



UNIVERSIDAD DE LOS ANDES
FACULTAD DE INGENIERÍA Y
CIENCIAS APLICADAS

MODELO PARA EXPLICAR EL VALOR
DEL SUELO PARA PROYECTOS
INMOBILIARIOS EN LA COMUNA DE
PROVIDENCIA

KATALINA ANDREA JORQUERA CÁRDENAS

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL INDUSTRIAL

PROFESOR GUÍA: SANTIAGO TRUFFA

SANTIAGO, FEBRERO DE 2023

UNIVERSIDAD DE LOS ANDES
FACULTAD DE INGENIERÍA Y
CIENCIAS APLICADAS

AGRADECIMIENTOS

En primer lugar, quiero agradecer a mis padres por darme la oportunidad de estudiar y por su apoyo incondicional a lo largo de la carrera universitaria. Por estar presentes en los momentos que los necesitaba y motivarme a seguir estudiando y dar lo mejor de mí.

En segundo lugar, agradecer a mis amigos que me han acompañado durante esta etapa universitaria y que hicieron de estos años un período inolvidable.

Por último, quiero agradecer a Christian Wegmann y Santiago Truffa, por haber confiado en mi trabajo y por brindarme diversos conocimientos a lo largo del año. Fueron un gran apoyo durante este período, sin ellos no hubiese logrado tan fácilmente este trabajo.

ÍNDICE

RESUMEN	VII
1. INTRODUCCIÓN	1
1.1. CONTEXTO	1
1.2. PROBLEMA	2
1.3. MOTIVACIÓN	3
1.4. OBJETIVOS	3
1.4.1. Objetivo general	3
1.4.2. Objetivos específicos	4
1.5. ALCANCES	4
1.6. METODOLOGÍA	5
2. MARCO TEÓRICO	6
2.1. INVESTIGACIÓN PREVIA	6
2.2. PRECIO DEL SUELO Y FACTORES QUE LO DETERMINAN	6
2.3. PROYECTO INMOBILIARIO	8
2.4. NORMAS URBANÍSTICAS	8
2.5. REGRESIÓN MÚLTIPLE L	9
2.6. R-CUADRADO, TEST-T Y P-VALUE	9
2.6.1. R-cuadrado y R-cuadrado ajustado	9
2.6.2. Test-T	10
2.6.3. P-value	11
2.7. ESTADÍSTICO SOFTWARE R	¡ERROR! MARCADOR NO DEFINIDO.
2.8. GRÁFICO BOXPLOT	¡ERROR! MARCADOR NO DEFINIDO.
3. LEVANTAMIENTO DE DATOS Y GENERACIÓN DE BASE DE DATOS	14
3.1. LEVANTAMIENTO DE DATOS	14
3.1.1. Levantamiento de los permisos	14
3.1.2. Roles	15
3.1.3. Transacción	17
3.2. GENERACIÓN DE LA BASE DE DATOS	19
3.2.1. Permisos de edificación	19
3.2.2. Análisis descriptivo de los precios unitarios	21
4. ANÁLISIS DE DATOS	24
4.1. ANÁLISIS DE COMPORTAMIENTO NORMAL	24
4.2. PAQUETE CARET	29
4.3. MATRIZ DE AUTOCORRELACIÓN	30
5. ANÁLISIS REGRESIÓN LINEAL MÚLTIPLE	33
5.1. EVALUACIÓN NORMALIDAD Y HOMOSCEDASTICIDAD DE LA REGRESIÓN	37
5.2. EVALUACIÓN DEL MODELO	40
5.3. EVALUACIÓN DE LOS COEFICIENTES	41
6. CONCLUSIONES Y RECOMENDACIONES	50
7. BIBLIOGRAFÍA	54

8.	ANEXOS.....	58
8.1.	ANEXO 1	58
8.2.	ANEXO 2.....	58
8.3.	ANEXO 3.....	60
8.4.	ANEXO 4.....	61
8.5.	ANEXO 5.....	61
8.6.	ANEXO 6.....	62

