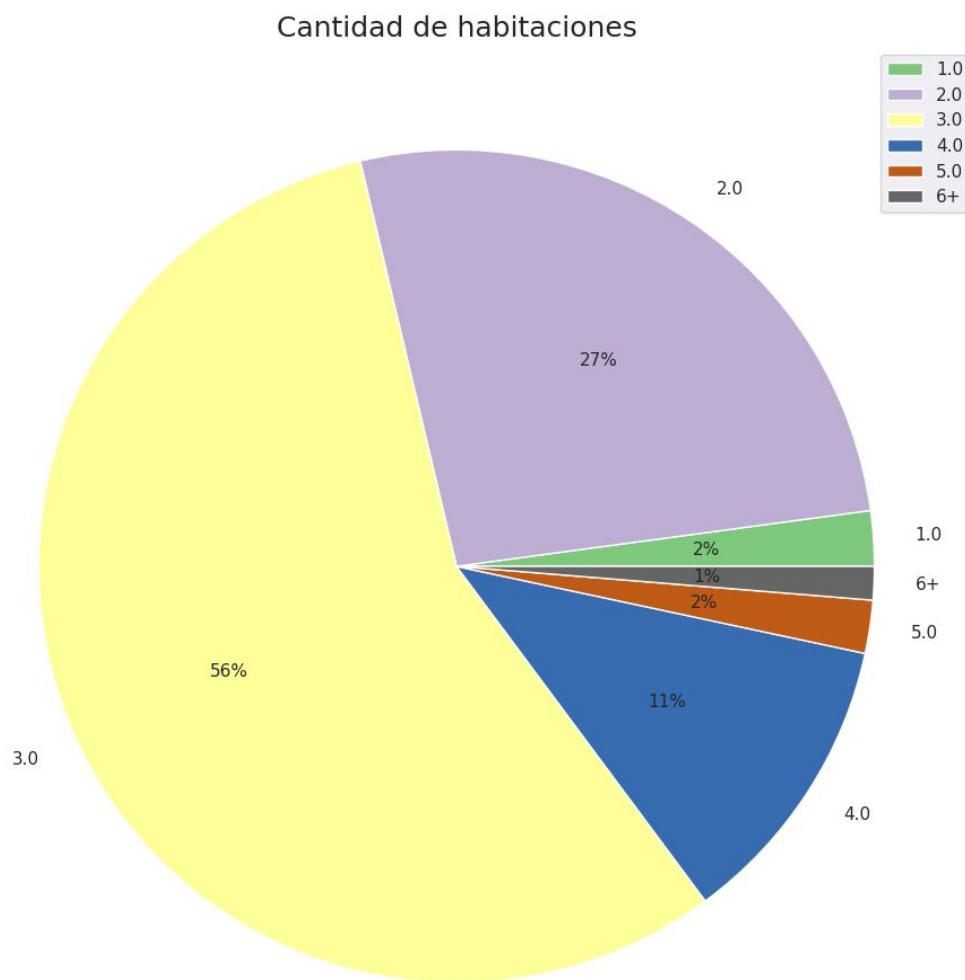


A partir del gráfico, se puede apreciar no sólo el gran volumen que ocupan las publicaciones de apartamentos y casas con respecto del total, sino también la distribución de los mismos según su año de construcción. Es notable el salto en la cantidad de publicaciones de apartamentos construidos a partir de mediados de la década del 60, así como el sostenido crecimiento de la cantidad de publicaciones de casas en condominio construidas a partir del 80. Los comerciales y edificios publicados, por su parte, suelen tener más años de antigüedad, al igual que los terrenos.

Análisis según cantidad de habitaciones, baños y garages

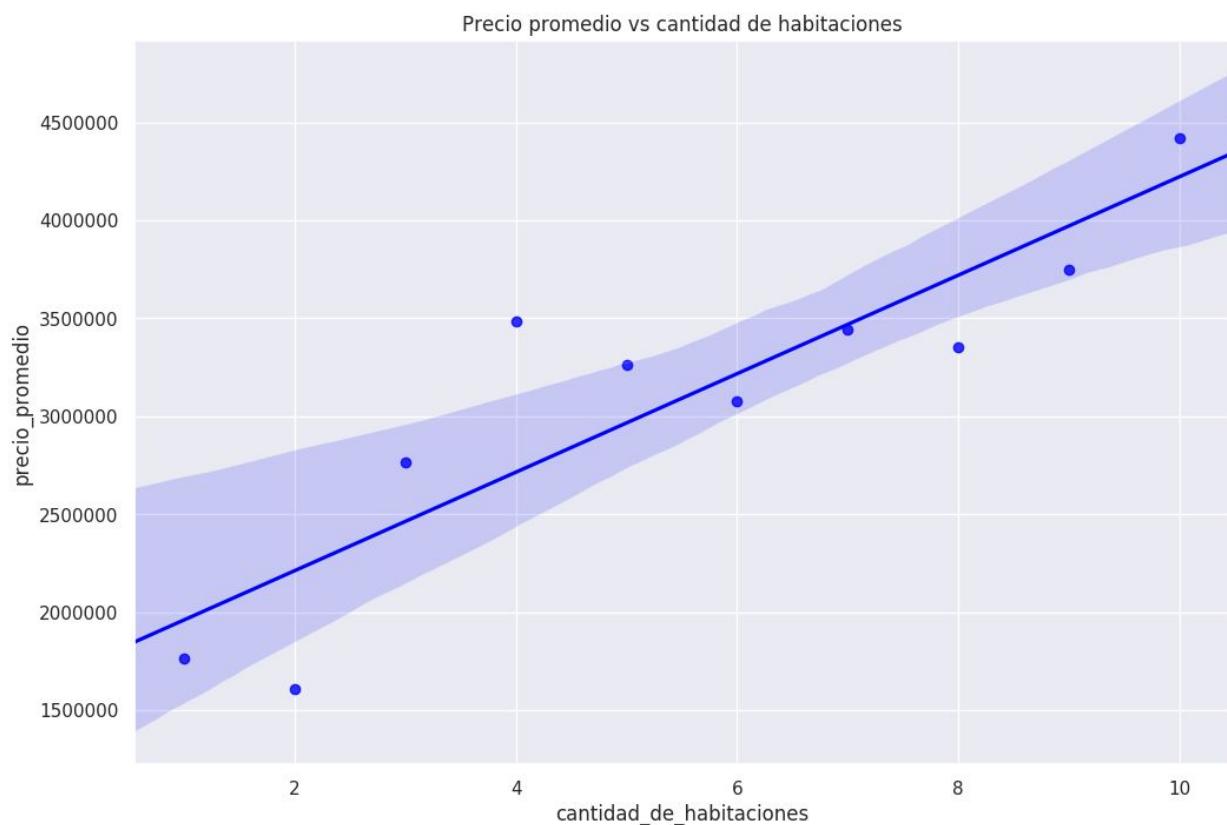
Esta etapa del análisis se realizará sobre los tipos: Casa, Apartamento y Casa en Condominio.

¿Cómo es la distribución según la cantidad de habitaciones?



Se puede observar que más de la mitad de las publicaciones son de inmuebles con 3 habitaciones.

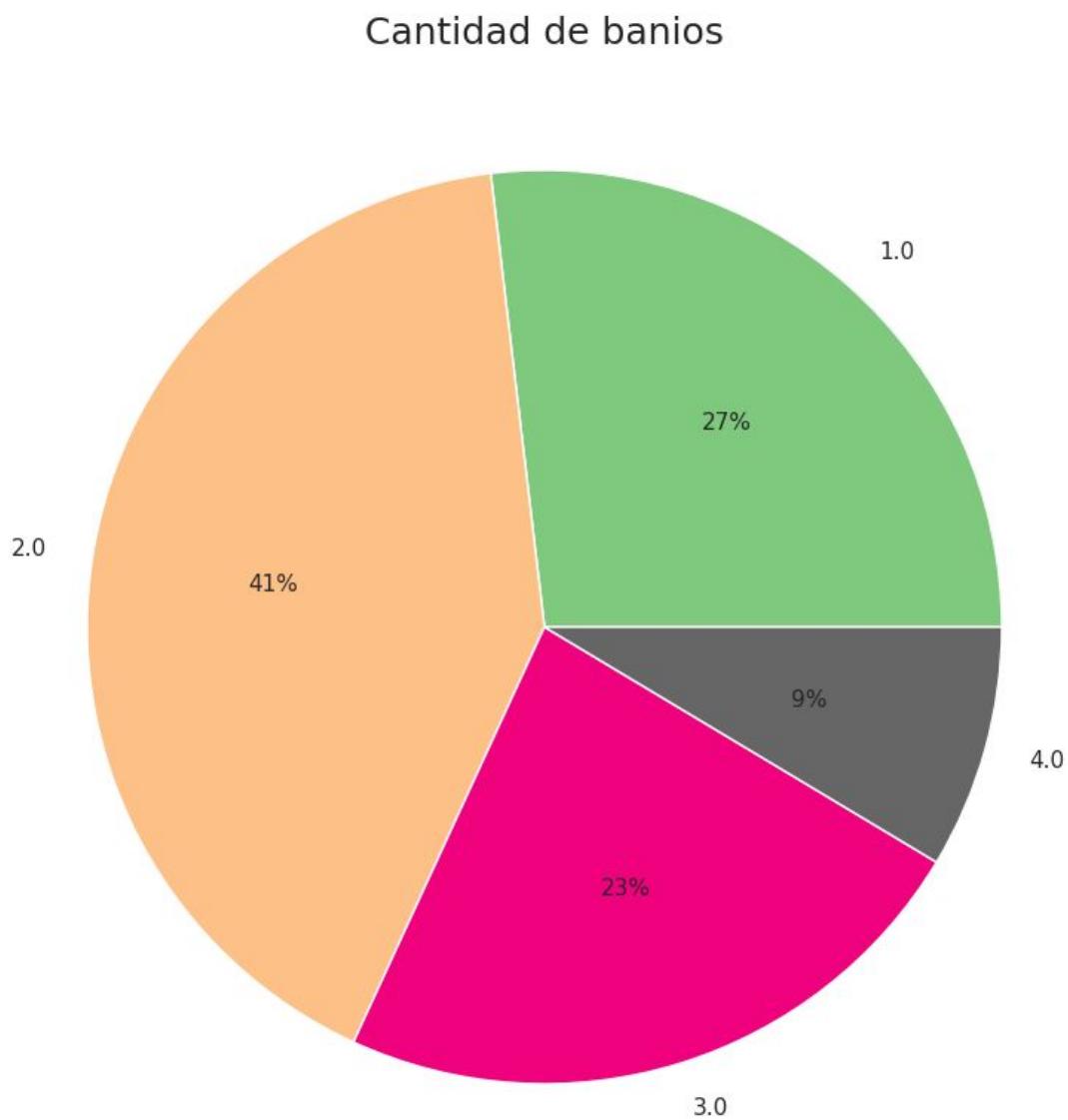
¿Cómo es la variación de precio promedio según la cantidad de habitaciones?



Naturalmente, el precio promedio de los inmuebles aumenta a mayor cantidad de habitaciones. Aquellos inmuebles que cuentan con cuatro habitaciones o más reflejan un incremento menor en su precio promedio, pero también representan una menor parte muestral, como hemos apreciado anteriormente en el gráfico tipo pie.

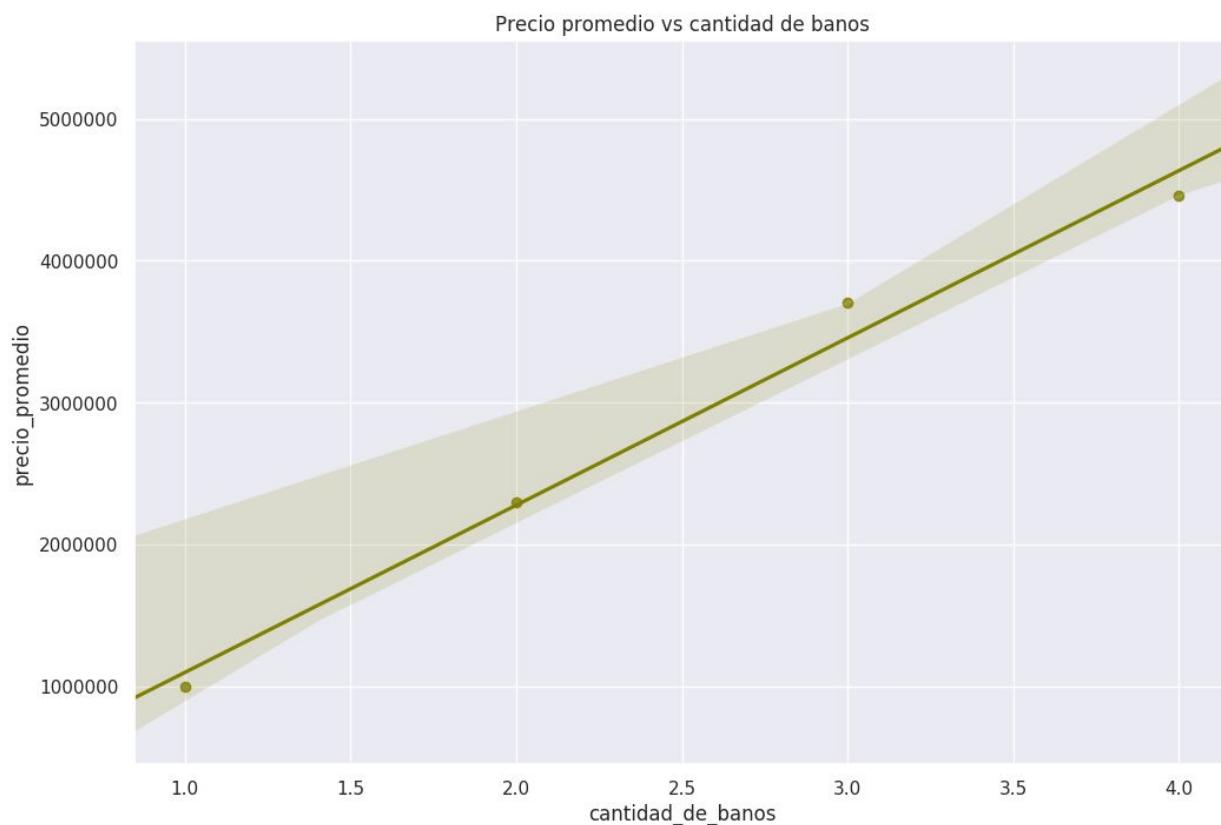
Los monoambientes, por su parte, son más valorados en promedio que los de dos habitaciones, lo cual se condice con el boom que tuvieron en los últimos tiempos.

¿Cómo es la distribución según la cantidad de baños?



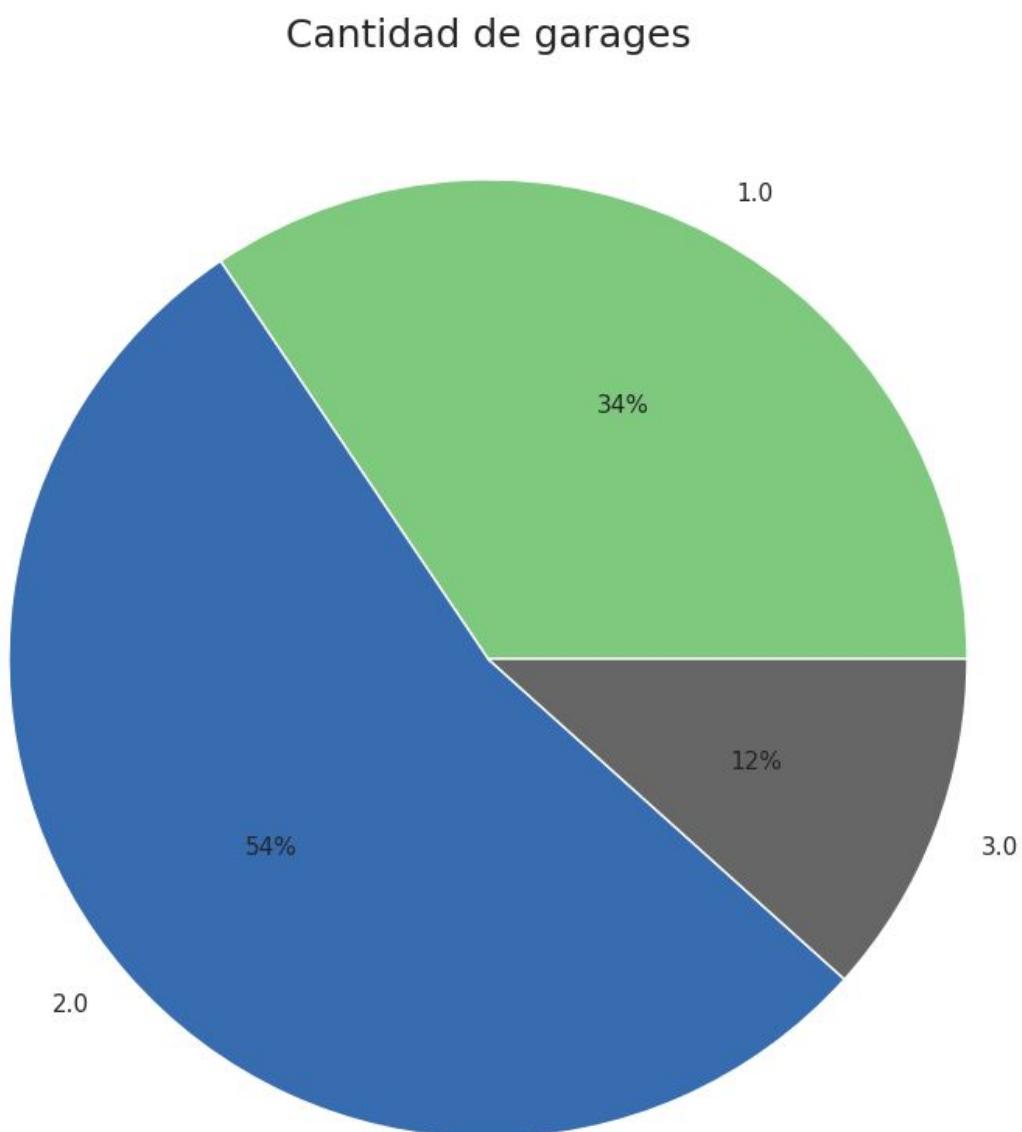
Se puede observar que un poco menos de la mitad de las publicaciones son de inmuebles con 2 baños.

¿Cómo es la variación de precio promedio según la cantidad de baños?



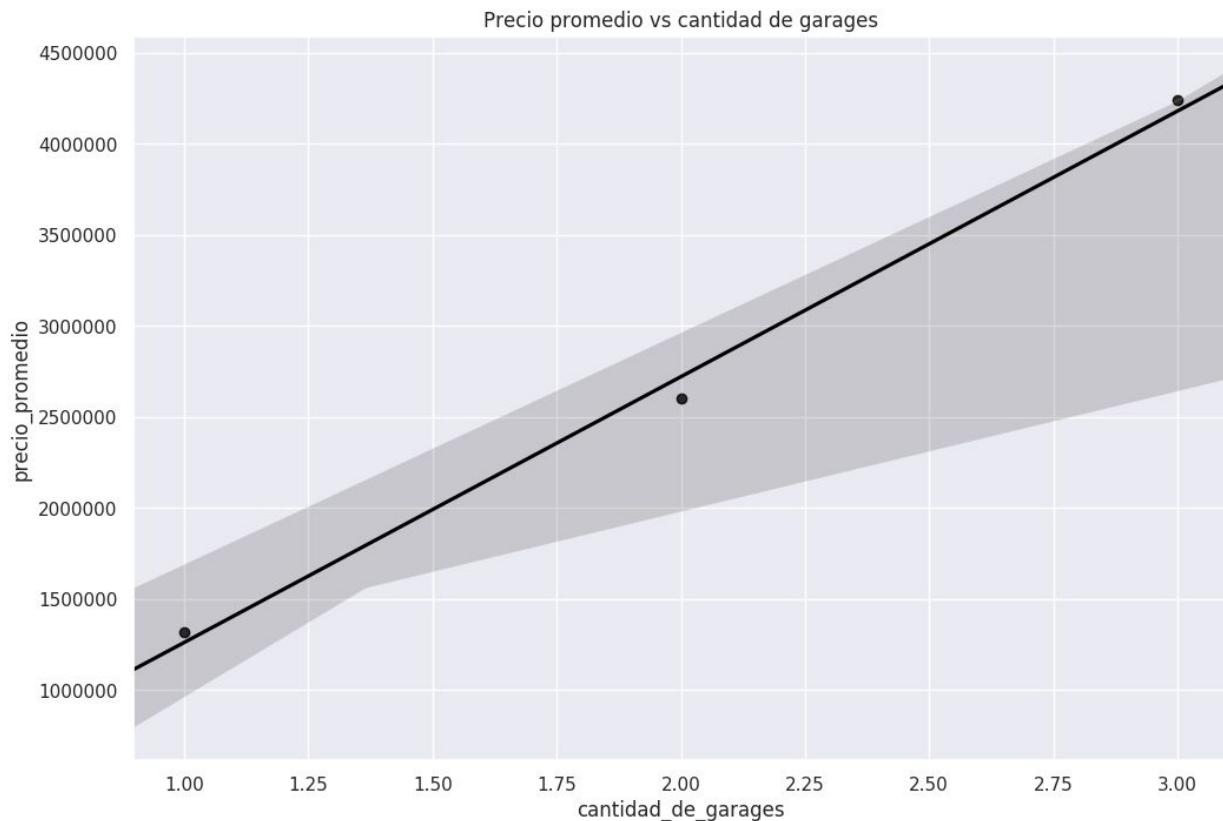
Como se ve reflejado en el gráfico, es claro el aumento del precio promedio de los inmuebles a medida que aumenta su cantidad de baños.

