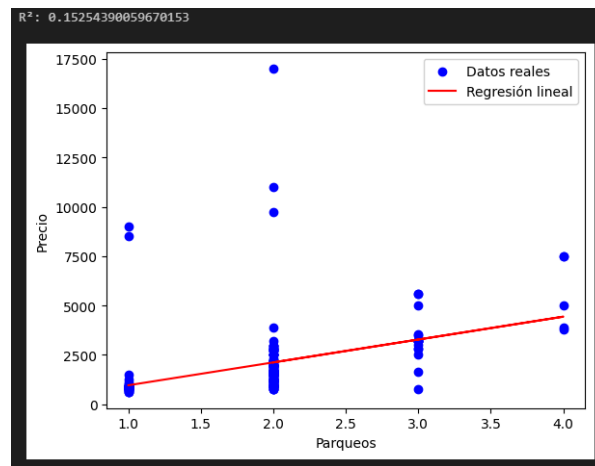


## Regresión Lineal Simple del Web Scraping

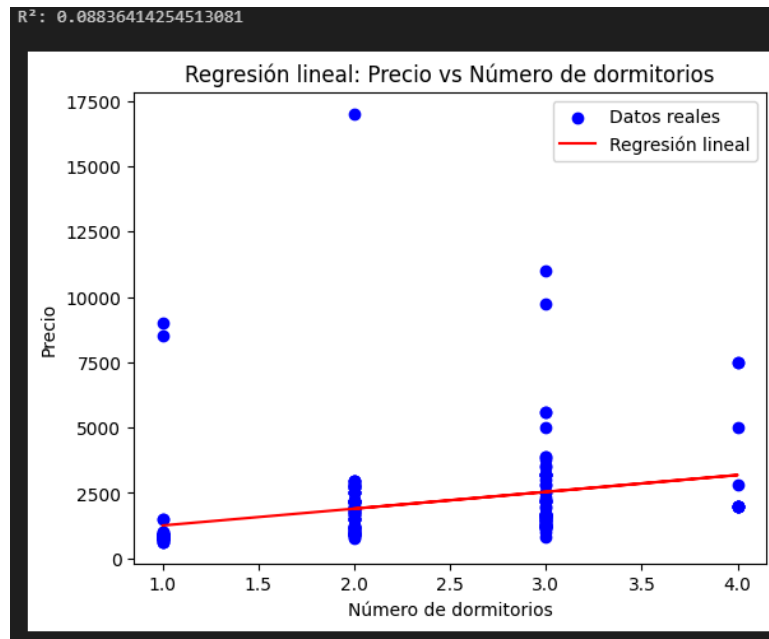
Previamente se realizó un Web Scraping de la página Mapainmuelle, específicamente de la sección de apartamentos en alquiler de la Zona 14. El Scraping se realizó para las primeras 25 páginas de la sección, resultando en un total de 214 propiedades. De estas se rescataron las siguientes características: precio, número de parqueos, número de dormitorios, número de baños, ID de la propiedad, tamaño, y el enlace. Para la parte que se desarrollará en este documento, las características importantes son el precio, número de dormitorios, número de baños, tamaño y número de parqueos. Esto porque se realizaron cuatro regresiones lineales simples, usando como dependiente al precio, e independiente a las otras 3 variables relevantes.

### Precio y Número de Parqueos



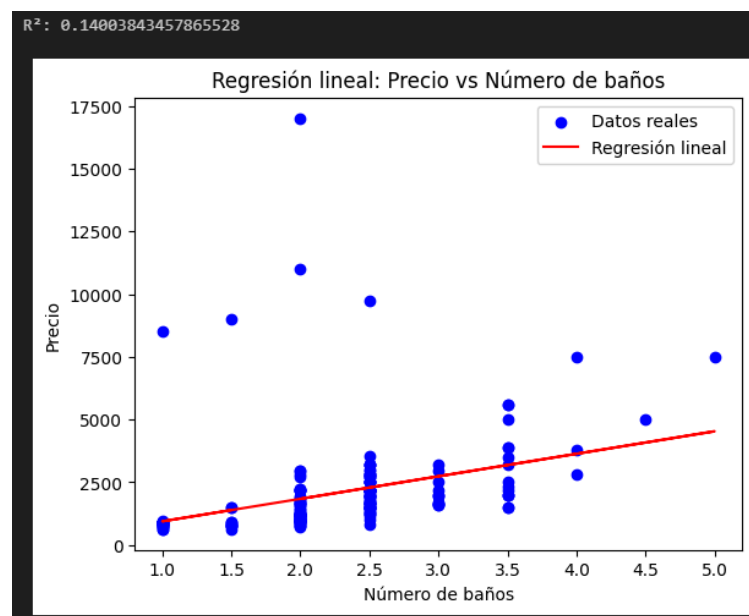
En este caso se puede ver que sí existe algo de relación pues la regresión es positiva. Esto indica que entre más cara es la propiedad más parqueos se pueden esperar, sin embargo se puede ver que hay varios casos en los que por diferentes precios hay la misma cantidad de parqueos. Por lo tanto, no se puede esperar una correlación bastante fuerte, al igual de como lo muestra el  $R^2$ , pues muestra que el modelo sólo representa el 15% de la varianza.