ÍNDICE DE TABLAS

TABLA 1: CANTIDAD DE PROYECTOS INMOBILIARIOS POR AÑO.....	15
TABLA 2: CANTIDAD DE ROL POR AÑO.....	16
TABLA 3: CANTIDAD DE TRANSACCIONES POR AÑO	18
TABLA 4: ANÁLISIS PRECIO UNITARIO POR AÑO.....	21
TABLA 5: INTERPRETACIÓN ESTIMACIÓN BETA	43
TABLA 6: COMPARACIÓN PROYECTO CON ALTO Y BAJO VALOR UNITARIO	48

ÍNDICE DE ILUSTRACIONES

ILUSTRACIÓN 1: REPRESENTACIÓN DIAGRAMA DE CAJAS Y BIGOTES; ERROR! MARCADOR NO DEFINIDO.	
ILUSTRACIÓN 2: VALOR DE LAS TRANSFORMACIONES.....	25
ILUSTRACIÓN 3: RESUMEN ESTADÍSTICO	25
ILUSTRACIÓN 4: RESULTADO DE LA REGRESIÓN LINEAL CON LAS VARIABLES IMPORTANTES	34
ILUSTRACIÓN 5: VARIABLES PROCESO DE SELECCIÓN	35
ILUSTRACIÓN 6: RESULTADO DE LA REGRESIÓN LINEAL CON LAS VARIABLES PREDICTORAS.....	35
ILUSTRACIÓN 7: RESULTADO FINAL DE LA REGRESIÓN LINEAL MÚLTIPLE.....	36
ILUSTRACIÓN 8: PRUEBA DE SHAPIRO.....	37
ILUSTRACIÓN 9: PRUEBA DE WHITE.....	38
ILUSTRACIÓN 10: REGRESIÓN CON ERRORES ESTÁNDARES ROBUSTOS	39
ILUSTRACIÓN 11: PLANO PROVIDENCIA.....	52
ILUSTRACIÓN 12: CAPTURA DE PANTALLA DE LOS DATOS DE TRANSACCIÓN.....	58
ILUSTRACIÓN 13: CAPTURA DE PANTALLA BASE DE DATOS	59
ILUSTRACIÓN 14:GRÁFICO IMPORTANCIA DE LAS VARIABLES	60
ILUSTRACIÓN 15: MATRIZ DE CORRELACIÓN VALOR UNITARIO VS COEFICIENTE DE CONSTRUCTIBILIDAD.....	61
ILUSTRACIÓN 16: MATRIZ DE CORRELACIÓN VALOR UNITARIO VS SUPERFICIE DEL TERRENO....	61
ILUSTRACIÓN 17: REGRESIÓN LINEAL SIN SUPERFICIE TOTAL DEL TERRENO.....	62

ÍNDICE DE GRÁFICOS

GRÁFICO 1: RELACIÓN PRECIO/AÑO.....	22
GRÁFICO 2: GRÁFICO DE CAJAS Y BIGOTES PRECIO POR AÑO	23
GRÁFICO 3: HISTOGRAMA DEL VALOR UNITARIO	26
GRÁFICO 4: HISTOGRAMA VALOR UNITARIO TRANSFORMADA A LOGARITMO	27
GRÁFICO 5: HISTOGRAMA VALOR UNITARIO TRANSFORMADA A CUADRÁTICA	27
GRÁFICO 6: VARIABLE CATEGÓRICA.....	29
GRÁFICO 7: MATRIZ DE CORRELACIÓN.....	31
GRÁFICO 8: RESIDUOS VS VALOR AJUSTADO	38

RESUMEN

El precio del suelo es un componente importante para tasar el valor de la vivienda, ayudando a comprender los incrementos que pueda presentar en su valor. Además, es un insumo para construir no solo viviendas, sino también oficinas y comercios. El objetivo de este trabajo es crear un modelo para explicar el valor del suelo para proyectos inmobiliarios entre los años 2017 hasta el 2021.