¿Cómo es la distribución según la cantidad de garages?

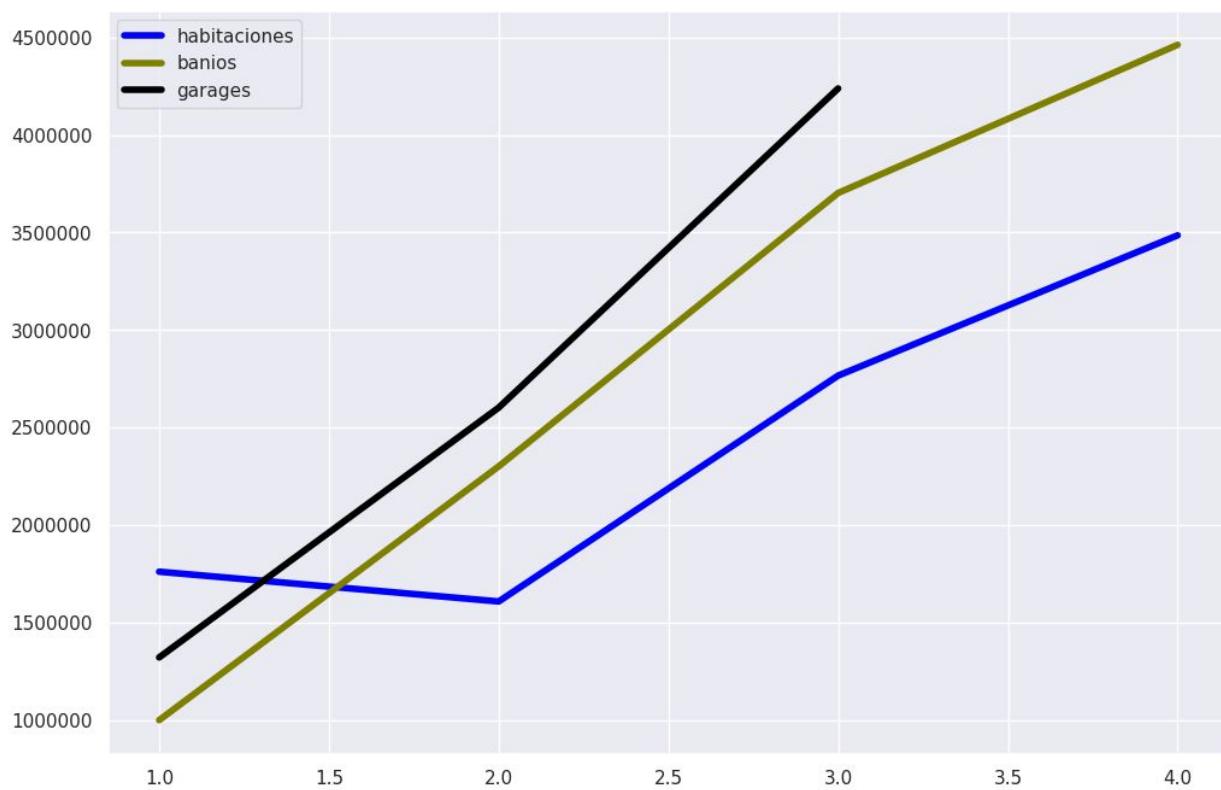


Se puede observar que más de la mitad de las publicaciones son de inmuebles que cuentan con 2 garajes.

¿Cómo es la variación de precio promedio según la cantidad de garages?



Como se ve reflejado en el gráfico, es claro el aumento del precio promedio de los inmuebles a medida que aumenta su cantidad de garages.

Comparación de la variación del precio promedio según cantidad de habitaciones, baños y garages

A Partir del gráfico, es fácil concluir que la función de precio tiene una pendiente positiva respecto a la cantidad de habitaciones, baños y garages. Pero una parte se resulta interesante que entre una habitación y 2 habitaciones hay una recta con pendientes negativa, es decir que los promedios de las propiedades con una habitacion resultan valer más que la de 2 habitaciones. También, es natural que a medida que aumenta las cantidad de habitaciones, baños y garages, aumentan los metros totales de la propiedad.

Este resultado es totalmente esperable, y se establece una relación muy fuerte entre el precio y la cantidad de habitaciones, baños y garages.

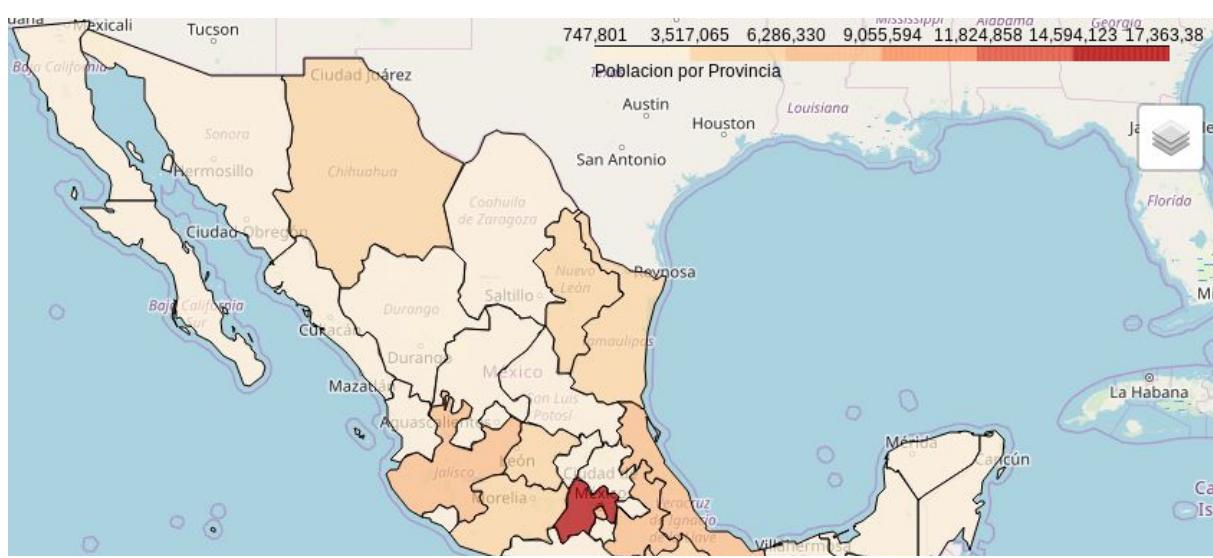
Análisis por provincia

¿Cómo se relaciona la cantidad de publicaciones por provincia y la cantidad de habitantes que tiene?

Cantidad de publicaciones por Provincia



Población por Provincia



La gran mayoría de las publicaciones está concentrada en el Distrito Federal y el Estado de México, seguida por los estados de Jalisco, Querétaro, Nuevo León y Puebla.

Además, un detalle que escapa al gráfico es que hay cinco provincias que cuentan con menos de 1000 publicaciones, y ellas son: Zacatecas, Campeche, Oaxaca, Tlaxcala y Tabasco.

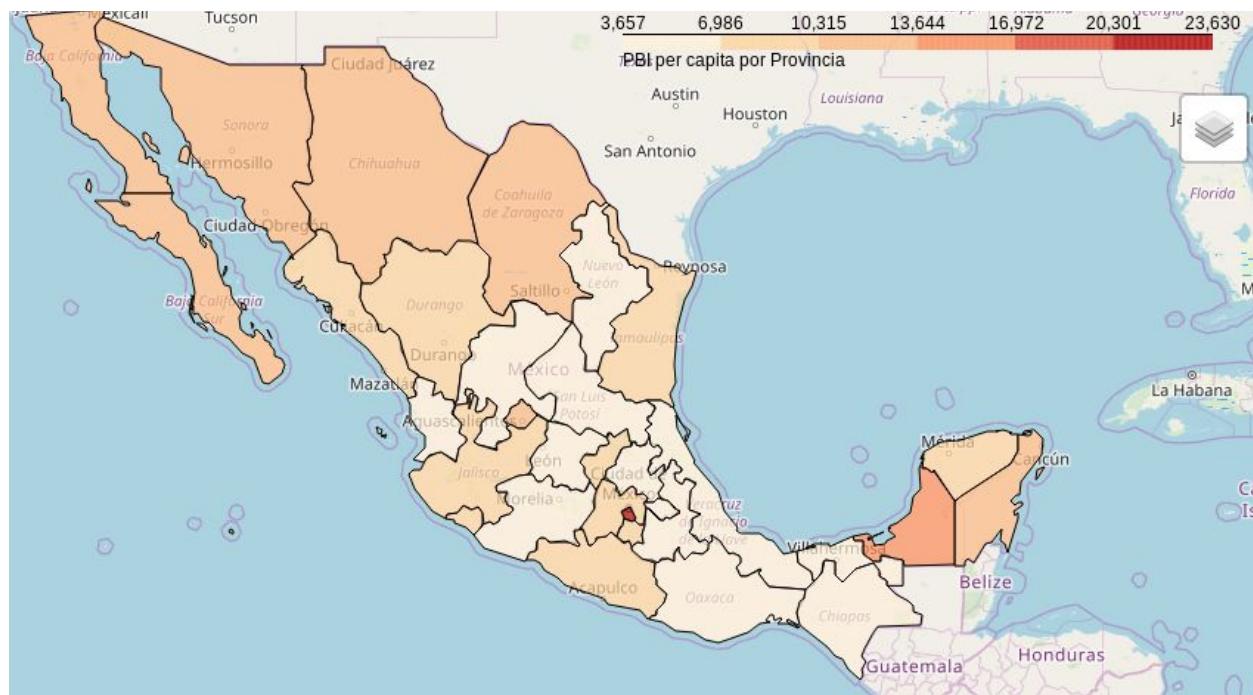
Comparando ambos gráficos, podemos concluir que hay mayor cantidad de publicaciones en los lugares donde hay más población, particularmente donde esta más se concentra, es decir, el Distrito Federal y sus alrededores.

Observación: De aquí en adelante -en esta sección- el análisis será efectuado en Casas, Apartamentos, y Casa en condominios.

¿Cómo se relaciona el precio promedio por m² con el PBI per cápita en cada provincia?

Precio promedio por m² por provincia



PBI per cápita por provincia

A raíz de ambos gráficos, vemos que no hay una relación fuertemente marcada entre el PBI per cápita por provincia y el costo del precio por m² en Casas, Apartamentos, y Casa en condominios.

Si bien el Distrito Federal y Quintana Roo presentan precios por m² acordes a su PBI per cápita, esto no se ve reflejado en otros casos, tales como los estados del Norte, cuyo PBI per cápita es significativo.

¿Cuáles son las 5 provincias de México con inmuebles más caros?

¿Cuáles son las 5 provincias de México con inmuebles más baratos?



Análisis por ciudad

En esta parte, se analizarán los promedios del precio/m² en distintas ciudades a partir de la información anterior.

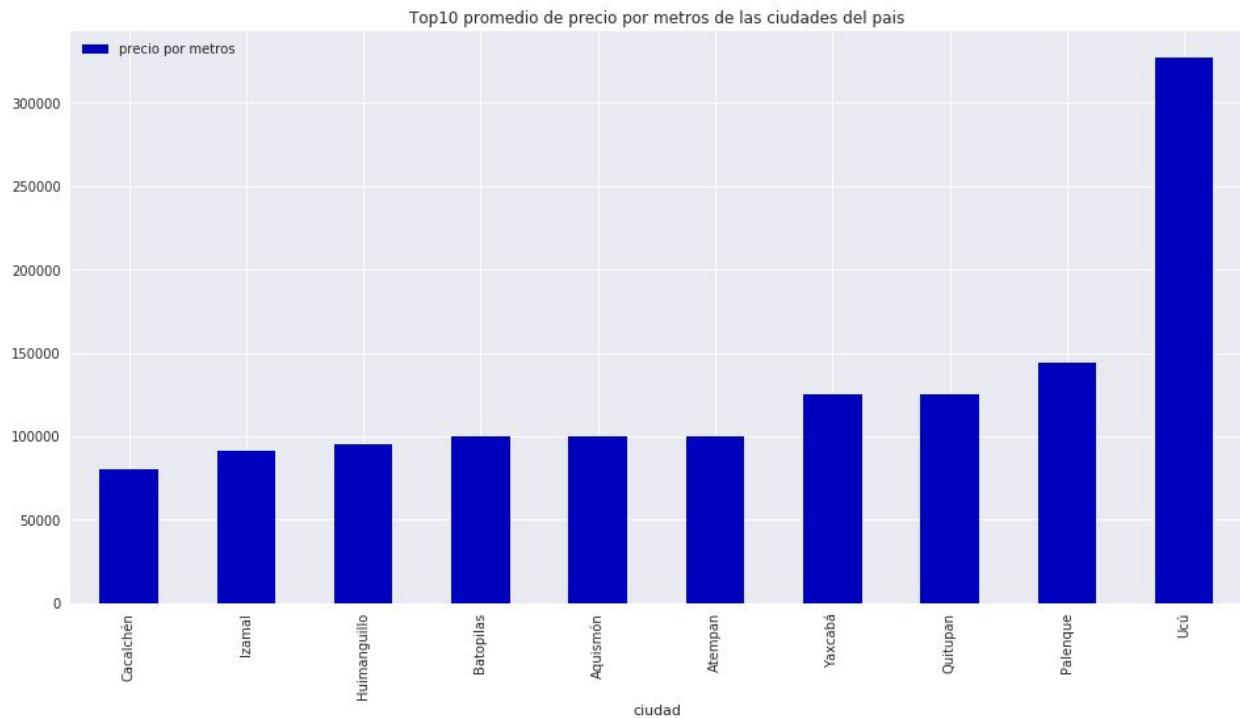
Top 10 ciudades más caras del país

En este caso, se observan los 10 ciudades con mayores promedios de precio/m² de todos los inmuebles.

Al realizar este análisis, se obtiene el siguiente resultado:

Ciudad	Precio/m ²	Ciudad	Precio/m ²
Ucú	326,923.08 \$MX/m ²	Batopilas	100,000.00 \$MX/m ²
Palenque	143,816.12 \$MX/m ²	Huimanguillo	100,000.00 \$MX/m ²
Yaxcabá	125,000.00 \$MX/m ²	Izamal	95,370.30 \$MX/m ²
Quitupan	125,000.00 \$MX/m ²	Aquismón	90,909.09 \$MX/m ²
Atempan	100,028.79 \$MX/m ²	Cacalchén	80,000.00 \$MX/m ²

En la tabla se observa que la primera ciudad, Ucú, tiene un precio/m² mucho mayor que los demás, y entre la segunda y la última las diferencias entre precios son menos significativas.



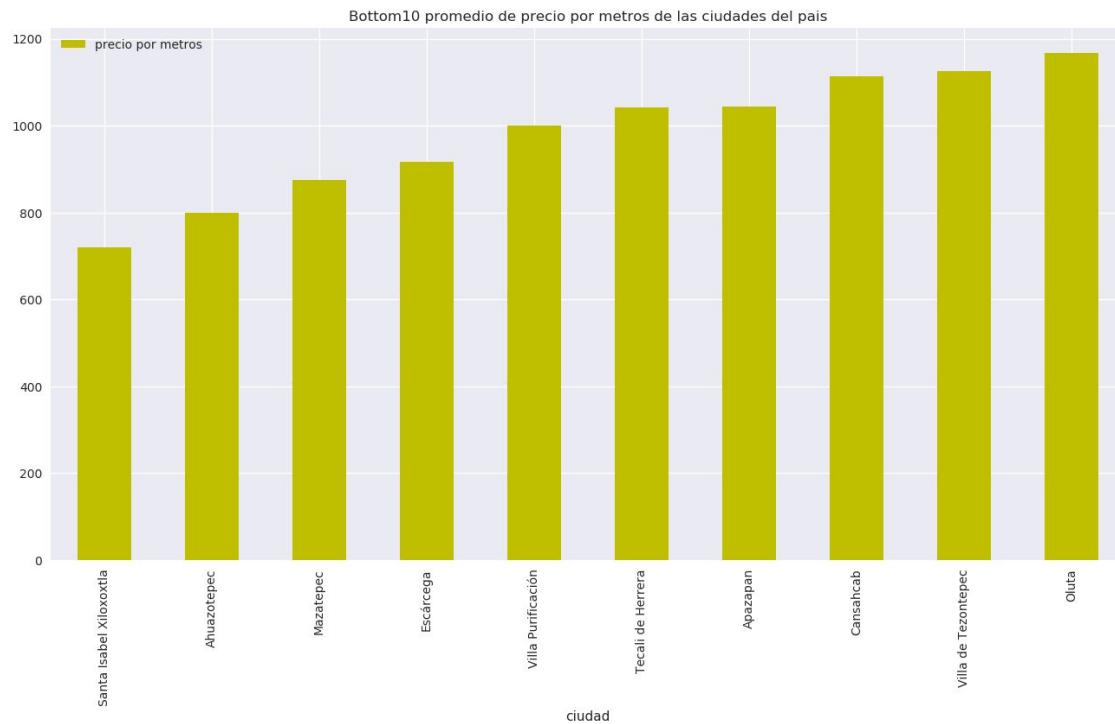
Top 10 ciudades más baratas del país

Para esta parte, se ordenan los datos para analizar cuáles ciudades tienen menor promedio de precio/m² de todas las propiedades.

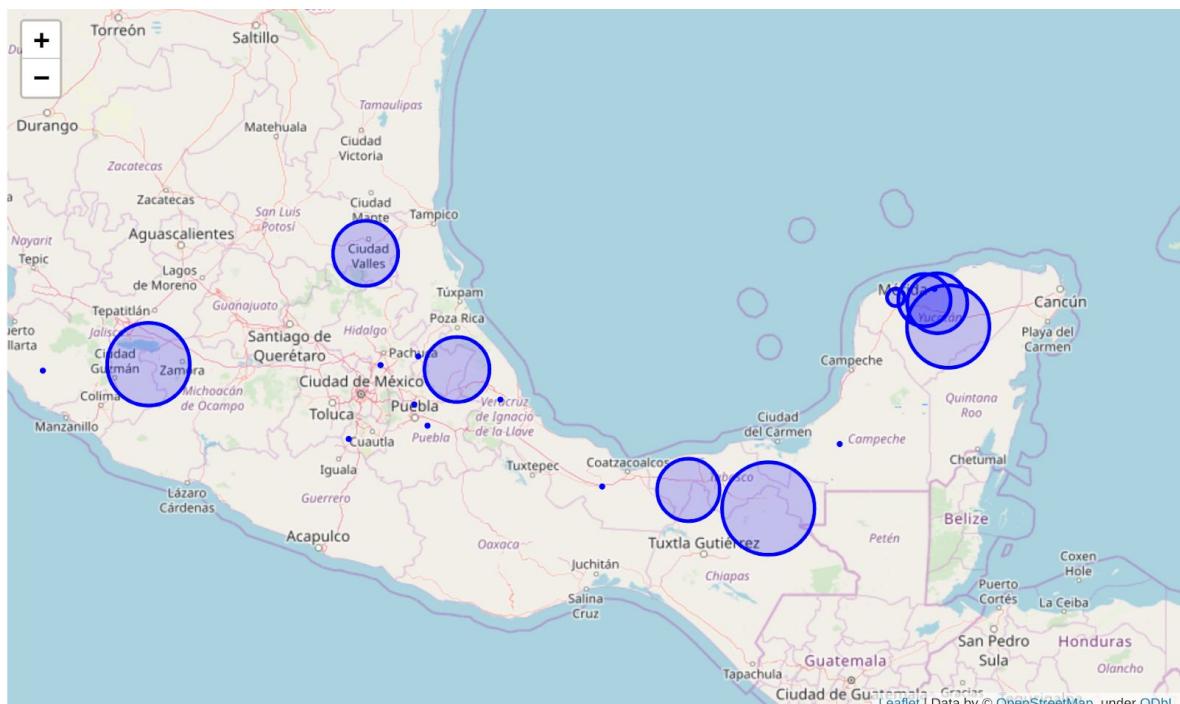
En este caso, el resultado obtenido fue:

Ciudad	Precio/m ²	Ciudad	Precio/m ²
Santa Isabel Xiloxoxtla	720 \$MX/m ²	Tecali de Herrera	1041.67 \$MX/m ²
Ahuazotepec	800.00 \$MX/m ²	Apazapan	1044.39 \$MX/m ²
Mazatepec	875.00 \$MX/m ²	Cansahcab	1113.33 \$MX/m ²
Escárcega	916.47 \$MX/m ²	Villa de Tezontepec	1125.00 \$MX/m ²
Villa Purificación	1000.00 \$MX/m ²	Oluta	1166.67 \$MX/m ²

Con la tabla se puede observar que hay poca diferencias entre los precios.