## Precio y Número de Dormitorios



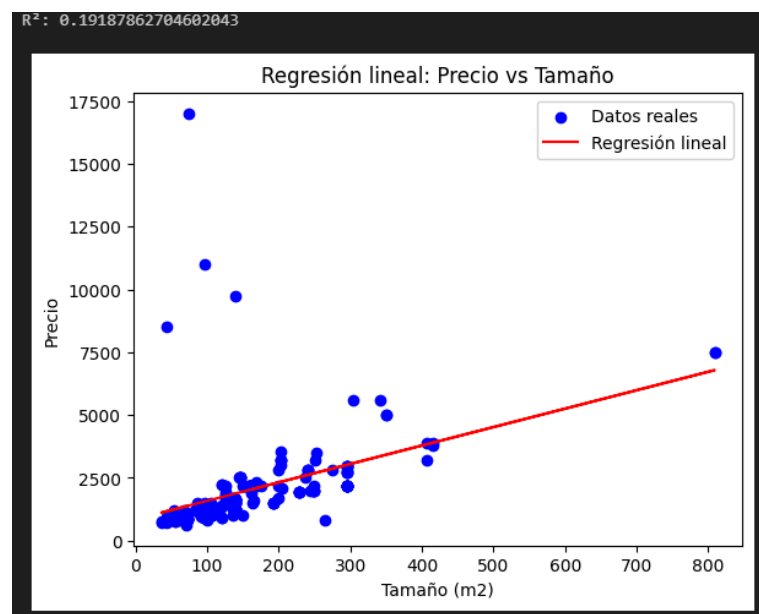
En este caso se puede ver que también existe algo de relación positiva, mientras más cara sea la propiedad sí se pueden esperar una mayor cantidad de dormitorios. Sin embargo, también hay varios casos en los que por un mismo precio hay la misma cantidad de dormitorios. Curiosamente, es bastante similar al de la cantidad de parqueos. Similarmente, se puede esperar algo de relación, pero no la más fuerte, pues el R<sup>2</sup> muestra que el modelo únicamente representa el 9% de la varianza.

## Precio y Número de Baños



En este caso se puede ver que hay una mayor variedad de cantidades, gracias a los baños sin ducha. Similarmente, la relación es positiva y en este caso se puede ver de manera más evidente la relación. Entre más cara sea la propiedad se puede esperar una mayor cantidad de baños. En este caso son menores los casos en los que una misma cantidad es compartida por un gran rango de precios, sin embargo aún existen varios casos. Por lo tanto, no se puede esperar una correlación demasiado fuerte, al igual de como el  $R^2$  lo muestra, al demostrar que el modelo sólo representa el 14% de la varianza.

## Precio y Tamaño



Este caso también tiene relación positiva, y esta es la relación que es más evidente a primera vista. Se puede ver más fácilmente como entre mayor es el precio, aumenta el tamaño de la propiedad. Sin embargo, aún se pueden ver *outliers* y algunos casos que entre rangos distintos de precios comparten un mismo tamaño. A pesar de ser el modelo con mayor correlación, el  $R^2$  muestra que el modelo sólo representa el 19% de la varianza.

Se puede ver que todas las regresiones muestran una relación positiva, por lo tanto todas estas variables se pueden considerar a la hora de determinar el precio de una propiedad. Sin embargo, es importante notar que una regresión simple no tiene la suficiente información como para crear un modelo lo suficientemente significativo. Por

lo tanto, para obtener mejores resultados, se podrían combinar estas cuatro variables que ya demostraron cierta correlación, para crear una regresión múltiple. En la cual se pueda obtener un modelo de regresión con una mayor precisión.