En este proyecto, se pretende investigar el impacto que tienen las variables más relevantes que se asocian al precio unitario. Para averiguar cuáles son estas, se hace un levantamiento de datos donde primero se evalúa los permisos correspondientes a proyectos inmobiliarios, para luego levantar la cantidad de roles de ese permiso, evaluar las transacciones y finalmente extraer los datos que se encuentran en un permiso de edificación como las normas urbanísticas, la superficie de los terrenos, entre otros.

Para determinar cuáles de los datos extraídos son las más importantes, se utiliza un paquete de R llamado Caret que nos muestra a través de un gráfico cuáles son las variables de mayor importancia para el modelo con respecto al precio unitario. A partir de estas variables se realiza una selección de los mejores predictores. Se emplea la estrategia de stepwise mixto, y el valor matemático empleado para determinar la calidad del modelo va a ser Akaike (AIC). Esto nos retorna las variables más relevantes y las que mejor predicen el modelo.

Teniendo en cuenta esto, se realiza un modelo de regresión lineal múltiple con estas variables, siendo el plazo entre la primera y última transacción, la superficie del terreno, el año en que se procedió la última transacción y la normativa DFLN°2 las más cruciales y significativas del modelo. Además de obtener como significante la zona censal 2_5 correspondiente al sector de Tobalaba y la zona censal 2_7 correspondiente al sector cercano al parque Inés de Suárez.

Se concluye que el modelo generado es significativo, ya que retorna un valor p menor a 0,05, por lo que se rechaza la hipótesis nula. Además de que el modelo se ajusta en un 74,09% con respecto al valor unitario, es decir, las variables predictoras explican en un 74,09% el precio unitario.

1. INTRODUCCIÓN

En este capítulo, se describe el contexto y problema de la comuna donde se desarrolla un modelo para explicar la evolución del precio del suelo para proyectos inmobiliarios, el objetivo general y específico para abordar el proyecto, los alcances y la metodología respectiva.

1.1. CONTEXTO

En los países desarrollados, el precio de las viviendas ha crecido considerablemente más rápido que los ingresos personales y rentas (Parrado, 2009). Un índice de los precios mundiales de la vivienda, elaborado por la consultora inmobiliaria Knight Frank, muestra que se incrementó en un 19% en todo el mundo entre el primer trimestre de 2020 y el primero del 2022, o un 10% luego de ajustar la inflación (The Wall Street Journal, 2022).

De acuerdo con los datos entregados por Knight Frank, entre los países que más han elevado sus precios en vivienda, Chile ocupa el lugar N.º 49 de 56 y registra un alza del 1,7% en 2020 (Generación M, 2021). El precio promedio de la vivienda nueva en Chile subió un 75% entre el 2011 y 2021, por lo que actualmente se sitúa en 3.980 UF. Por lo tanto, el dividendo promedio en el país en UF para comprar propiedad ascendió, en igual lapso, en 20,8 UF (Tinsa, 2021).

Desde la industria afirman que el aumento de precio de las viviendas se debe a las mayores regulaciones a la densificación y uso del suelo. Lo anterior estaría restando capacidad de reacción a la oferta, lo que presiona al alza el valor del suelo y el de las viviendas. Además, existiría un mayor grado de incertidumbre normativa, lo que aumenta el riesgo de ejecución efectiva de los proyectos y, por lo tanto, los encarece. Estos factores, sumados al aumento de costos de construcción junto a nuevas exigencias técnicas, presionaría al alza el precio de venta de las viviendas (Cepchile, 2019).