Top 10 ciudades más caras y Bottom 10 ciudades más baratas

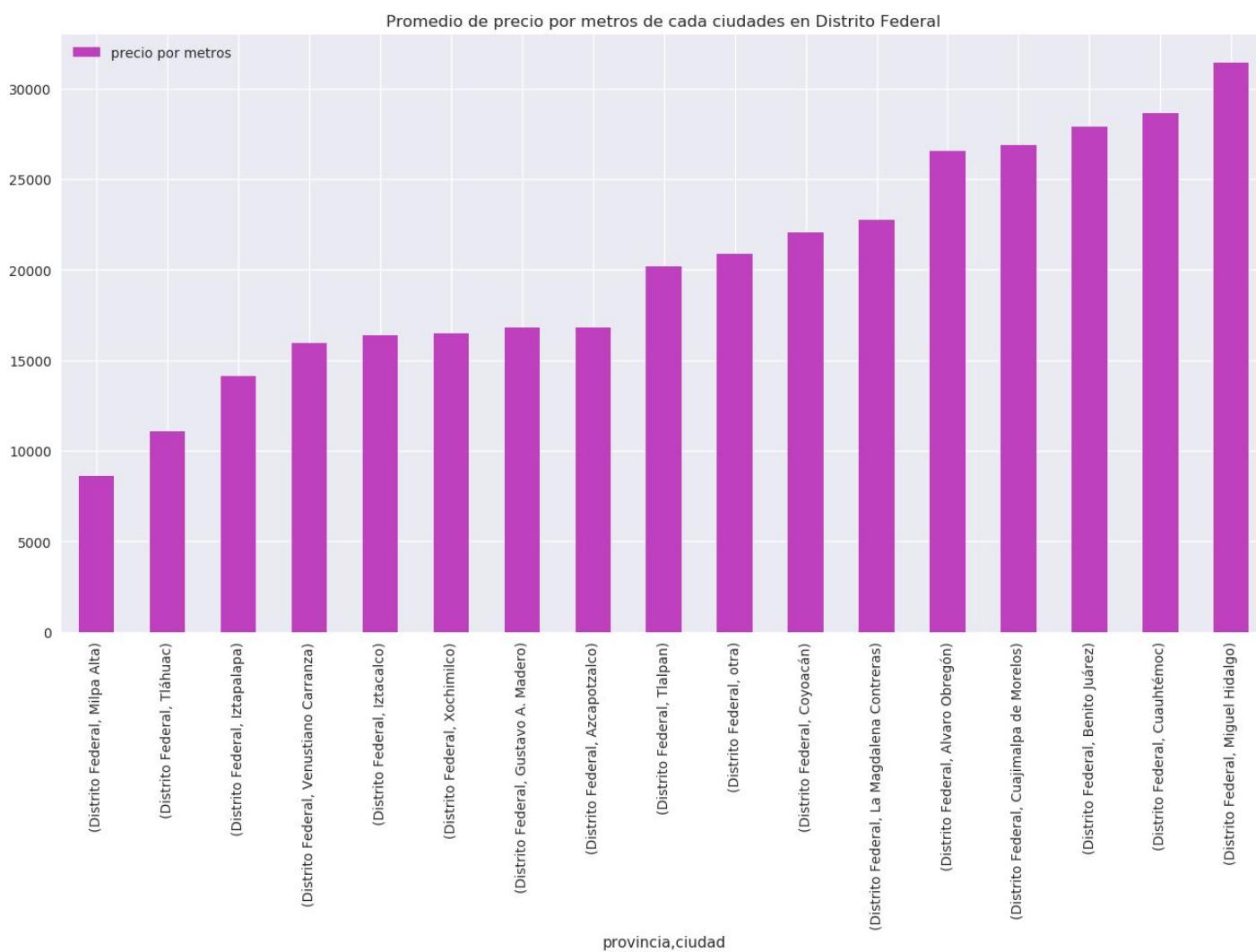


Observación: El tamaño del bubble es función del precio/m², y los mismos se encuentran centrados en las ciudades que representan. Así, las bottom 10 ciudades más baratas quedan reducidas a puntos en el mapa.

Precio/m² de las ciudades de las provincias con mayor publicaciones de ventas

Después de observar la distribución de publicaciones de ventas por provincia, se puede afirmar que sólo hay 5 provincias que tienen cantidades de publicaciones que superen el 5% del total de publicaciones. Por lo tanto, se analizarán los top 10 y bottom 10 de promedio de precio/m² de las ciudades de dichas provincias.

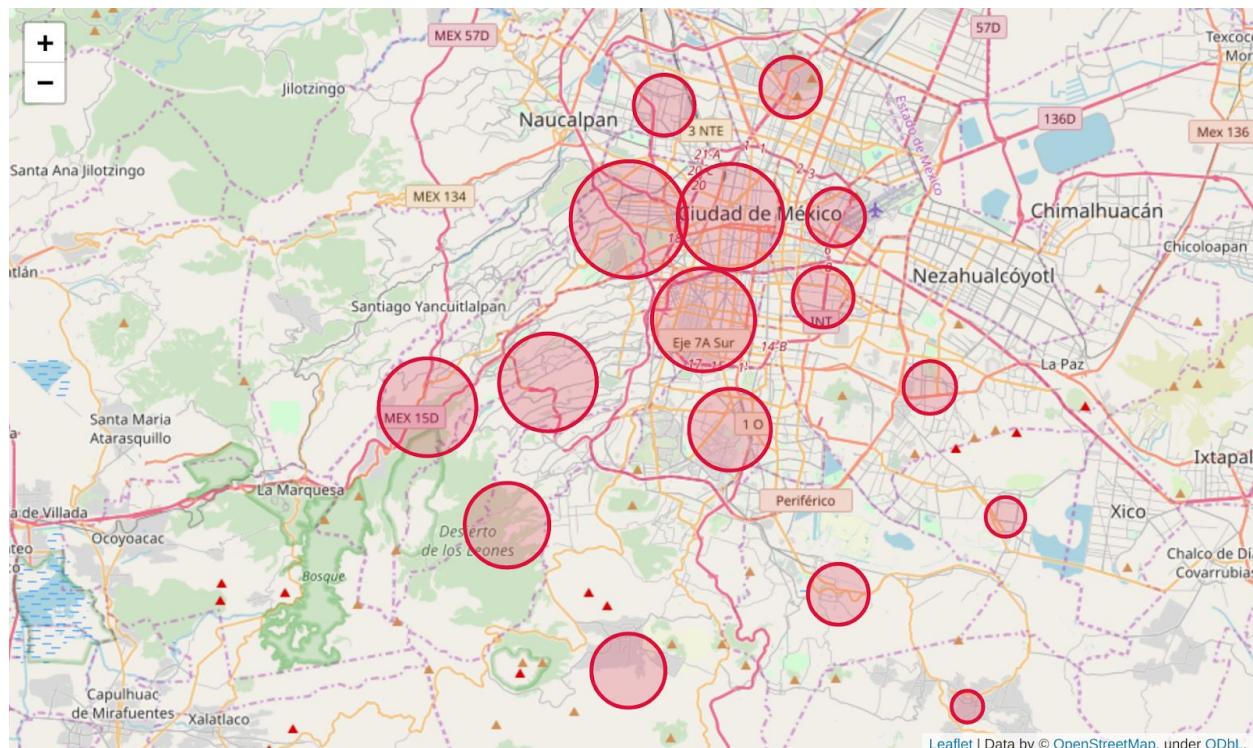
Los precios promedios de propiedades de las ciudades de Distrito Federal



Como únicamente hay información de 17 ciudades del Distrito Federal, se analizaron en el mismo gráfico.

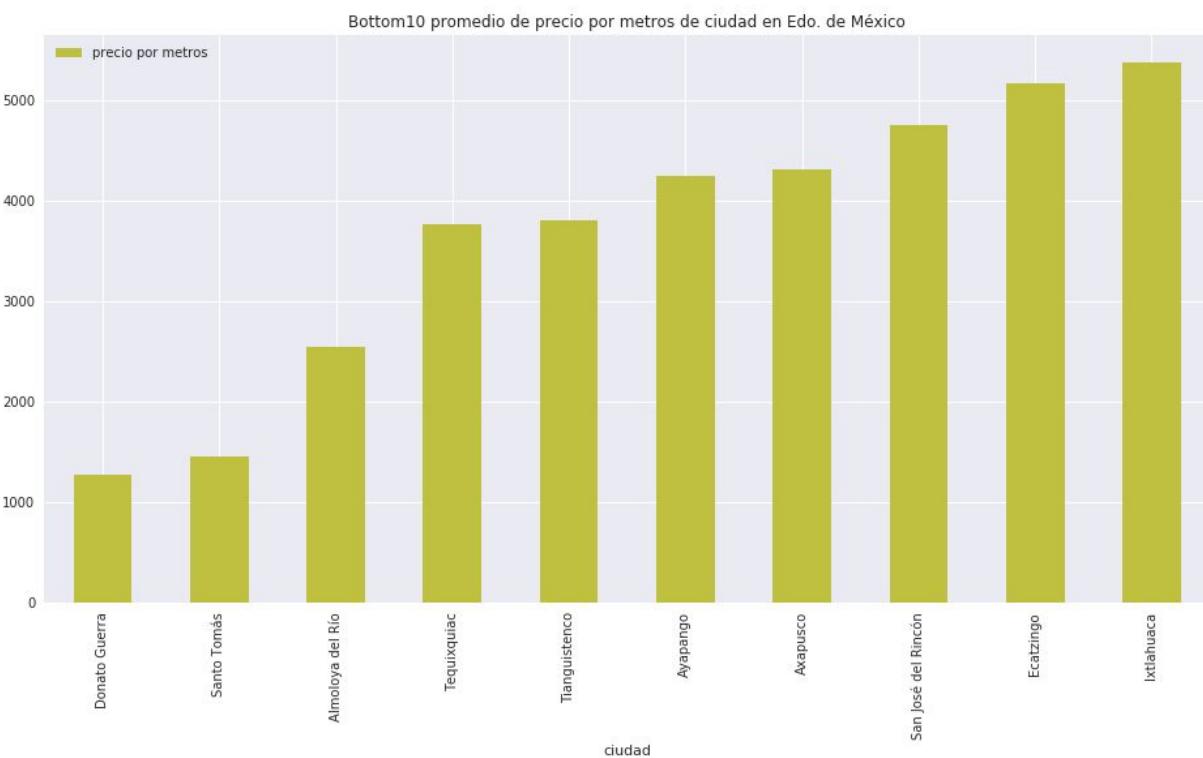
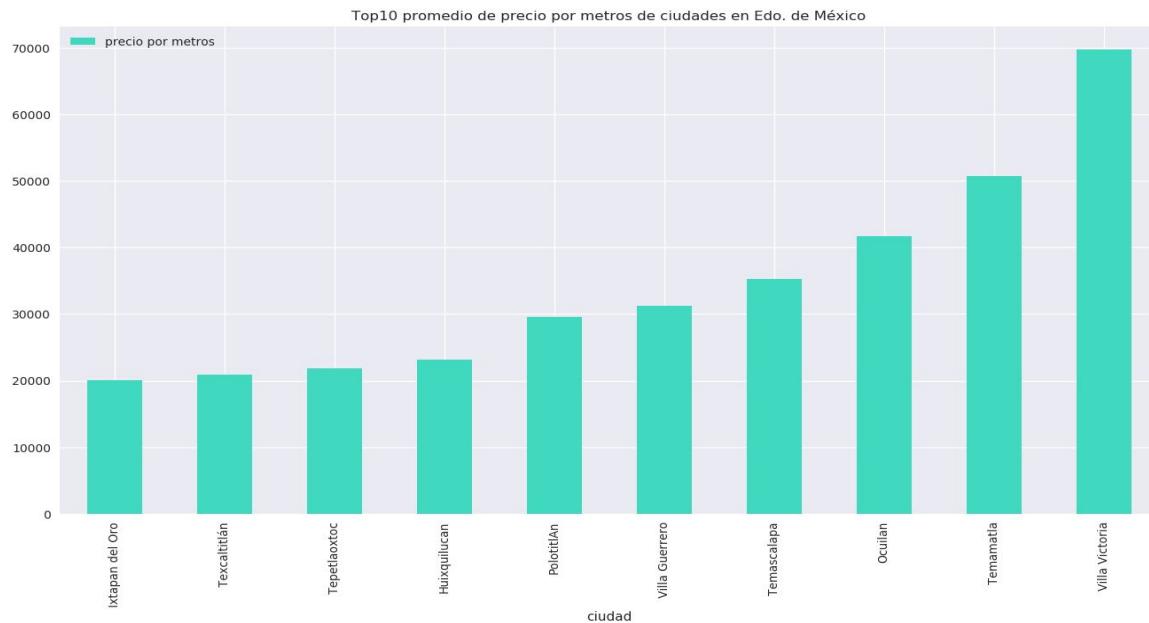
A partir del mismo, se puede observar que hay una gran diferencia de precio entre el primero 'Miguel Hidalgo' y el último 'Milpa Alta'.

Para ver más claro las distribuciones, se representa los datos en el bubble map:



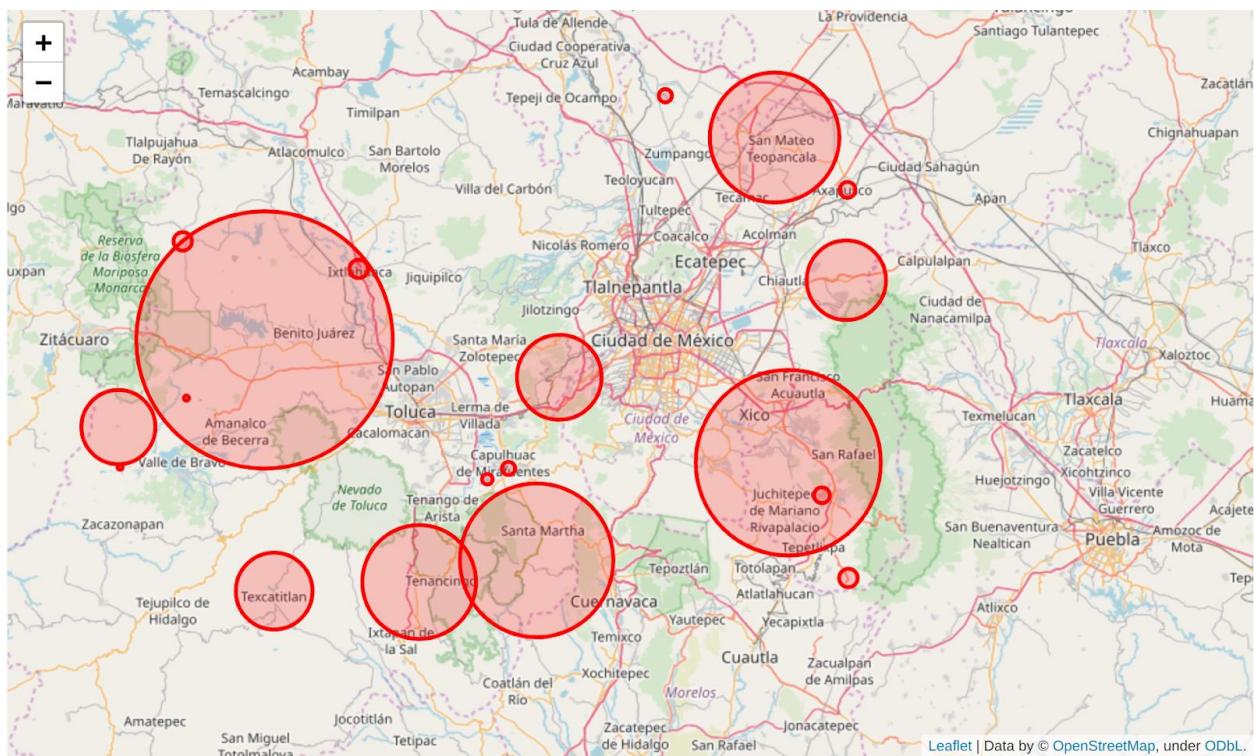
Promedio de precio/m² de propiedades de las ciudades del Edo. de México

En Edo. de México hay más de 20 ciudades, por lo tanto se realiza un gráfico de top 10 y bottom 10, en función de su precio promedio por metro cuadrado:



Se puede observar que en top 10 y bottom 10 hay una gran diferencia entre la primera y la última.

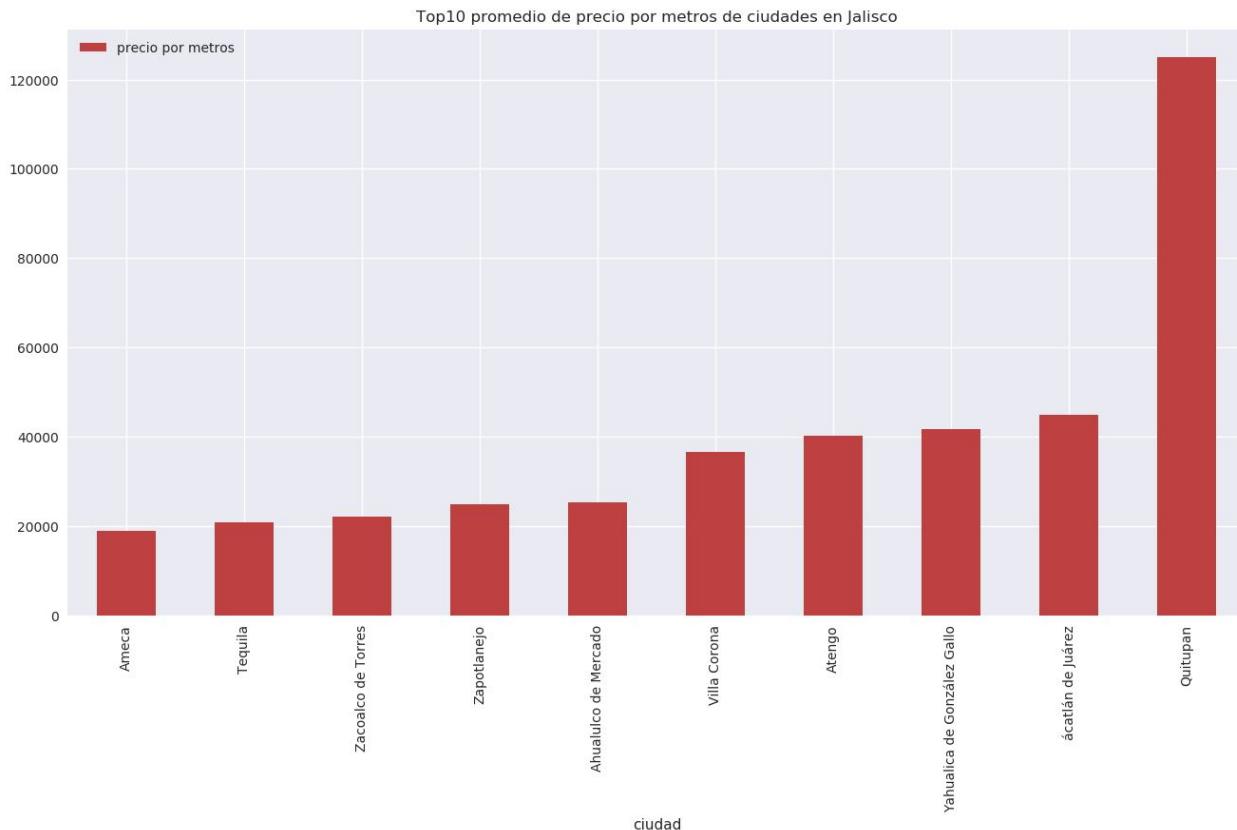
Se utiliza el bubble map a representar esas diferencias geográficamente:

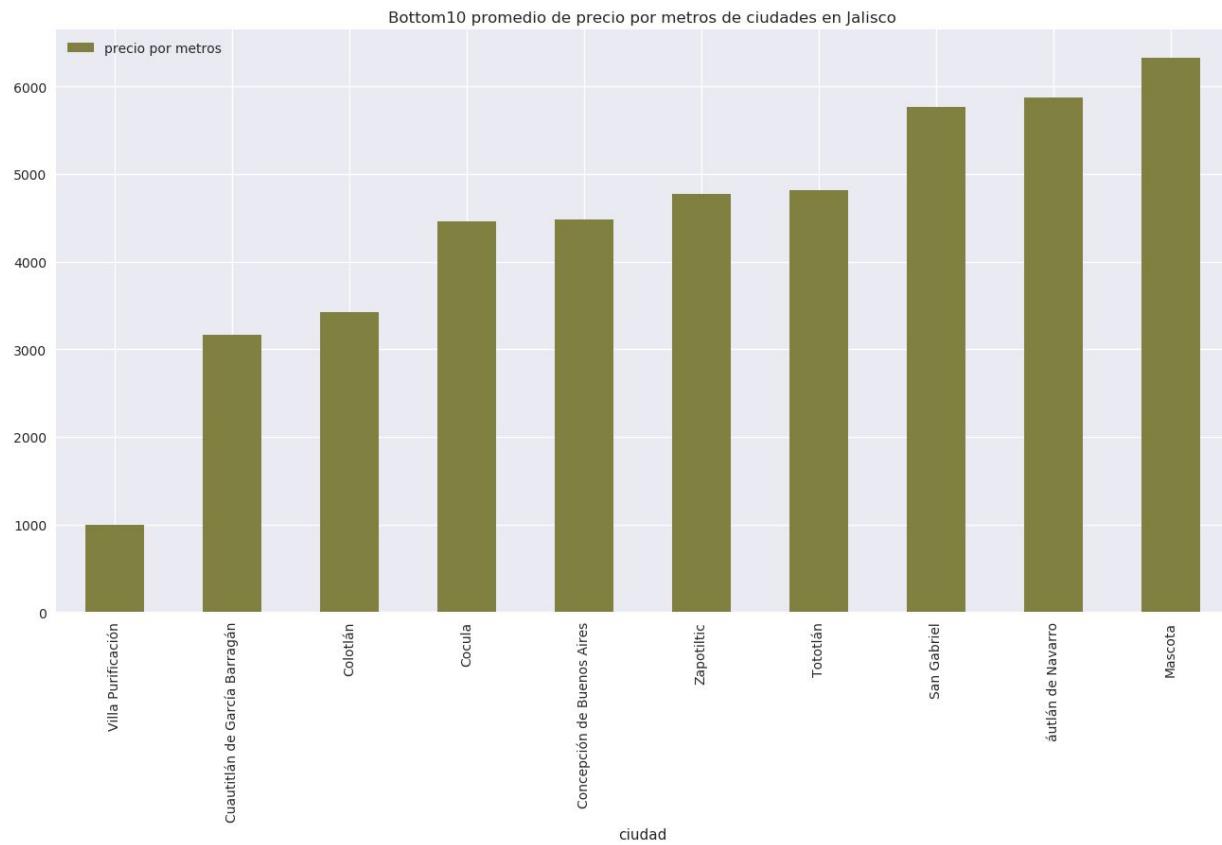


A partir del gráfico, se puede observar que las ciudades más caras no están concentradas en el centro de la ciudad de México, sino en los suburbios de ella.

Promedio de precio/m² de propiedades de las ciudades de Jalisco

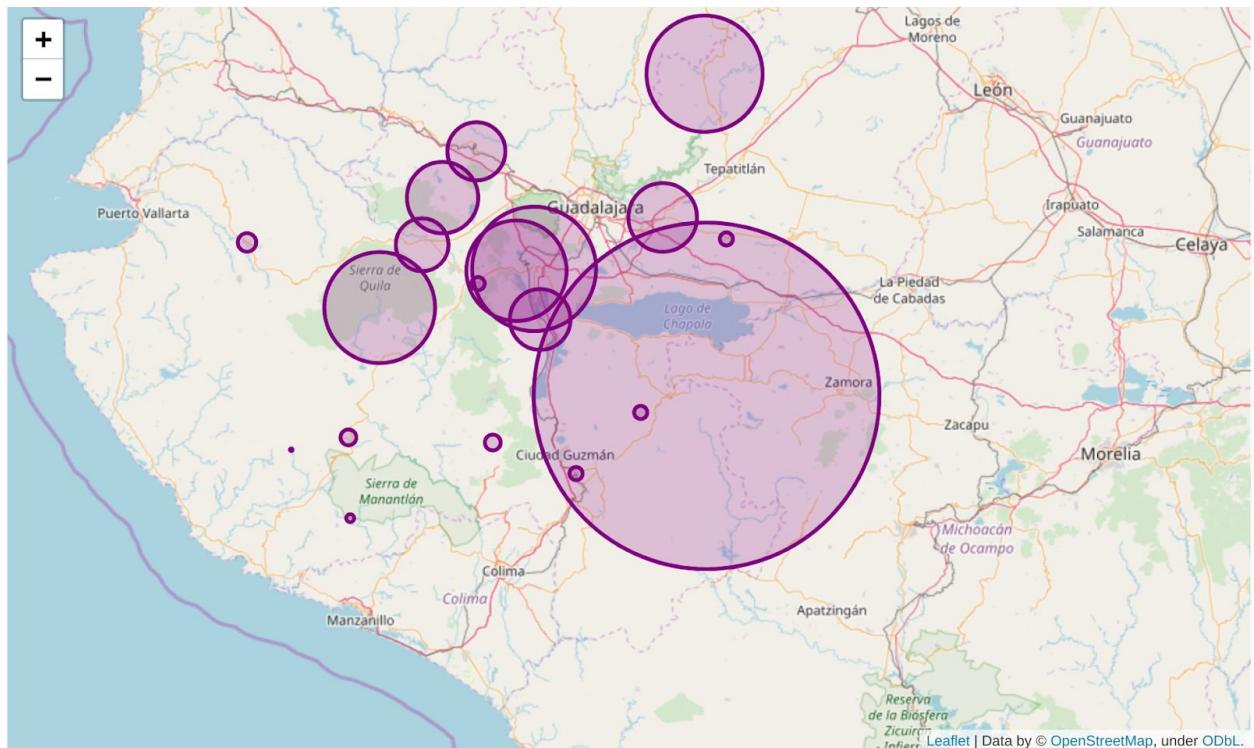
Con más de 20 ciudades en Jalisco, se analiza los top 10 y bottom 10 de los promedios de precios/m²:





Se puede observar que en el top 10 hay una gran diferencia entre el primero y el resto y en bottom 10 ocurre lo mismo pero con el último, con lo cual se ve una importante diferencia entre la ciudad más cara -Quitupan- y el resto.

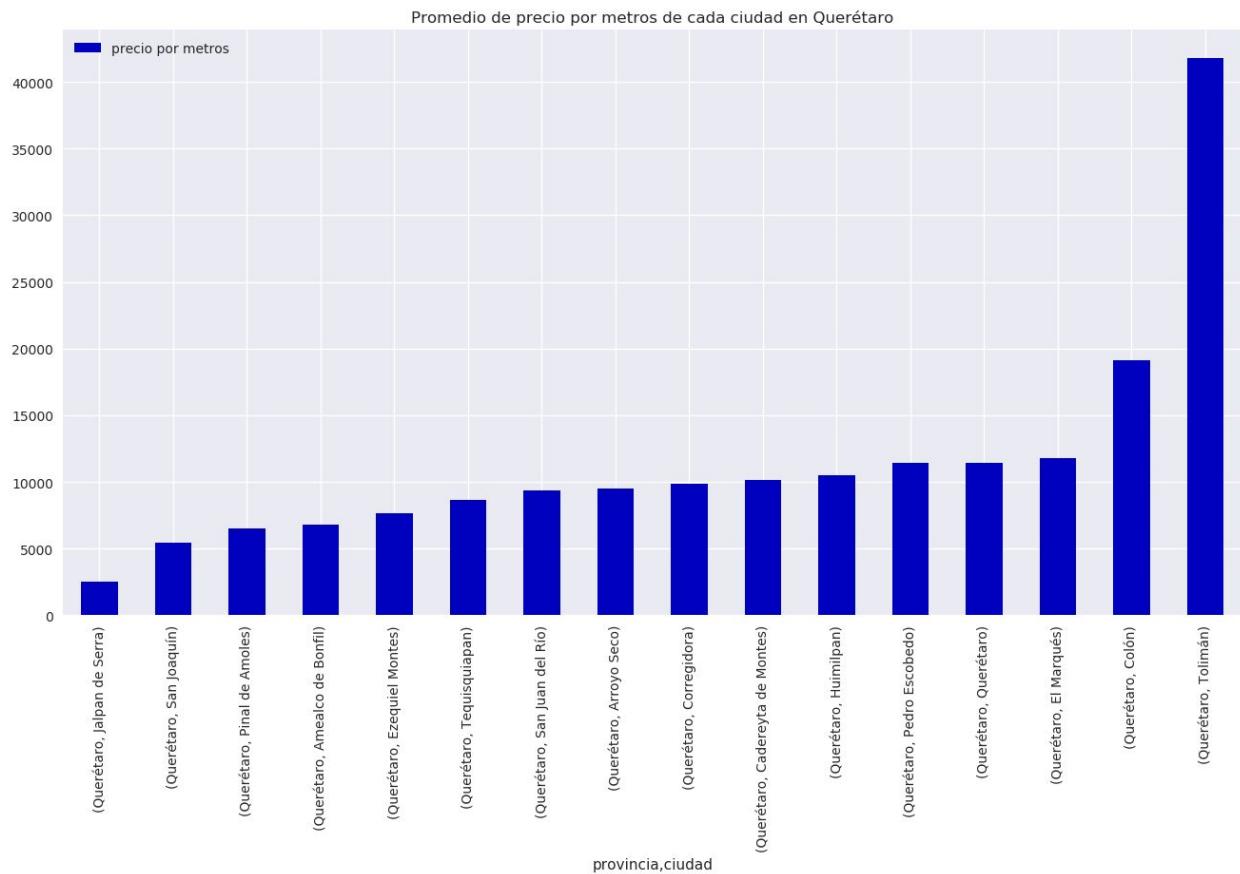
Se presentan los datos en el mapa para ver la distribución de los precios:



En el mapa se puede ver los precios más caros están bien concentrados, salvo la ciudad más cara -Quitupan-. Además, la mayor parte de las ciudades más baratas están lejos del centro de la provincia lo cual resulta razonable.

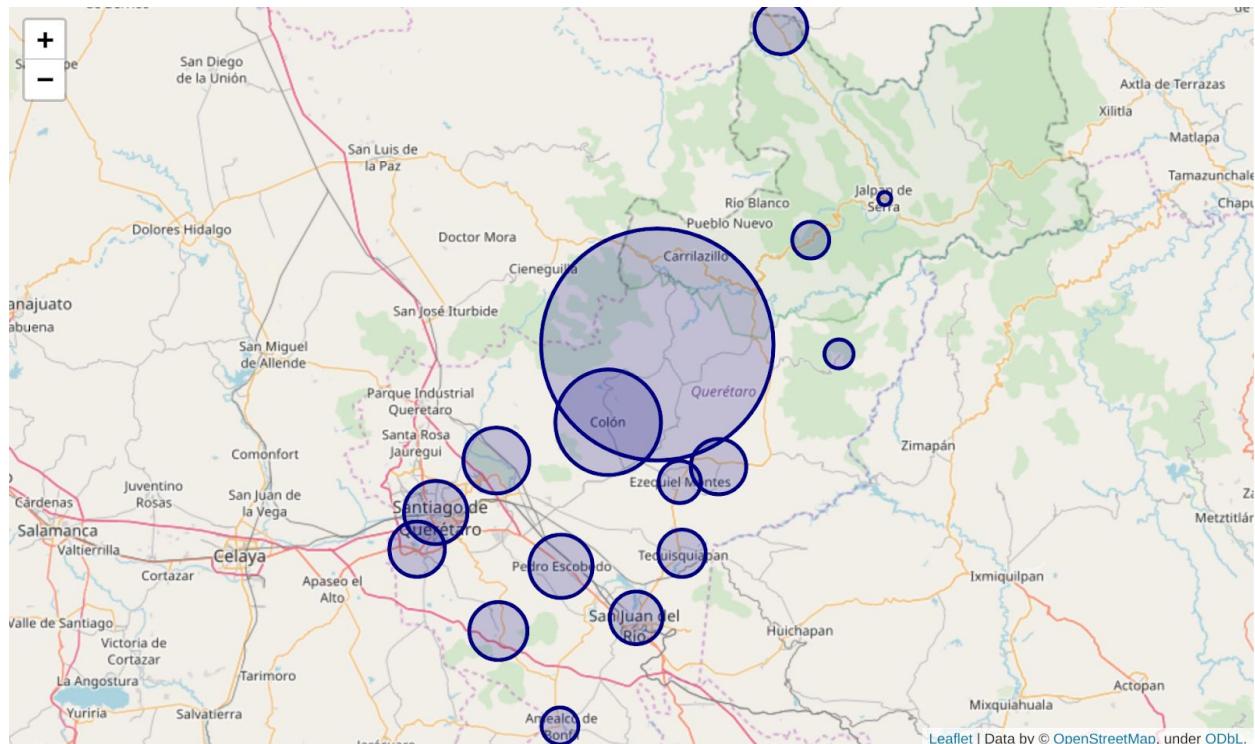
Promedio de precio/m² de propiedades de las ciudades de Querétaro

En la provincia Querétaro hay menos de 20 ciudades con los datos de precios promedios, por lo tanto, se representan las susodichas en un mismo gráfico:



Con el resultado de gráfico, se puede ver que el precio de la ciudad más cara es muy grande con respecto al resto.

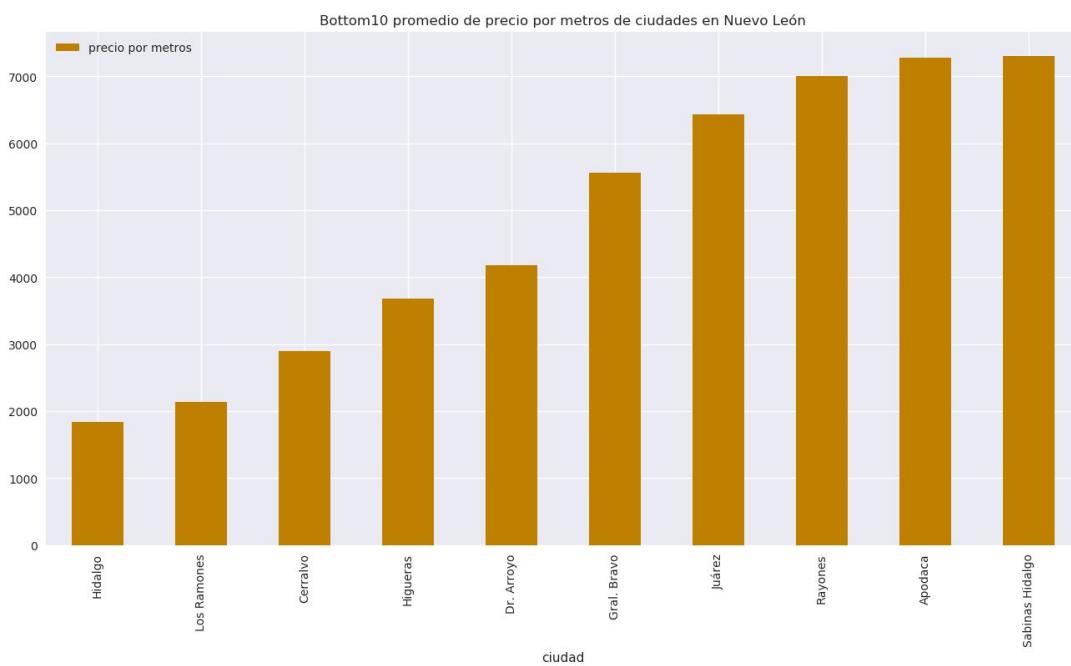
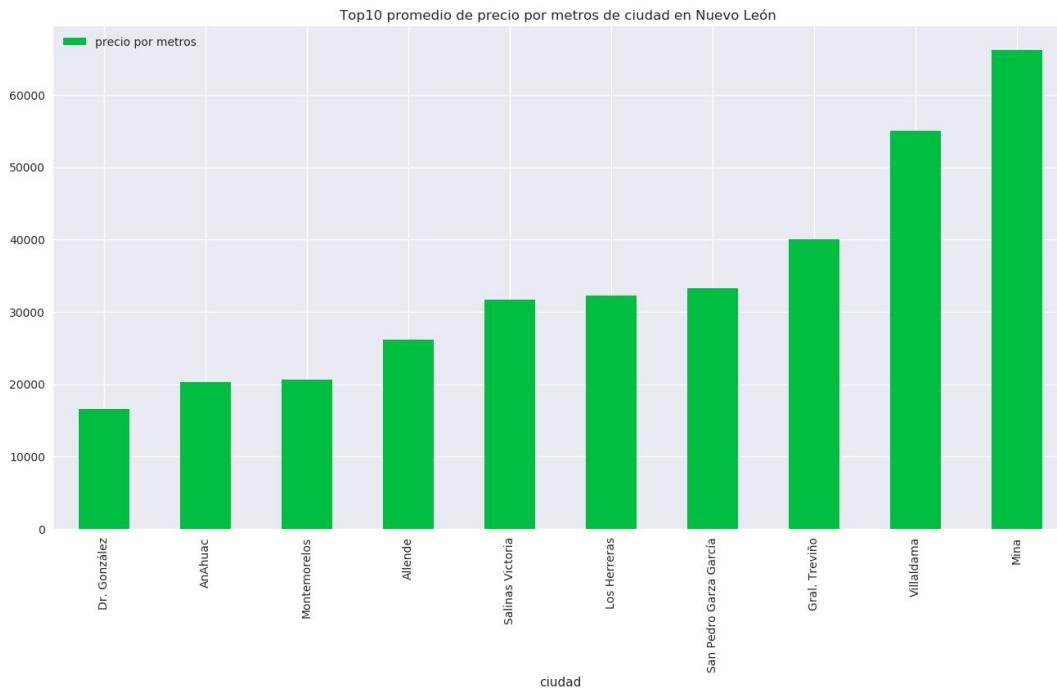
En el mapa debe notarse dichas diferencias:



En la provincia de Querétaro no se ve una relación muy clara entre las ubicaciones y los precios, ya que, a excepción de la ciudad Tolimán (la más cara), la mayoría de las ciudades no tiene una diferencia muy marcada de precio.

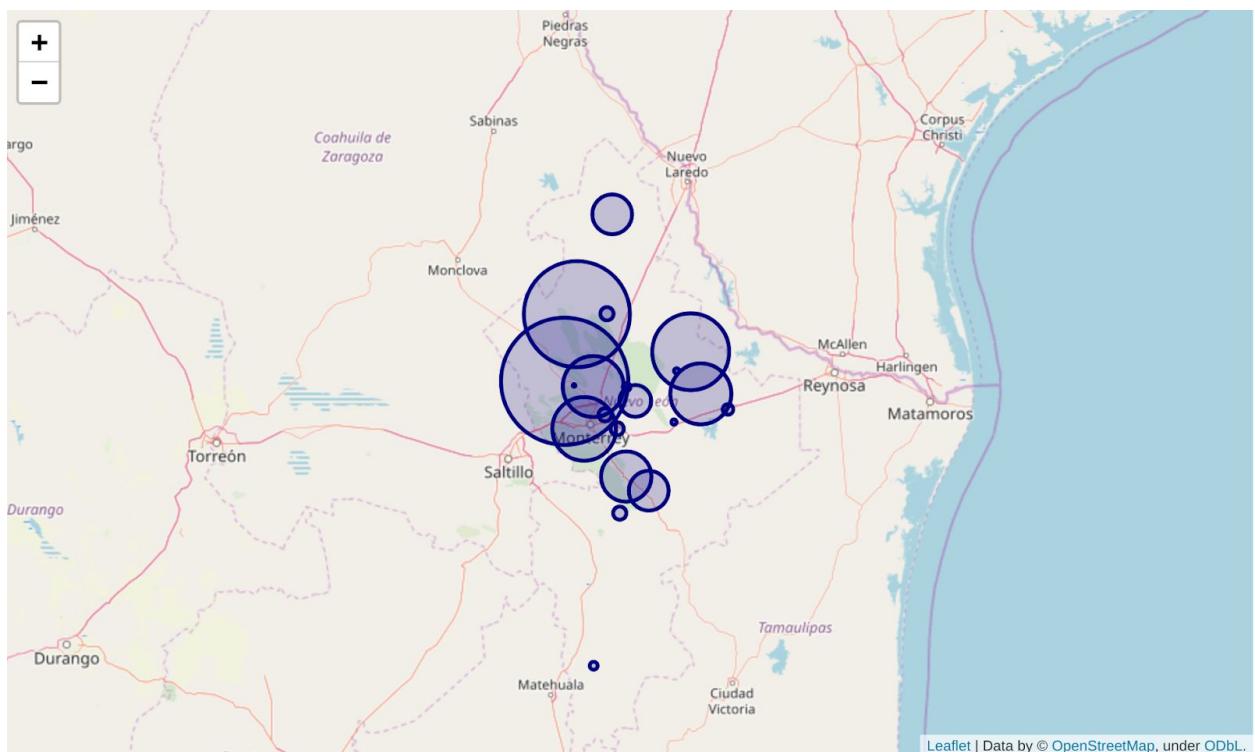
Promedio de precio/m² de propiedades de las ciudades de Nuevo León

Se representa los gráficos obtenidos de top 10 y bottom 10:



Se puede observar que en ambos gráficos, existe una diferencia grande entre el primero y el último. Los primeros tres precios de bottom 10 son parecidos.

Y se puede ver los resultados de datos en el mapa:



Se puede ver que la mayoría de las ciudades más cara y más baratas están bien concentradas en el centro de la provincia Nuevo León.

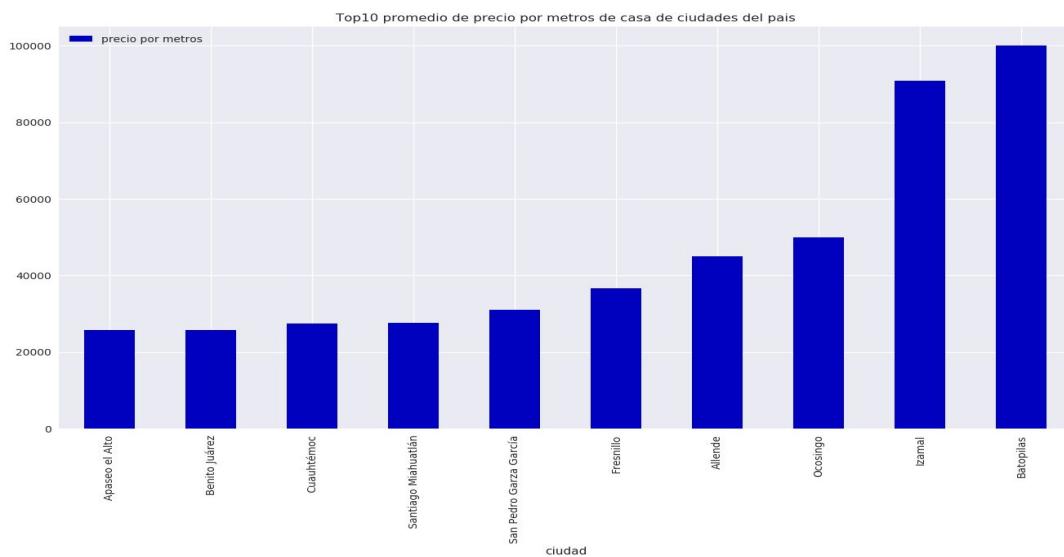
Comparación de precio/m² entre las casas y los apartamentos de las ciudades

Top 10 ciudades con casas más caras del país

Para el caso de casas, se realizó un top 10 obteniendo de los siguientes resultados:

Ciudad	Precio/m ²	Ciudad	Precio/m ²
Batopilas	100000.00 \$MX/m ²	San Pedro Garza García	31091.84 \$MX/m ²
Izamal	90909.09 \$MX/m ²	Santiago Miahuatlán	27588.24 \$MX/m ²
Ocosingo	50000.00 \$MX/m ²	Cuauhtémoc	27411.00 \$MX/m ²
Allende	45008.48 \$MX/m ²	Benito Juárez	25747.42 \$MX/m ²
Fresnillo	36666.67 \$MX/m ²	Apaseo el Alto	25675.00 \$MX/m ²

Se puede observar que en la tabla la diferencia entre las últimas 4 ciudades es pequeña, pero entre la primera y las últimas son enormes.