En la Región Metropolitana el precio del suelo ha aumentado un 218% en la última década (2010-2019). En el año 2010, costaba en promedio 18,4 UF/m² y aumentó a 58,5 UF/m² en el año 2019, lo que conlleva a un crecimiento promedio de 15% anual. Al igual que el precio, la cantidad de terrenos transados en la última década fue en crecimiento, experimentando un aumento

significativo en el año 2015, en donde se pasó de 248 transacciones de suelo en el año anterior a 716 en ese año (Osorio, 2021).

La escasez de suelo en la Región Metropolitana y las mayores regulaciones municipales para el desarrollo de proyectos inmobiliarios, son algunos de los principales componentes que han contribuido al alza en el precio del suelo, lo que se ha reflejado en el aumento reciente en el valor de casas y departamentos nuevos en los últimos años. El efecto del suelo puede representar hasta el 35% del precio de una vivienda, reflejando los mayores impactos en el sector oriente (Gutiérrez, 2015).

Una de las comunas con mayor variación en el precio de las viviendas entre el 2012 al 2017 es Providencia. El precio de los departamentos varió de un 64,9 UF/m² en el 2012 a 90,6 UF/m² en el 2017. La variación anual en esos cinco años alcanzó el 6,9%. Una de las causas de este aumento de precio es la escasez de terreno, además de los costos de construcción (Peña, 2018).

La carencia de series sistemáticas de precios de viviendas, oficinas y terrenos ha dificultado el seguimiento y análisis de las tendencias del mercado inmobiliario y sus consecuencias. Debido a esto, se elaboran diversos indicadores de precios de vivienda utilizando tanto precios de venta como de transacción, y se contrastan con variables macroeconómicas que explicarían su comportamiento (Parrado, 2009).

1.2. PROBLEMA

A raíz de lo dicho anteriormente, no existe un modelo que ayude a entender el precio del suelo, en particular, para proyectos inmobiliarios. Como el suelo es un componente del precio de la vivienda, su valor ayudaría a evaluar con mayor precisión el incremento de esta. Una de las razones por las cuales no hay un registro objetivo del costo del suelo, se debe a que no es un producto homogéneo, por lo tanto, es complejo comparar suelos con diferentes tamaños, formas y condiciones normativas.

1.3. MOTIVACIÓN

En la economía de un hogar, la vivienda representa, generalmente, la tenencia más importante, tanto porque es una medida de riqueza como porque constituye el principal costo que debe enfrentar una familia: gastos de arriendo, mantenciones, servicios asociados, etc. (Banco Central de Chile, 2014).

Una caída abrupta en los precios de viviendas genera, por un lado, una pérdida de riqueza en los hogares, lo que genera una percepción negativa de la situación económica y esto, a su vez, afecta las decisiones de consumo y pago. Por otro lado, aumenta el riesgo de la cartera de créditos de los oferentes de crédito, dado que la vivienda representa un respaldo para los créditos hipotecarios. Entonces, contar con un adecuado monitoreo del sector inmobiliario y de la dinámica de precios de viviendas es crucial para el análisis de estabilidad financiera (Banco Central de Chile, 2014).

Debido a lo mencionado anteriormente, es relevante explicar el precio del suelo, ya que, como es un factor del precio de la vivienda, explica los incrementos de este. Además, es un insumo para construir no solo viviendas, sino también oficinas y comercios.

1.4. OBJETIVOS

En el siguiente capítulo, se presentan los objetivos generales y específicos de esta memoria.

1.4.1. Objetivo general

El objetivo general de esta memoria es evaluar, con herramientas econométricas, un modelo para entender el precio del suelo pagado por los proyectos inmobiliarios aprobados por la municipalidad de Providencia entre los años 2017 y 2021, controlado por los principales factores normativos que rigen en esta comuna.