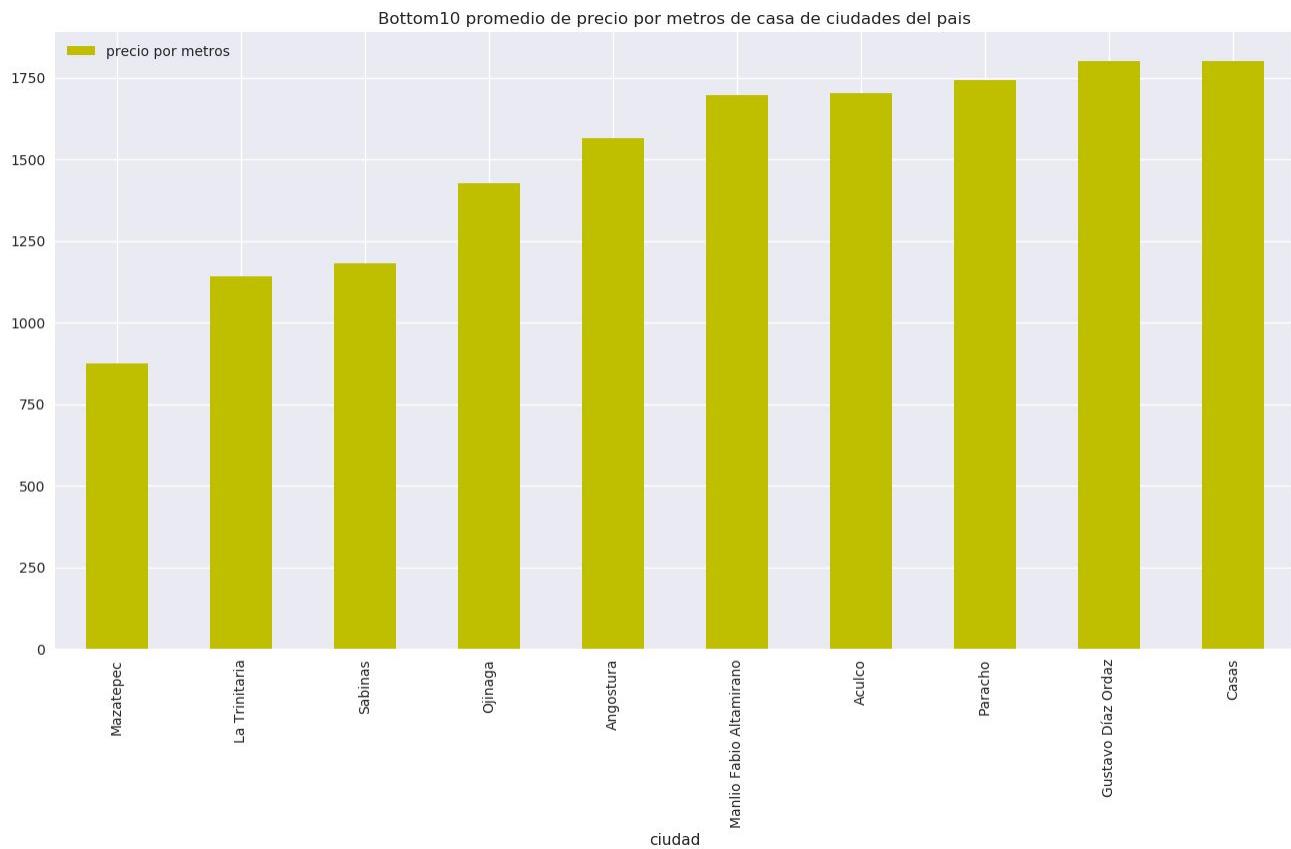


Top 10 ciudades con casa más barata del país

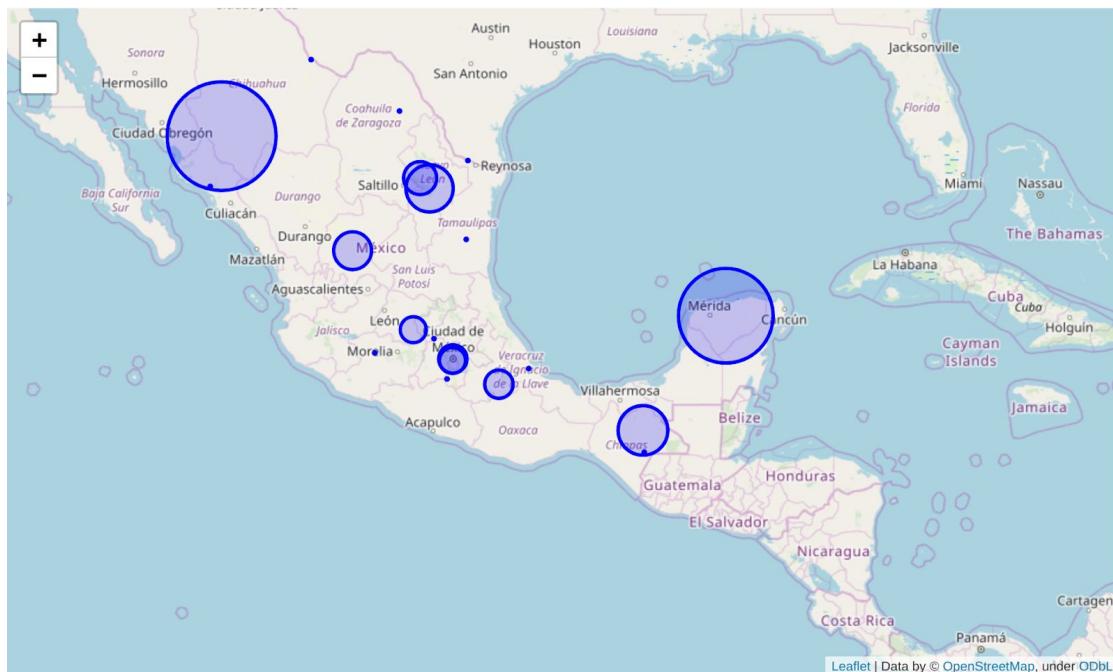
Con la misma forma, se realizó un bottom 10 obteniendo de los siguientes resultados:

Ciudad	Precio/m ²	Ciudad	Precio/m ²
Mazatepec	875.00 \$MX/m ²	Manlio Fabio Altamirano	1697.67 \$MX/m ²
La Trinitaria	1142.86 \$MX/m ²	Aculco	1702.76 \$MX/m ²
Sabinas	1180.17 \$MX/m ²	Paracho	1743.59 \$MX/m ²
Ojinaga	1427.98 \$MX/m ²	Gustavo Díaz Ordaz	1800.00 \$MX/m ²
Angostura	1564.52 \$MX/m ²	Casas	1800.00 \$MX/m ²

Según la tabla, se nota que las diferencias entre las ciudades son mucho menor que las de top 10.



Finalmente, se representa dichos resultados en el mapa:

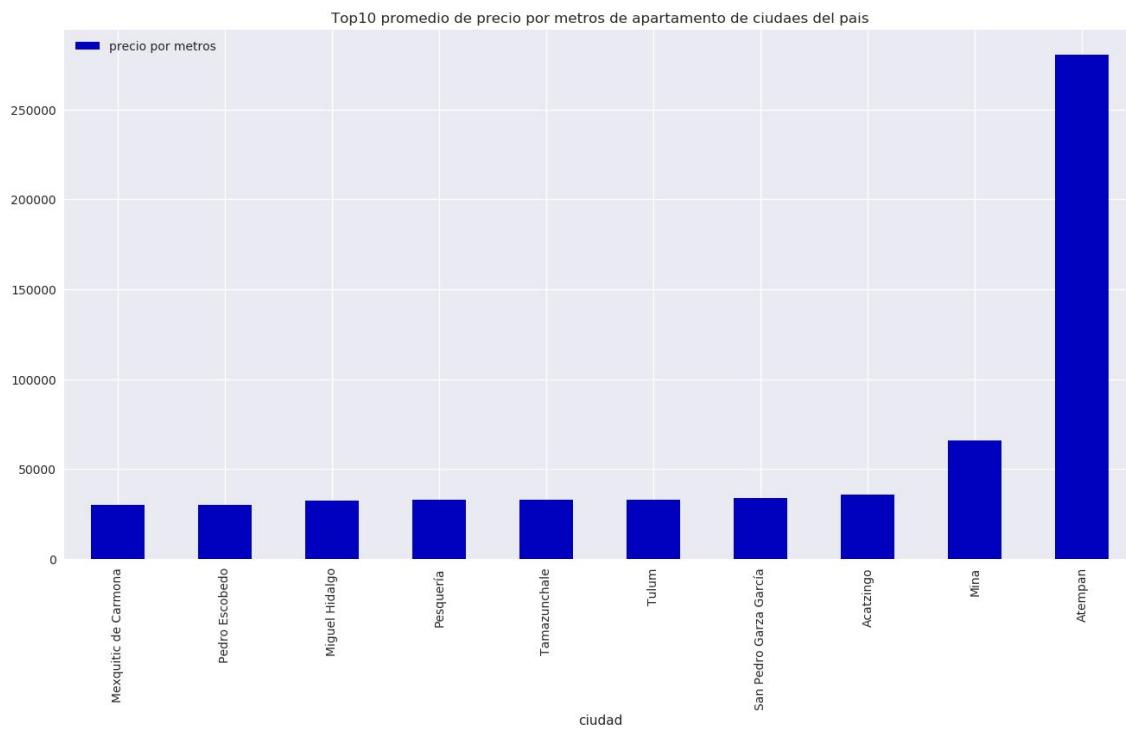


Top 10 ciudades con apartamentos más caro del país

Para el caso de apartamento, se realizó un top 10 obteniendo de los siguientes resultados:

Ciudad	Precio/m ²	Ciudad	Precio/m ²
Atempan	280303.03 \$MX/m ²	Tamazunchale	32894.74 \$MX/m ²
Mina	66153.85 \$MX/m ²	Pesquería	32833.33 \$MX/m ²
Acatzingo	35757.09 \$MX/m ²	Miguel Hidalgo	32643.64 \$MX/m ²
San Pedro Garza García	34055.92 \$MX/m ²	Pedro Escobedo	30195.00 \$MX/m ²
Tulum	33146.86 \$MX/m ²	Mexquitic de Carmona	30188.91 \$MX/m ²

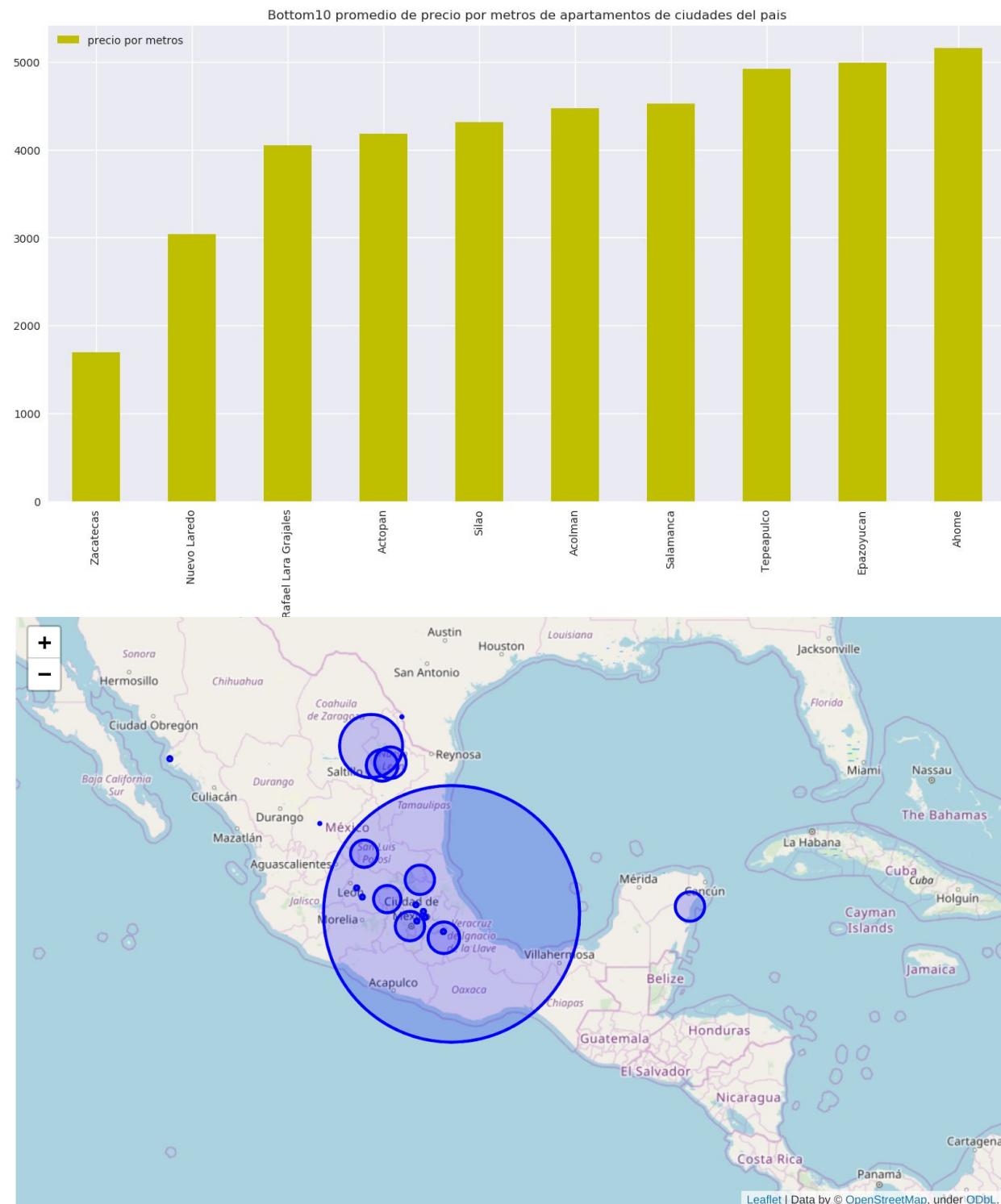
Se puede observar que entre la ciudad con apartamentos más caros del país y el resto hay una gran diferencia de precio.



Top 10 ciudades con apartamentos más baratos del país

De la misma manera, se realizó un bottom 10 obteniendo los siguientes resultados:

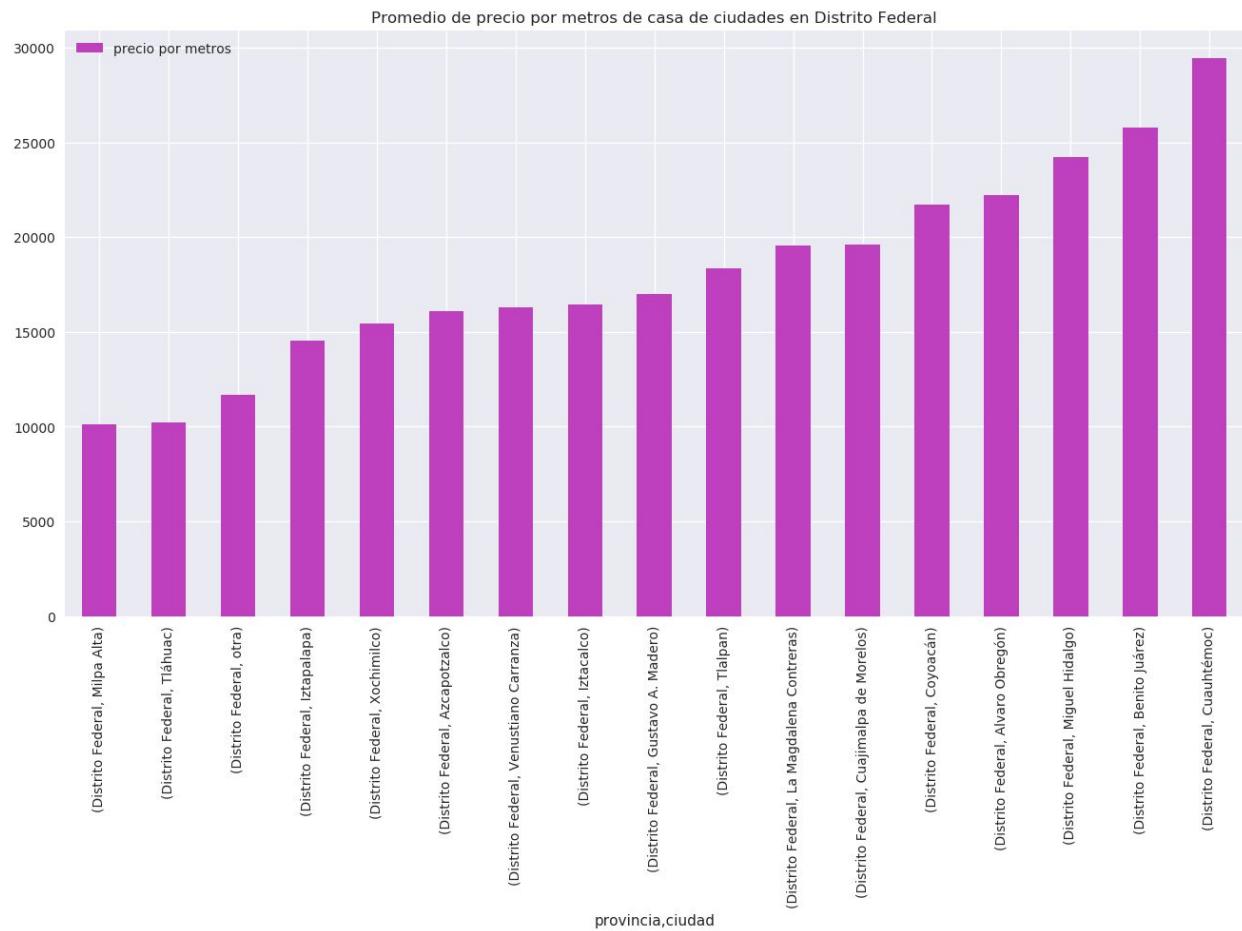
Ciudad	Precio/m ²	Ciudad	Precio/m ²
Zacatecas	1696.11 \$MX/m ²	Acolman	4474.9 \$MX/m ²
Nuevo Laredo	3038.38 \$MX/m ²	Salamanca	4521.28 \$MX/m ²
Rafael Lara Grajales	4050.63 \$MX/m ²	Tepeapulco	4923.08 \$MX/m ²
Actopan	4177.22 \$MX/m ²	Epazoyucan	4992.13 \$MX/m ²
Silao	4314.38 \$MX/m ²	Ahome	5157.34 \$MX/m ²



Al comparar los mapas de distribución de las casas y apartamentos, no presentan una distribución similar.

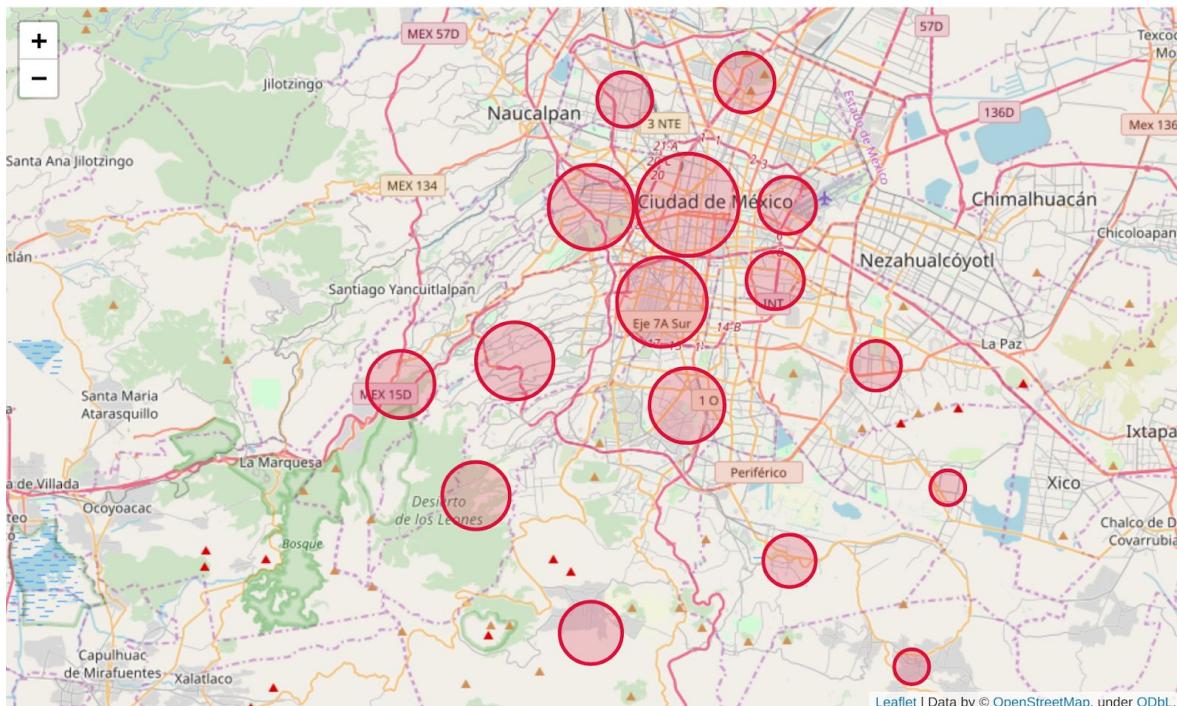
Promedio de precio/m² de casas de las ciudades de Distrito Federal

Se representan los resultados en la misma tabla ya que hay menos de 20 ciudades con datos realizados:

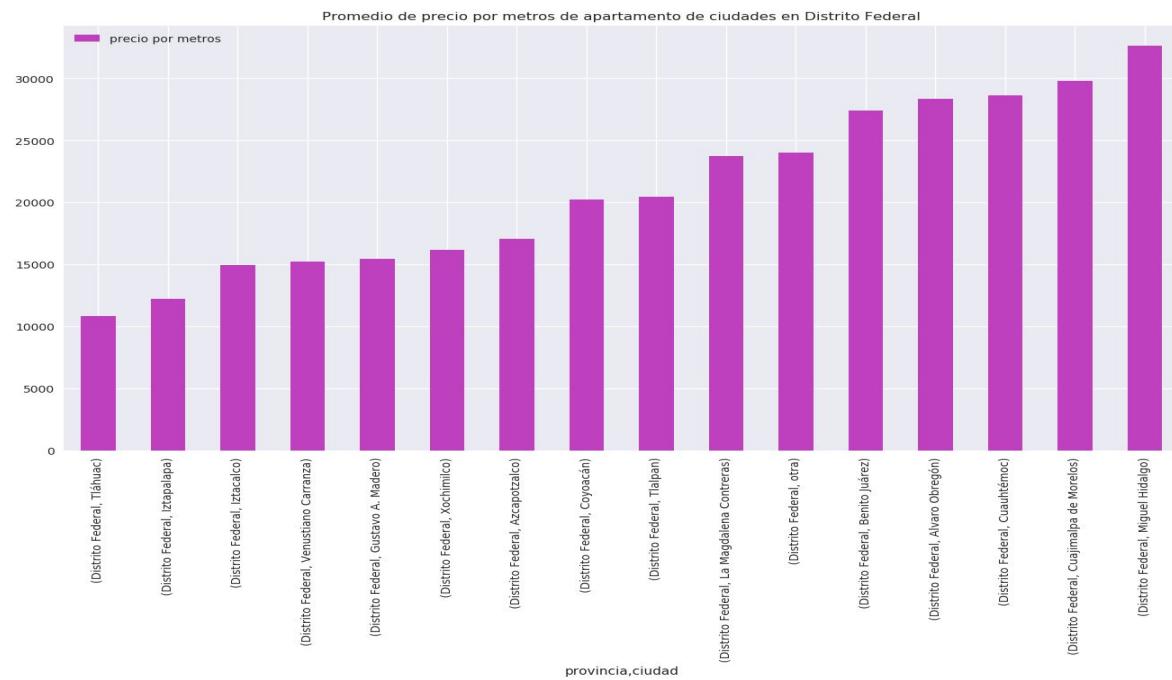


A partir del gráfico, se puede ver que el primero es casi triple que el último. Entre cada una hay una diferencia de precio parecida.