1.4.2. Objetivos específicos

Con el propósito de conseguir el objetivo general, se definieron los siguientes objetivos específicos:

- a) Obtener la base de datos de permisos de edificación de obra nueva de la comuna de Providencia desde el año 2017 hasta el año 2021, a través de la sección de transparencia activa de la web.
- b) Analizar y clasificar los permisos de edificación de obra nueva para identificar los que corresponde a proyectos inmobiliarios.
- c) Identificar las propiedades presentadas por las inmobiliarias para la solicitud del permiso y encontrar las transacciones asociadas con ellas.
- d) Analizar cuán diverso es el escenario del precio del suelo a través del tiempo.
- e) Realizar un modelo de regresión lineal y analizar cuáles son los principales componentes que explican el precio del suelo entre el 2017 al 2021, en particular, enfocado en las distintas restricciones de edificación actualmente existentes en la comuna.
- f) Realizar una conclusión sobre las variables que más inciden en el precio de proyectos inmobiliarios.

1.5. ALCANCES

El alcance de esta memoria se vincula con analizar las variables elegidas que ayudan a explicar el precio del suelo. El propósito es obtener conclusiones importantes para evaluar la evolución del suelo de la comuna de Providencia entre el año 2017 y el 2021. Con los datos de estos límites es suficiente para realizar un análisis en profundidad.

Se busca crear un modelo que explique el precio del suelo de proyectos inmobiliarios, cuyo destino es comercializar productos inmobiliarios. Esto se efectúa con la identificación inicial de los permisos de edificación de los proyectos contenidos en la web del municipio, para luego identificar las transacciones en la página oficial del Conservador de bienes raíces. Además, se concluye y se responde a la pregunta: ¿qué es lo que influye para que el precio del suelo de un proyecto inmobiliario aumente?

1.6. METODOLOGÍA

Con el fin de alcanzar los objetivos generales y específicos, se utiliza la base de datos extraída de la página de transparencia de Providencia y se aplica una metodología de la siguiente forma:

- a) Levantamiento y procesamiento de datos
 - i) Ordenación de los datos, para luego clasificar los permisos en dos categorías: son o no proyectos.
 - ii) Obtención de los roles de cada permiso del proyecto.
 - iii) Obtención y ordenación de los datos de transacción de la página del conservador.
 - iv) Levantamiento de variables que caracterizan los terrenos, las restricciones normativas y el proyecto aprobado en cada permiso.

- b) Análisis de las variables
 - i) Definir las variables que más inciden en el incremento de la evolución del precio.
 - ii) Crear un modelo de regresión lineal múltiple que explique el precio del suelo con las variables más importantes.
 - iii) Interpretar el modelo realizado.

- c) Recomendaciones y conclusiones
 - i) Determinación y desarrollo de conclusiones.

2. MARCO TEÓRICO

En este capítulo, por una parte, se da a conocer investigaciones previas que se consideran para el análisis de los datos y las posteriores conclusiones del trabajo. Por otra parte, se abordan los conceptos claves necesarios para el desarrollo de este trabajo.

2.1. INVESTIGACIÓN PREVIA

El precio de las viviendas en Chile ha crecido en torno al 35% entre 2002 y 2007. Esto lo determina la antigüedad, la superficie construida y la superficie del terreno. Además del acceso a estaciones del Metro, la comuna, el ingreso promedio del vecindario y el ingreso real de los hogares. Entre el 68% y 71% de los precios, es explicado por los determinantes asociados con el acceso a bienes públicos, el entorno y los atributos físicos de la propiedad. Similarmente, el 68% del crecimiento de precios entre los años 1990 y 2007 son consecuencia de los factores macrofinancieros (Sagner, 2011).

Las inmobiliarias son las delegadas de crear y comercializar los proyectos hacia los compradores finales, además de efectuar las interacciones legales de un proyecto, la contratación de la empresa constructora y la supervisión de la construcción. Un concepto clave para estimar precios inmobiliarios es el cálculo del precio del suelo. En Chile este participa de la mano con la plusvalía, que es el aumento del valor de un bien por causas extrínsecas a él. Ambos se ven frecuentados del avalúo fiscal, el cual tiene que ver con la tasación con fines tributarios y constituido por el Servicio de impuestos Internos (SII), y el avalúo comercial, que es el valor de mercado. Por último, existe el avalúo del terreno que es el que emplea la inmobiliaria y que también esta denominado por el SII (Alejandro Cifuentes, 2020).