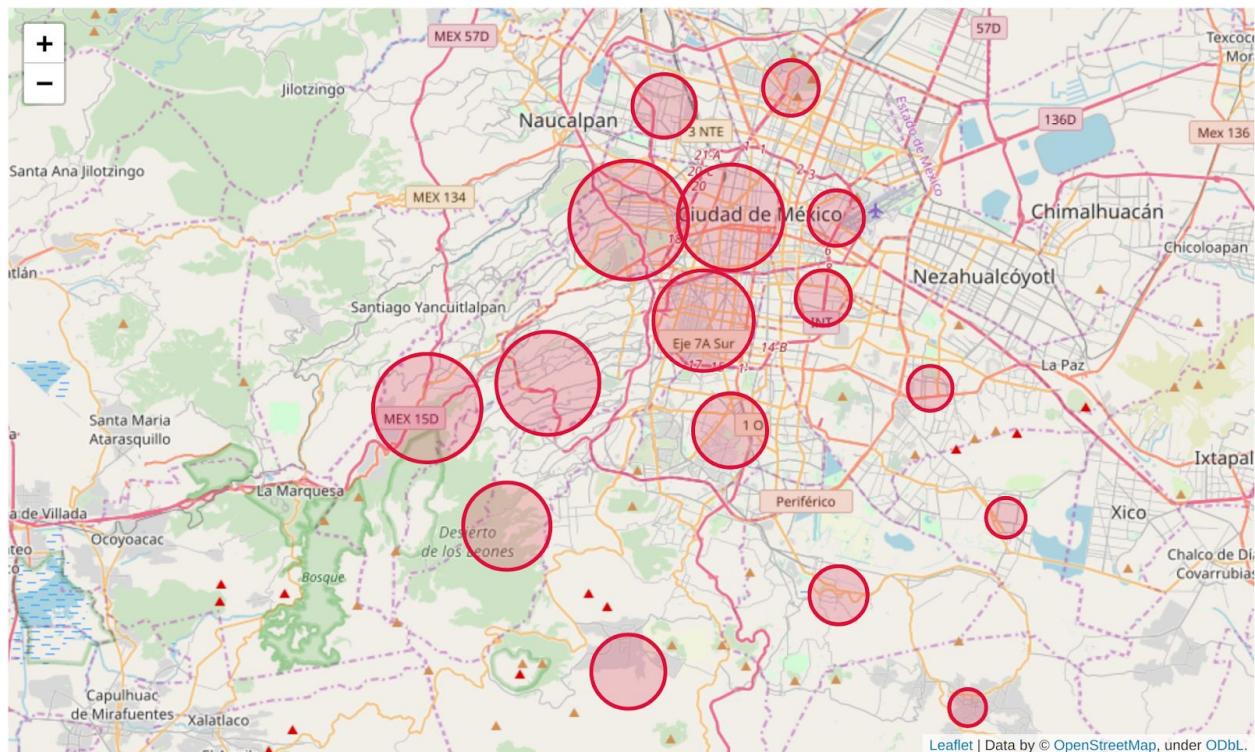
A partir del siguiente mapa, se puede notar que hay una gran concentración en el centro de la capital nacional, donde el precio de las casas resulta ser más alto.



Promedio de precio/m² de apartamentos de las ciudades del Distrito Federal



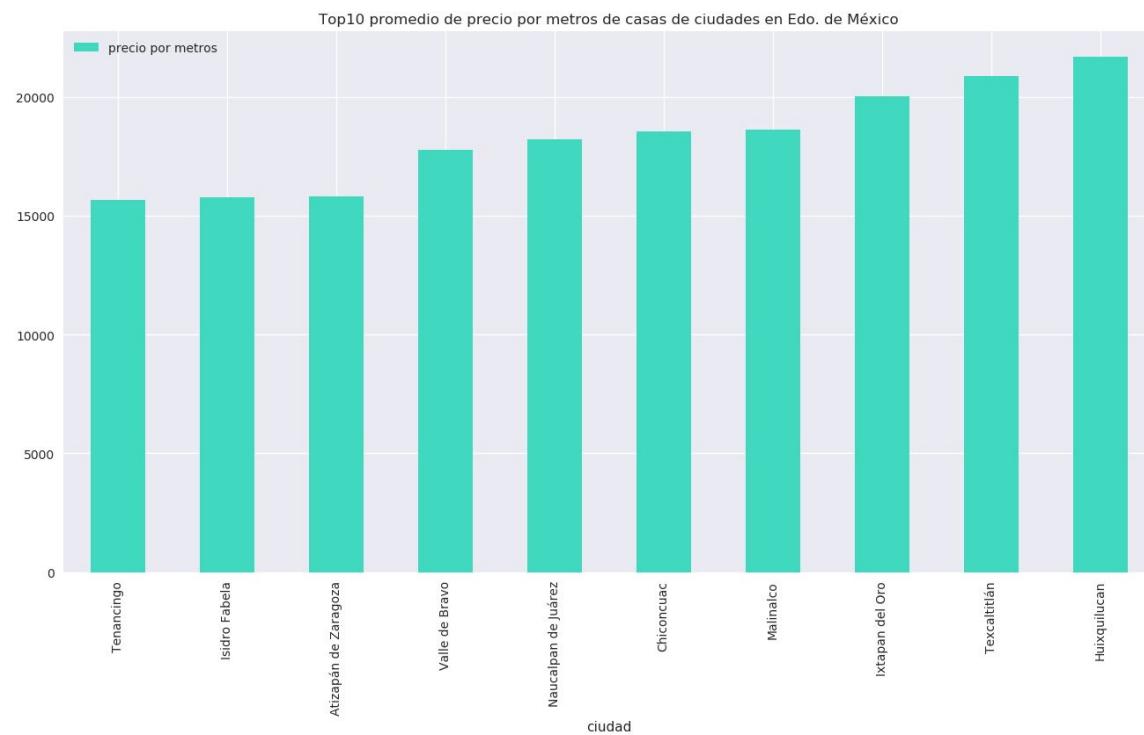
En el gráfico se puede observar que los resultados de los precios promedios por metros de apartamentos son parecidos que los de casas, ya que el primero es triple que el último.

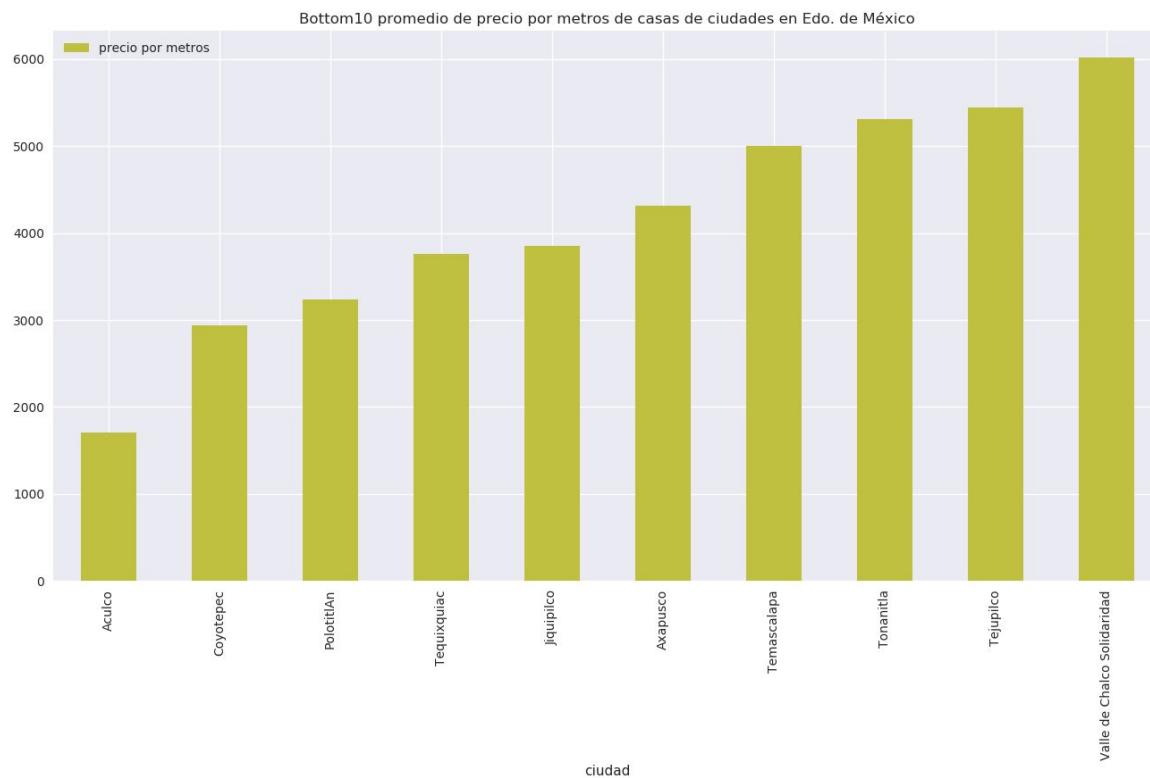


Después de comparar ambos mapas, se puede decir que ambos mapas tienen una distribución bastante parecida, la mayor parte de las ciudades más caras está concentrada en el centro de la provincia, y particularmente, el centro de la capital nacional.

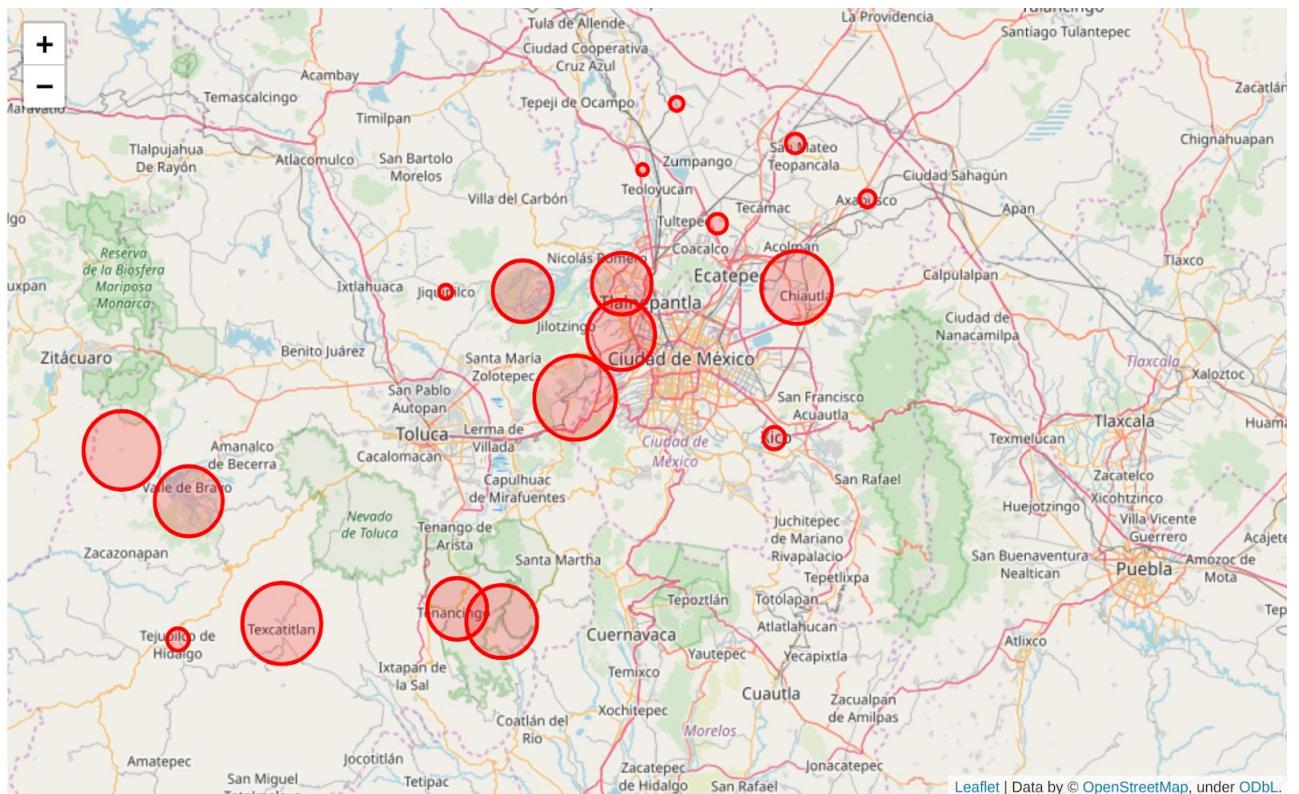
Promedio de precio/m² de casas de las ciudades del Edo. de México

Después de analizar los top 10 y bottom 10 de los promedios de precios/m² de casas en las ciudades del Edo. de México se obtuvo el siguiente resultado:

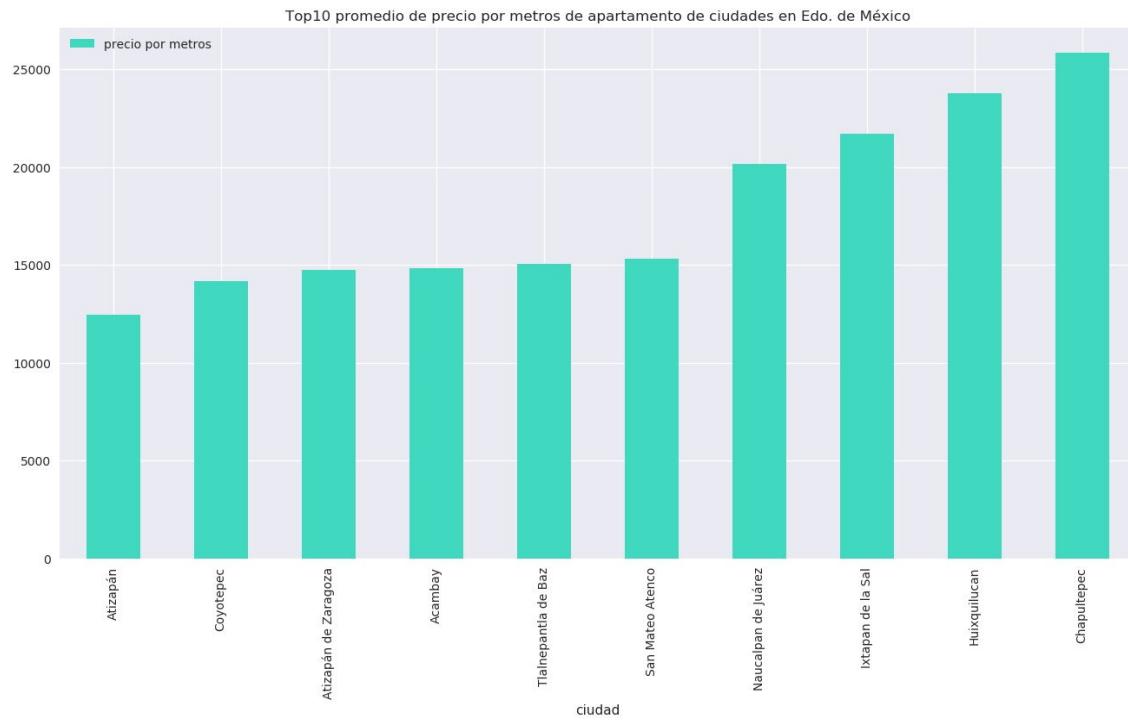


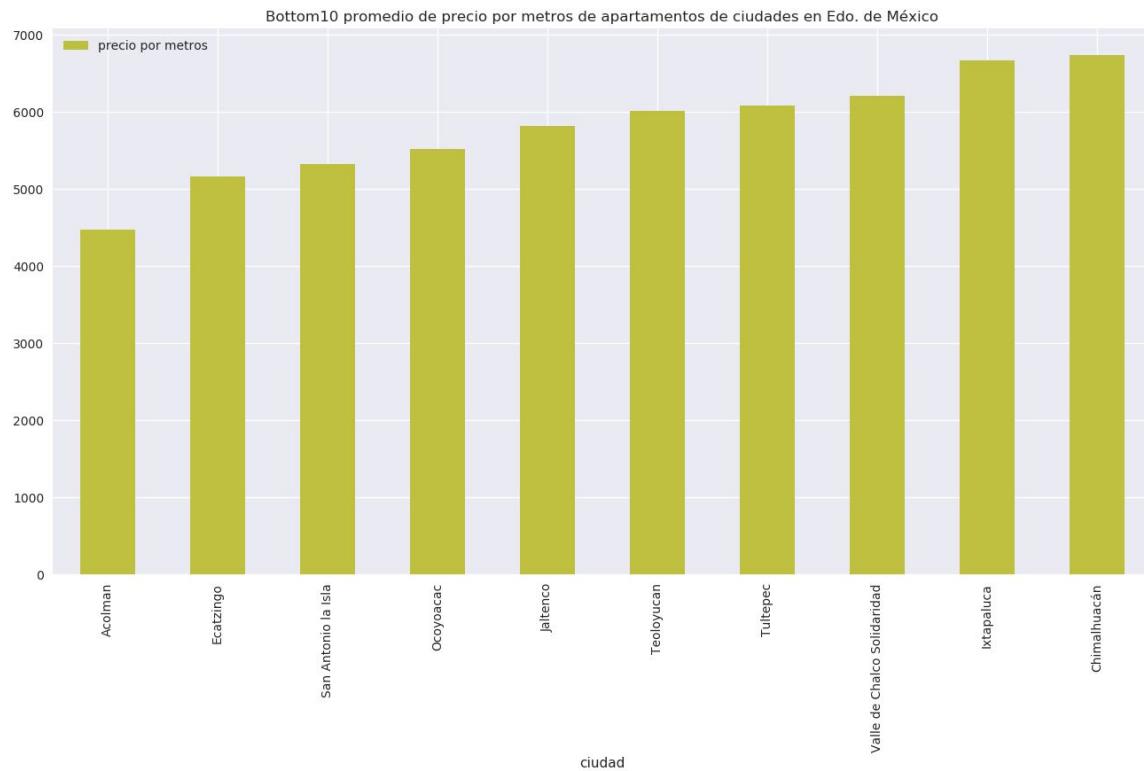


En el gráfico del top 10, se puede observar que las diferencias de los precios entre las ciudades son pequeñas, pero en bottom 10, se ve una diferencia de precio significativa.

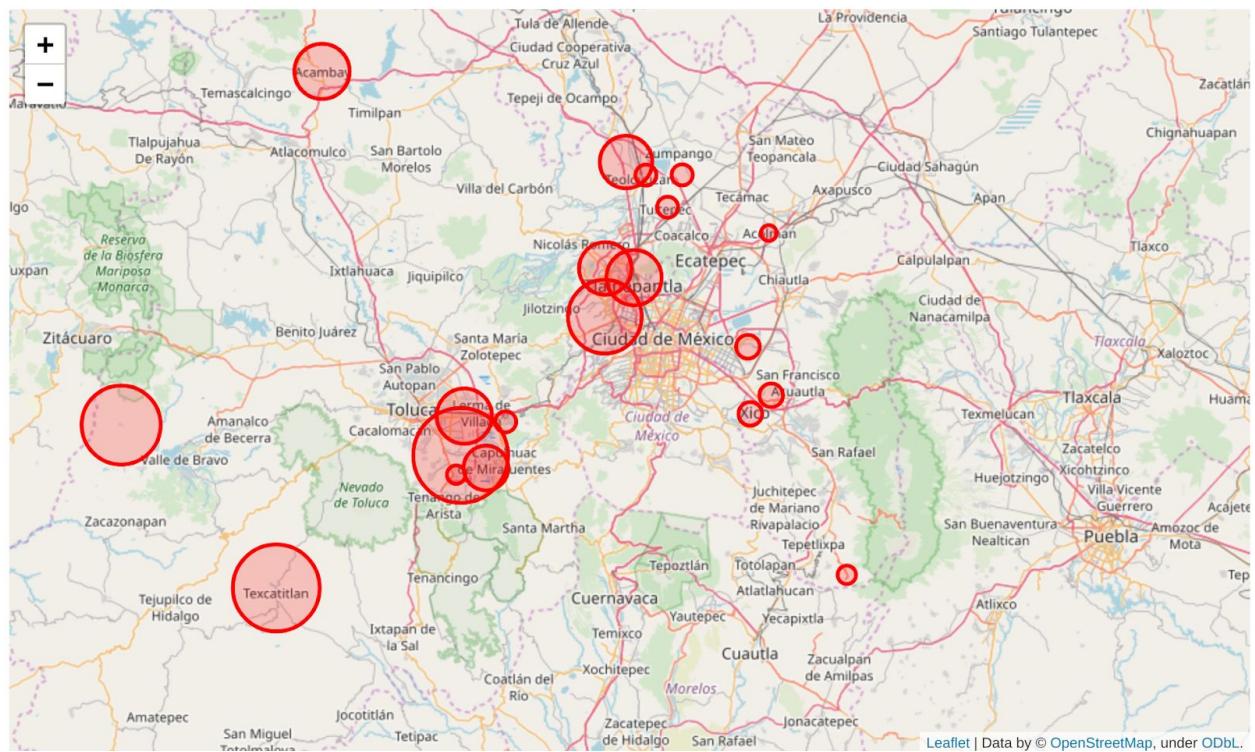


En el mapa, se nota que no tiene una distribución centrada en la provincia. Pero se puede ver que las ciudades con casas más caras están en la parte sureste de la provincia.

Promedio de precio/m² de apartamentos de las ciudades del Edo. de México



Comparando con los gráficos de casa, los gráficos de apartamentos tienen un resultado distinto, en top 10 se ve una diferencia clara entre los precios de las ciudades, pero en el bottom 10, las diferencias entre sí son pequeñas.



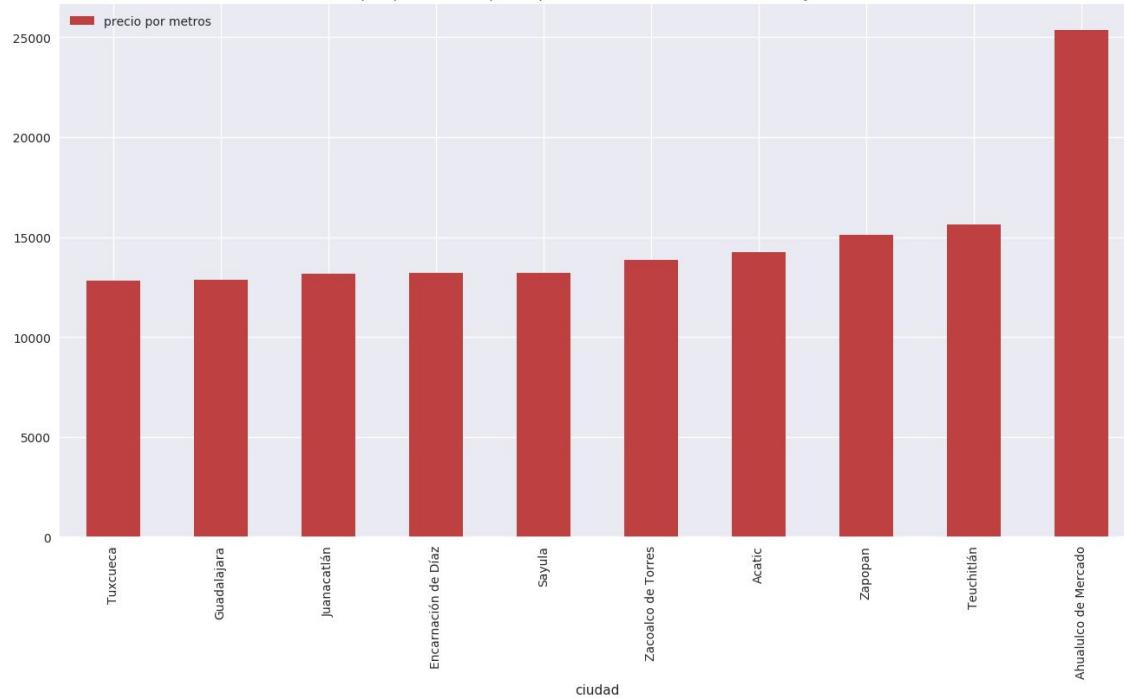
A partir del mapa obtenido, se puede ver que las ciudades con apartamentos más caros están centrados en dos zonas, y las ciudades con apartamentos más baratas también están concentrados en dos zonas, una de ellas al noroeste del DF.

Promedio de precio/m² de casas de las ciudades de Jalisco

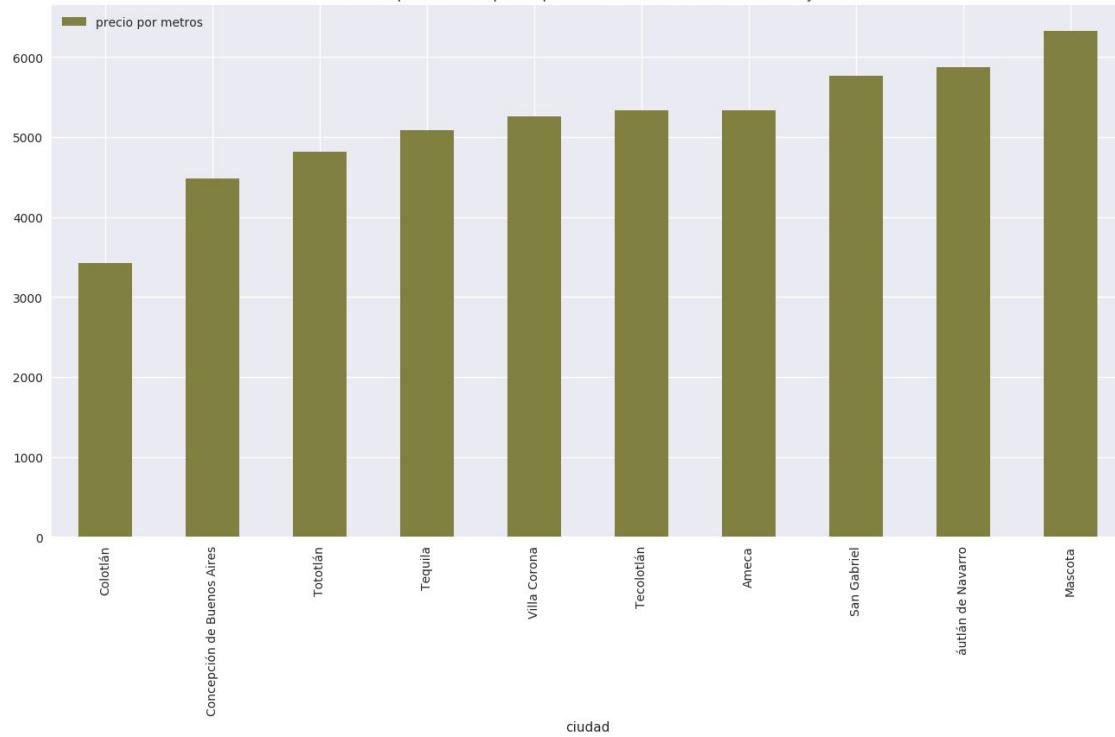
ORGANIZACIÓN DE DATOS: Trabajo Práctico N°1

GRUPO 10
Colavecchia - Huang - Klein

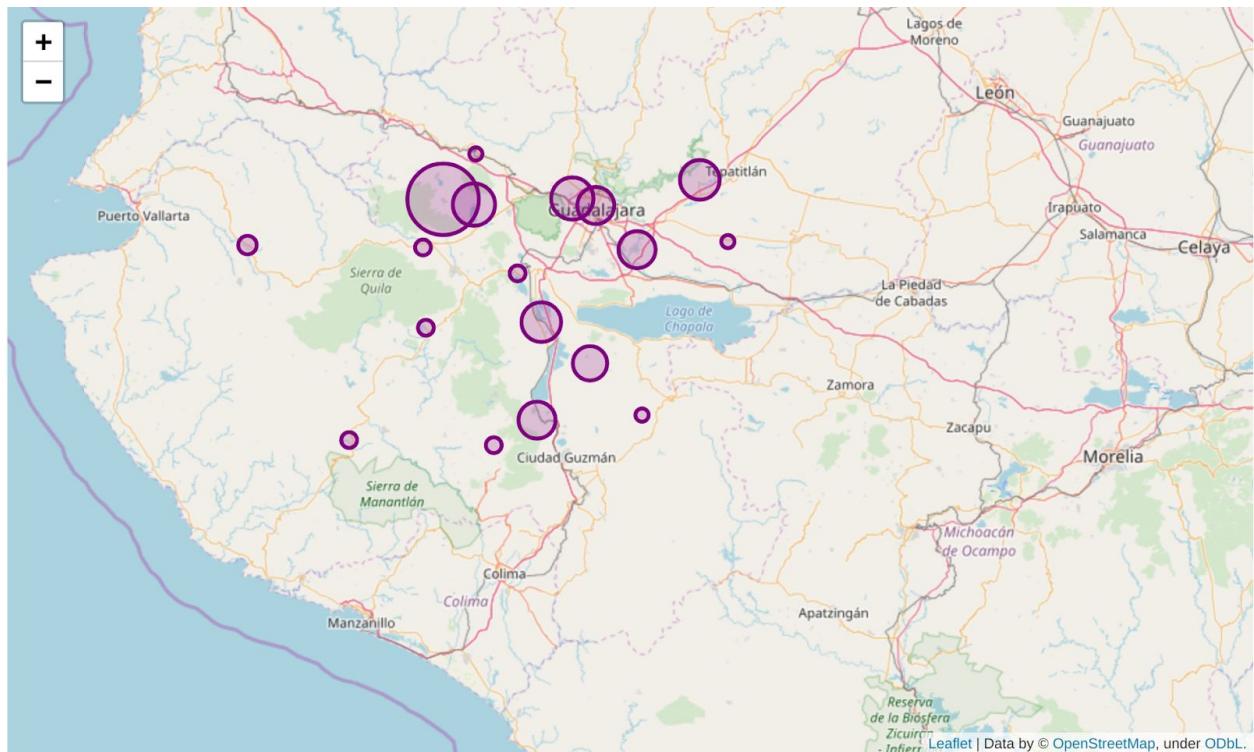
Top10 promedio de precio por metros de casas de ciudades en Jalisco



Bottom10 promedio de precio por metros de casas de ciudades en Jalisco



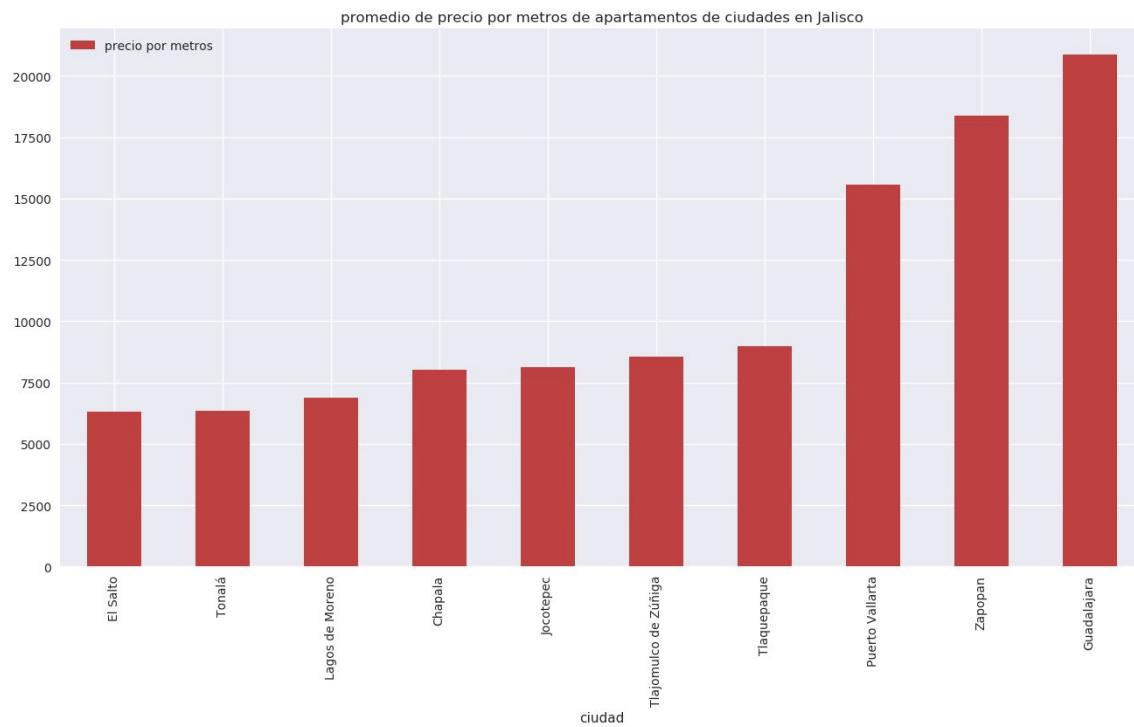
Al comparar los gráficos de top 10 y bottom 10, en el top 10 se puede ver que existe una gran diferencia entre el precio de las casas de la ciudad más cara y las demás. También, en el bottom 10, el precio de la primera es casi doble que el de la última.



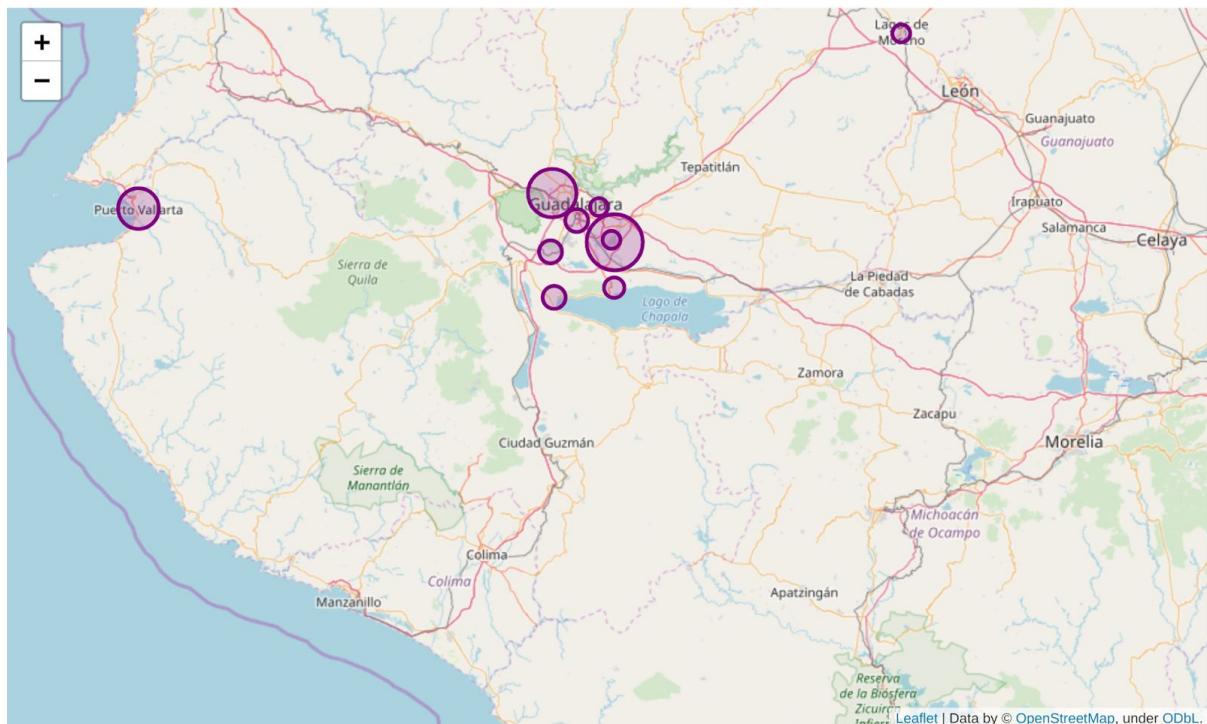
Después de observar el mapa, se puede decir que no existe una distribución muy clara.

Promedio de precio/m² de apartamentos de las ciudades de Jalisco

Después de analizar los datos de apartamentos de la provincia Jalisco, se obtuvieron sólo 10 ciudades, y se representan el mismo gráfico:



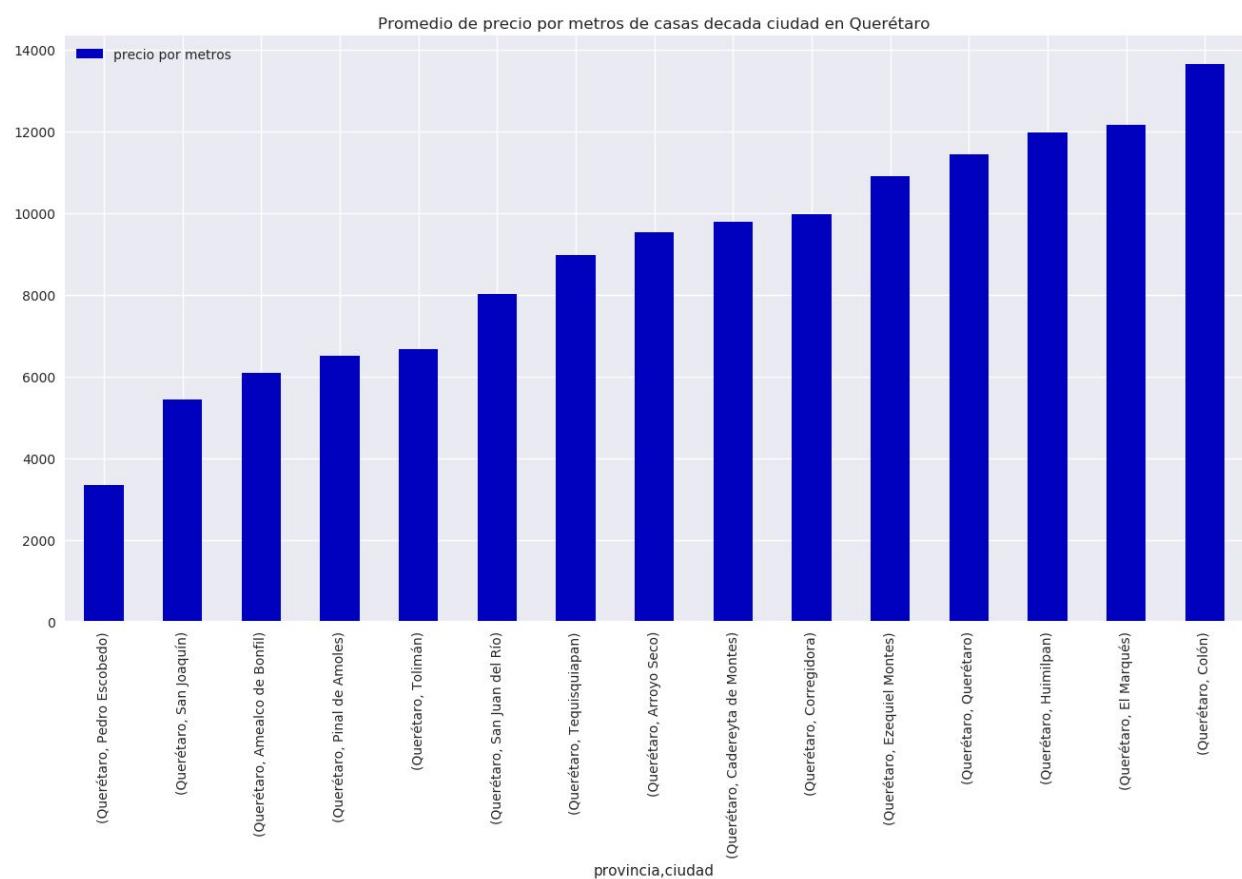
Se puede Observar que los primeros tres precios de apartamentos de las ciudades más caras son mucho mayores que el del resto.