2.2. PRECIO DEL SUELO Y FACTORES QUE LO DETERMINAN

El precio del suelo constituye el “valor de cambio” de la renta de una localización, el equivalente monetario por el cual un propietario está dispuesto a abandonar los valores o las virtudes de un lugar en el territorio (Díaz, 2014). El valor del suelo se determina dependiendo de la actividad

económica para la que se usa, la intensidad de estas y los factores contextuales de su ubicación como: salud, infraestructura, seguridad, entre otros, que lo hacen más o menos atractivo (Sañudo, 2020).

La ciudad se conforma de suelo urbano con distintos usos y clasificaciones. Dentro de la ciudad, existen usos habitacionales, comerciales, servicios e industriales, cada uno de ellos se clasifica según su calidad. En este contexto, la calidad corresponde a las características inmobiliarias del suelo preferidas por los consumidores, por lo tanto, existen suelos de alta, media y baja preferencia (Quintana, 2018). La determinación del precio del suelo se basa en diferentes aspectos como: expectativas, accesibilidad, externalidades, productividad, valor de uso, entre otro (Wartenberg, 2018).

Los diferentes terrenos urbanos ofrecidos en el mercado suelen presentar variaciones de precio ocasionadas por varios factores. Algunos factores son la ubicación, la cual no cambia según la valoración de compra de todo inmueble. El uso del suelo, que está relacionada con el Plan Regulador Municipal vigente y las normas que dictamina. Otro factor es la forma y extensión de la superficie, las cuales pueden favorecer o perjudicar el precio de venta de un terreno. Y, por último, la Normativa donde se encuentran elementos importantes tales como el coeficiente de constructibilidad, coeficiente de ocupación de suelo, densidad, entre otros (portal inmobiliario, s.f.)

Como fue mencionado anteriormente, para calcular el precio de un terreno se necesita conocer el plan regulador comunal del sector, donde se localiza el terreno, para conocer las normas constructivas que se observan en esa zona. Parte de las normas son el uso de suelo que determina qué tipo de construcción está permitida en la zona, densidad, altura, vialidad y servicios, limite urbano y constructibilidad. A continuación, se hace el estudio de cabida del terreno el cual permite determinar cuántos departamentos se pueden construir allí. Luego se hace el estudio de mercado, se decretan los costos inmobiliarios y finalmente se da el valor final del terreno (Koporcic, 2019).

2.3. PROYECTO INMOBILIARIO

Un proyecto inmobiliario es el resultado tangible de un estudio financiero destinado a una obra de construcción sobre un bien raíz o inmueble. Recibe el nombre de proyecto porque incluye elementos que se deben considerar como, por ejemplo, el terreno, la ubicación, la factibilidad económica, estudio de mercado, tiempo de ejecución, entre otros (Maximilianos, 2021).

Con el propósito de gestionar un proyecto inmobiliario, se requiere evaluar varios factores que convergen e intervienen en el proceso de desarrollo del proyecto. Con este fin, el proyecto inmobiliario se divide en 6 etapas que permiten el control adecuado de los objetivos de estas empresas las cuales son: factibilidad del proyecto, diseño del proyecto, ingeniería, modalidades de venta, proceso de construcción y colocación en venta (Maximilianos, 2021).