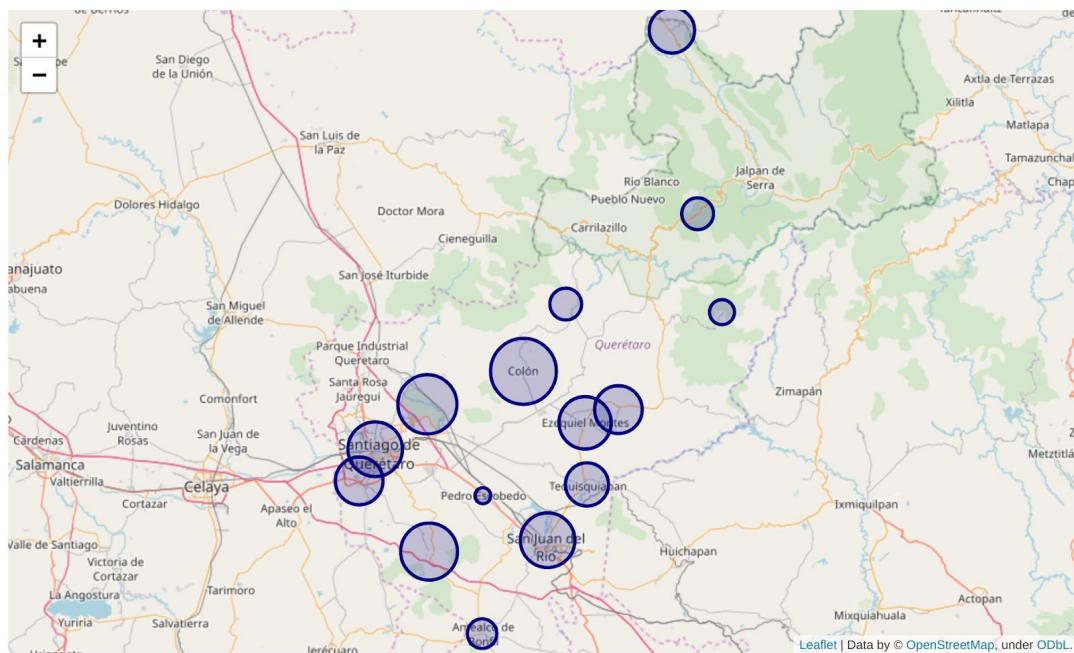
Con el mapa se puede ver que la mayoría está concentrada en el centro de la provincia, a diferencia del mapa de los precios de las casas, en el que se encuentran ubicadas de manera más dispersa. Esto obedece a que los grandes centros urbanos tienden a concentrar la cantidad de apartamentos..

Promedio de precio/m² de casas de las ciudades de Querétaro

Se analiza los precios promedios de las casas de cada ciudad en Querétaro, se obtuvo el siguiente resultado:

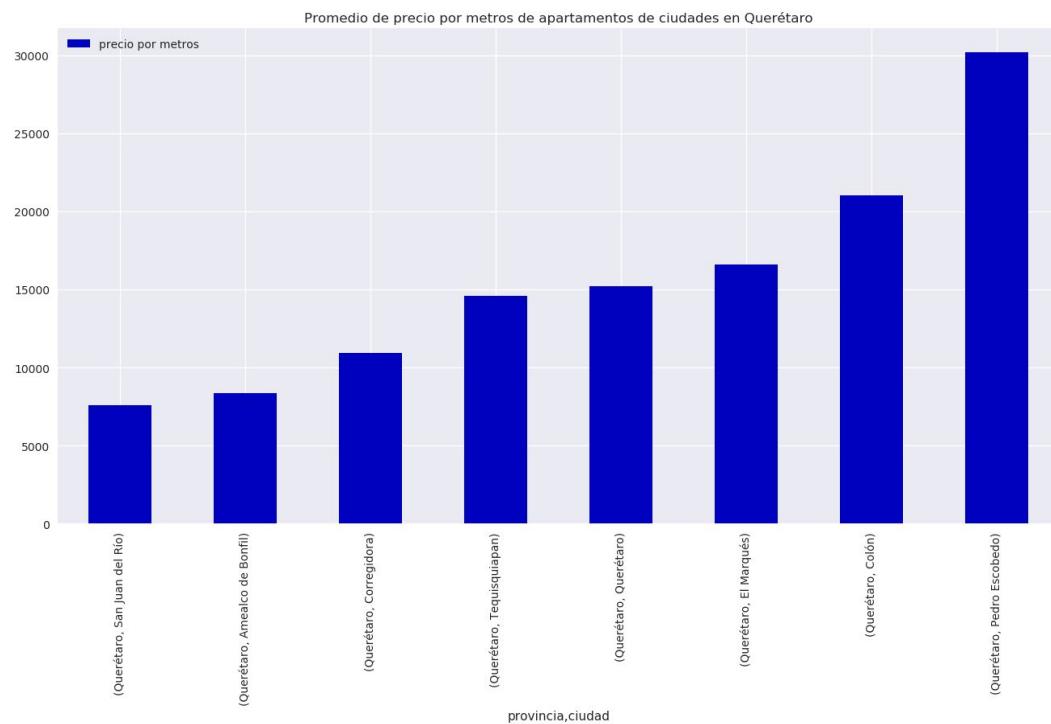


Según el gráfico, se puede ver que hay una diferencias entre las ciudades, salvo entre la segunda y tercera donde la diferencia de precio es mucho menor al resto de la tabla.

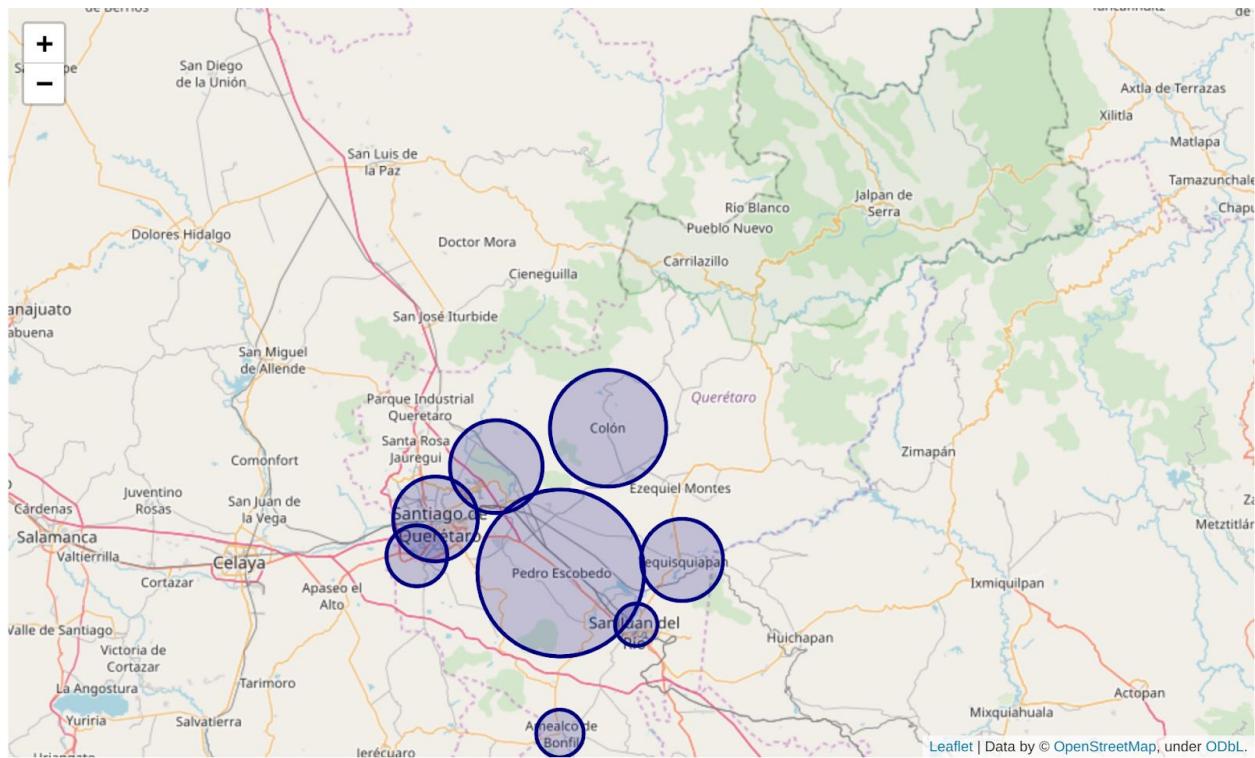


Se puede observar que las ciudades con casas más caras están nuevamente alrededor del centro de la provincia.

Promedio de precio/m² de apartamentos de las ciudades de Querétaro

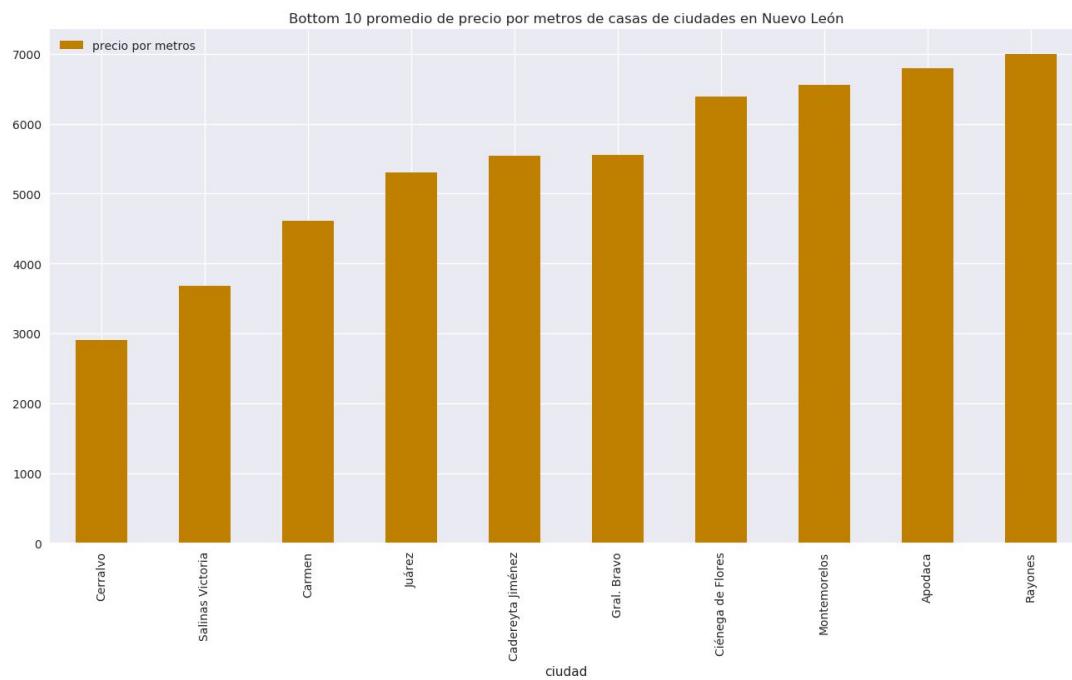
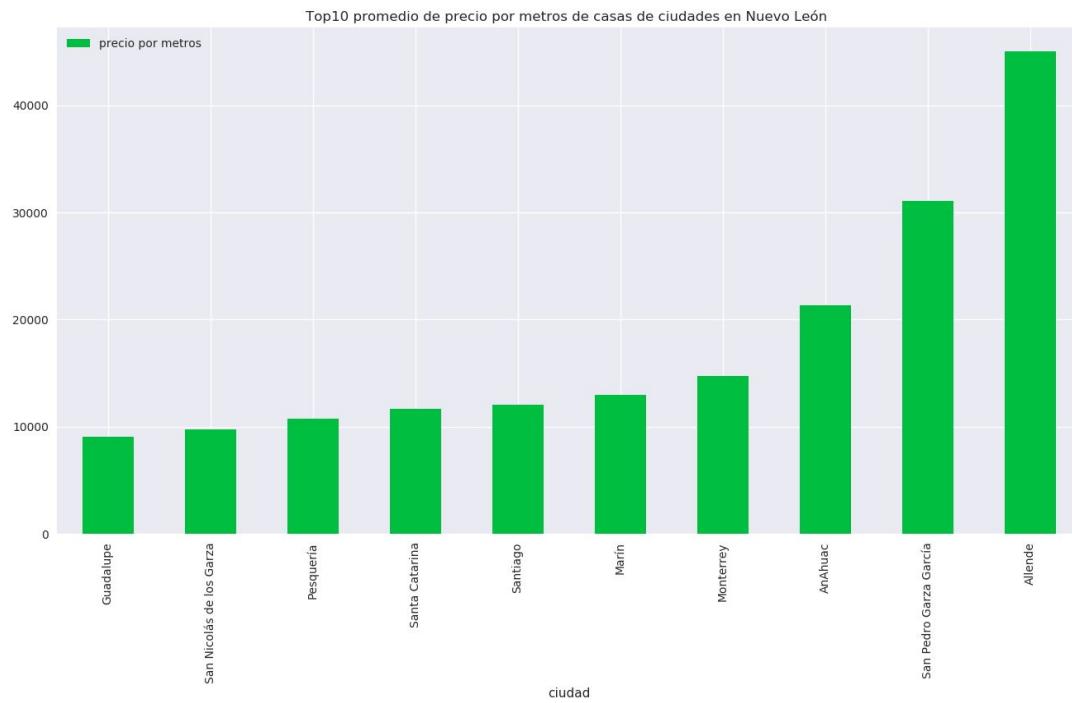


Con el gráfico se puede ver que entre la primera y la segunda existe una diferencia mayor que los restos.

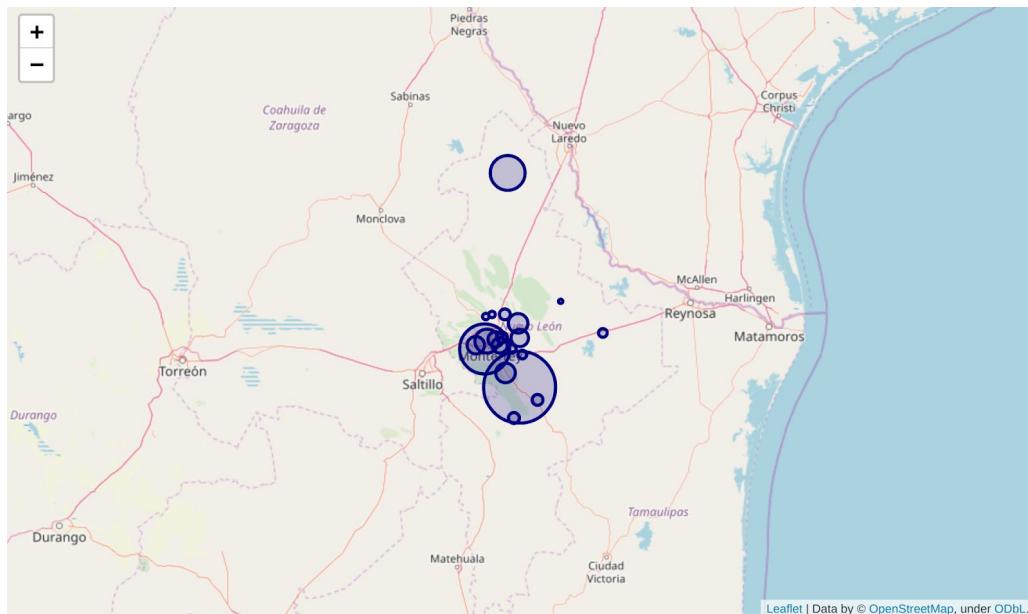


Comparando con el mapa de casas, los apartamentos más caros se encuentran ubicados en la zona Sureste.

Promedio de precio/m² de casas de las ciudades de Nuevo León



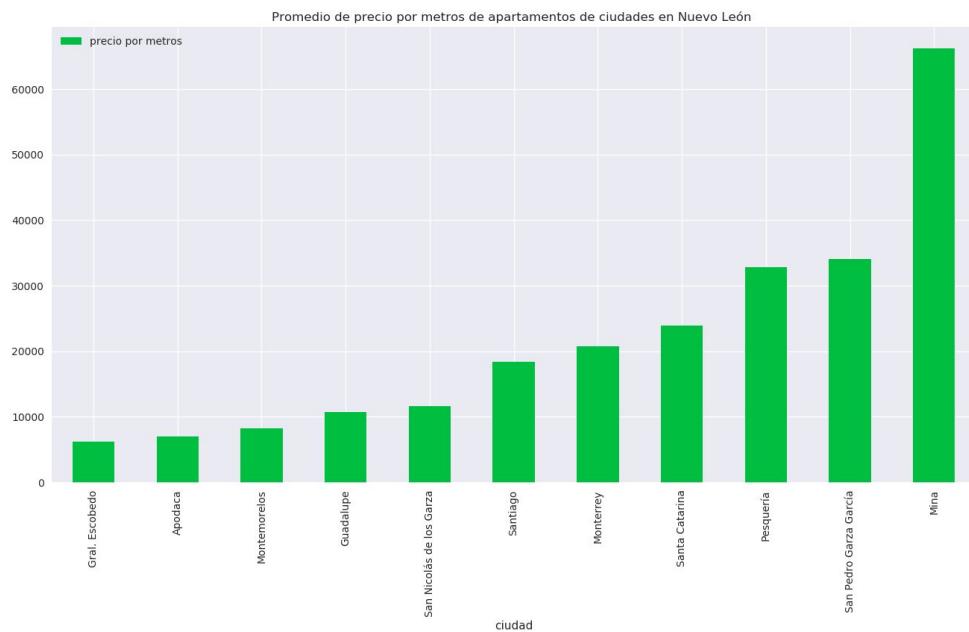
En el top 10, las dos ciudades con casas más caras de Nuevo León presentan un precio promedio/m² que se destaca por sobre el resto. La ciudad más barata, asimismo, lo es en gran medida.



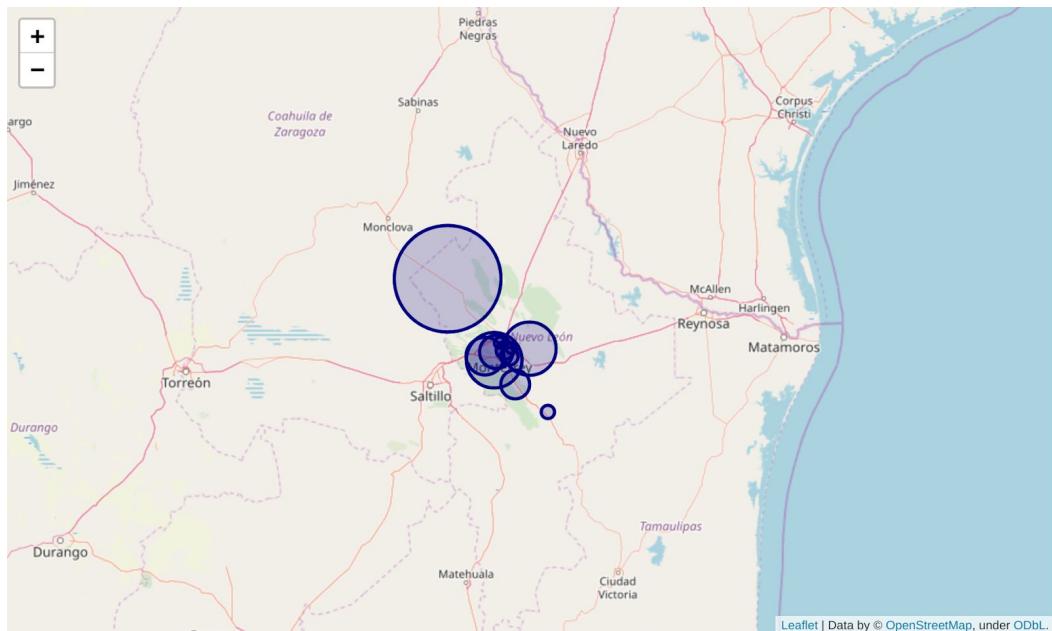
En el mapa se puede ver que la mayoría de las ciudades se encuentran ubicadas en el centro de la provincia.

Promedio de precio/m² de apartamentos de las ciudades de Nuevo León

En la provincia Nuevo León, solo en 11 ciudades hay publicaciones de ventas de apartamentos, por lo tanto, se deciden analizar en el mismo gráfico:



En el gráfico se puede ver, entre las ciudades hay varios grupos con precio totalmente distinto, y también existe una diferencia enorme entre la primera y el último



Comparando los dos mapas, se puede ver que los apartamentos de las ciudades se encuentran ubicados en la misma zona, la cual está cerca del centro de la provincia.

Conclusión respecto al análisis de ciudades

Después de haber analizado precio promedio de todas las propiedades , de las casas y de los apartamentos, se puede ver que la mayoría de los grupos se encuentran ubicados en zonas distintas del país. También, es apreciable que los distintos tipos de las propiedades influyen mucho en el precio.

Por último, es clara la tendencia en las provincias (salvo pocas excepciones), de concentrar apartamentos y casas (en menor medida estas últimas), en los grandes centros urbanos que corresponden a las capitales de las mismas.

Análisis de las descripciones

¿Qué palabras fueron las más utilizadas?



A partir de la realización del WorkCloud, tomamos dimensión de que las descripciones están contaminadas con símbolos y caracteres de formato.

Conclusiones

A partir del análisis exploratorio realizado, llegamos a las siguientes conclusiones:

- La mayor cantidad de publicaciones son de inmuebles de tipo Casa, Apartamento y, en menor medida, Casa en Condominio.
- Aquellas propiedades que cuentan con amenities resultaron ser más caras que aquellas que no.
- Existe una relación directa entre las cantidades de habitaciones, baños, y garajes con el precio/ m^2 , salvo monoambientes.
- En general, aquellas propiedades que ofrecen más comodidades son más costosas.
- La mayoría de las publicaciones constan de inmuebles de entre 50 m^2 y 150 m^2 , el precio de las mismas está concentrado entre 1000000 \$MX y 3500000 \$MX.
- Es notoria la pérdida de valor de un inmueble con el correr de los años, a excepción de aquellos que presentan un valor histórico, donde repunta su precio.
- El precio de las propiedades de tipo casa , apartamentos y casa en condominio tiene una relación directa con la cercanía de las escuelas y centros comerciales, siendo más marcada para estos últimos.
- Es clara la tendencia en las provincias de concentrar apartamentos y casas en los grandes centros urbanos que corresponden a las capitales de las mismas. Dicha tendencia es muy significativa en especial para los apartamentos.
- El Distrito Federal y el Edo. de México concentran la gran mayoría de inmuebles, en proporción a sus densidades poblacionales. El resto de los estados concentran en sus respectivas capitales la mayor cantidad de publicaciones.
- No se observó una correlación entre el PBI per cápita por provincia y el precio por m^2 por provincia.