2.4. NORMAS URBANÍSTICAS

Las Normas Urbanísticas son *“todas aquellas disposiciones de carácter técnico derivadas de la Ley General de Urbanismo y Construcciones de esta Ordenanza y del Instrumento de Planificación Territorial respectivo aplicables a subdivisiones, loteos y urbanizaciones o a una edificación tales como usos de suelo, sistemas de agrupamiento, coeficientes de constructibilidad, coeficientes de ocupación de suelo o de los pisos superiores, alturas máximas de edificación, adosamientos, distanciamientos, antejardines, y rasantes, densidades máximas, exigencias de estacionamientos, o cualquier otra norma de este mismo carácter.”* (Arévalo, 2018). En otras palabras, estas normas constituyen el documento donde se definen las condiciones generales y particulares de las diferentes clases y categorías de suelo.

Dentro de estas normas se puede encontrar el coeficiente de constructibilidad, que nos dice cuanto es lo máximo que se puede construir en el terreno, el coeficiente de ocupación de suelo, el cual fija el máximo de metros cuadrados posibles de construir en el nivel del primer piso. También se encuentra el coeficiente de ocupación de los pisos superiores, que fija el máximo de superficie edificada que se pueda construir en cada uno de los pisos superiores al primero (Serrano, 2018).

Otras normas que se pueden encontrar son la altura de edificación que puede ser expresada en metros o número de pisos, el adosamiento, distanciamiento, antejardín, densidades y rasantes, el cual es “*recta imaginaria que, mediante un determinado ángulo de inclinación, define a envolvente dentro de la cual puede desarrollarse un proyecto de edificación*” (Serrano, 2018).

2.5. REGRESIÓN MÚLTIPLE

El modelo de regresión múltiple permite estimar el efecto sobre la variable dependiente Y_i de la variación de una variable (X_{1i}) manteniendo constantes el resto de las variables independientes ($X_{2i}, X_{3i}, etc.$).

El modelo de regresión múltiple se expresa de la siguiente forma:

$$Y_i = \beta_0 + \beta_1 * X_{1i} + \beta_2 * X_{2i} + \dots + \beta_k * X_{ki} + u_i, i = 1, \dots, n, \quad (1)$$

donde Y_i es la i-ésima observación de la variable dependiente; $X_{1i}, X_{2i}, \dots, X_{ki}$ son las i-ésimas observaciones de cada uno de los k variables independientes; y u_i es conocido como error.

Habitualmente se utilizan tres estadísticos en la regresión múltiple que corresponden al error estándar de la regresión, R^2 ajustado y R^2 de la regresión. Estos estadísticos nos dicen en qué medida la recta de regresión múltiple describe o se ajusta a los datos (Watson, 2012).

2.6. R-CUADRADO, T-VALUE Y P-VALUE

2.6.1. R-cuadrado y R-cuadrado ajustado

El R^2 , también conocido como coeficiente de determinación, es una medida estadística que mide lo bien que un modelo de regresión se ajusta a los datos reales. Es la proporción de la variación total de la variable dependiente que es interpretada por su relación con la variable independiente (Franco, 2017).

$$R^2 = \frac{SSR}{SST} \quad (2)$$

Donde:

SSR: suma de cuadrados explicada por la regresión.

SST: suma total de cuadrados.

El coeficiente de determinación varía entre 0 y 1, cuanto más cerca de 1 este su valor, mayor será el ajuste del modelo a los datos. En cambio, mientras más cerca de 0, menos ajustado estará el modelo a los datos, por tanto, será menos fiable (López, 2023).

Al añadir una variable al modelo, el R^2 aumenta, pero esto no significa que la adición de una variable mejore realmente el ajuste del modelo. En este aspecto, el R^2 proporciona una estimación excesiva acerca de la bondad con la que la regresión se ajusta a los datos. Una forma de corregirlo es reducir el R^2 y esto es lo que hace el R^2 ajustado (Watson, 2012).

El R^2 ajustado es una versión modificada de R^2 que no necesariamente acrecienta al añadirse una nueva variable. Se representa de la siguiente forma:

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1} \frac{SR}{ST} \quad (3)$$

Esta ecuación nos indica que el R^2 ajustado es 1 menos el cociente entre la varianza muestral de los residuos MCO y la varianza muestral de Y (Watson, 2012).

2.6.2. T-Value

El valor t es el valor que toma el estadístico de prueba. Mide el tamaño de la diferencia en relación con la variación en los datos de la muestra. En otras palabras, T es la diferencia calculada simplemente representada por la unidad estándar de error. Cuanto mayor sea el tamaño de t, más grande será la evidencia en contra de la hipótesis nula. Esto significa que hay mayor certeza de

que existe una discrepancia significativa. Mientras más cerca este T del 0, mayor será la probabilidad que no exista una diferencia significativa (Minitab, 2019).

Se denominan pruebas t, ya que todos los resultados de las pruebas se basan en valores t. Estos son un ejemplo de estadísticas de prueba, que es un valor estandarizado calculado a partir de datos de muestra durante una prueba de hipótesis. El método que calcula la estadística de prueba compara sus datos con lo que se espera bajo la hipótesis nula. Cada tipo de prueba t emplea un procedimiento específico para disminuir todos los datos de muestra a un valor, el valor t. Los cálculos subyacentes a los valores t comparan las medias de la muestra con la hipótesis nula e integran el tamaño de la muestra y la variabilidad de los datos. Un valor t de 0 indica que los resultados de la muestra son iguales a la hipótesis nula y a medida que aumenta la diferencia entre los datos de la muestra y la hipótesis nula, incrementa el valor absoluto del valor t (Editor de blogs de Minitab, 2016).

2.6.3. P-value

El *p-value* se utiliza en la estadística, desde las pruebas t hasta el análisis de regresión. Es la probabilidad de conseguir los resultados observados de una prueba, considerando que la hipótesis nula es verdadera. Es el nivel de significancia marginal dentro de una prueba de hipótesis estadística que representa la posibilidad de que ocurra un evento particular. El valor p se usa como una opción a los puntos de rechazo para proveer el nivel de significancia más pequeño al que se rechazaría la hipótesis nula. Un valor p menor significa que hay pruebas más fuertes a favor de la hipótesis alternativa (Data Science Team, 2020).

El valor p evalúa la hipótesis nula que significa que el coeficiente es igual a cero, no hay efecto. Un valor p bajo, normalmente menor a 0,05, quiere decir que se rechaza la hipótesis nula. Dicho de otra manera, es probable que un predictor que tenga un valor p pequeño sea una adición significativa al modelo debido a que los cambios en el valor del predictor se relacionan con cambios en la variable dependiente. Al contrario, un valor p mayor a 0,05 (insignificante) aconseja que los cambios en el predictor no están asociados con cambios en la variable dependiente (Minitab, 2019).

2.7. CLASSIFICATION AND REGRESSION TRAINING (CARET) (explicar el método)

Classification And **RE**gression Training (CARET) es un paquete de R que proporciona herramientas para informar sobre la relevancia y la importancia de las variables, además de seleccionar las características más importantes para el modelo (Brownlee, 2019).

El paquete CARET intenta acelerar el proceso de creación de modelos predictivos. Contiene herramientas para división de datos, preprocesamiento, selección de características, ajuste del modelo mediante remuestreo y estimación de importancia de las variables (Kuhn, 2019).

VER SI DEFINO IMPORTANCIA VARIABLE

2.8. TEST DE SHAPIRO-WILKS Y TEST DE BREUSCH-PAGAN test de kolmogorov-smirnov

2.8.1. Prueba de Shapiro

La prueba de Shapiro-Wilks aborda la hipótesis nula que dice que una muestra proviene de una distribución normal y la hipótesis alternativa sostiene que la distribución no es normal. Esto con un nivel de significancia generalmente de 0,05 (Dietrichson, 2019).

2.8.2. Prueba de Breusch-Pagan

La regresión lineal cuenta con un supuesto clave que dice que los residuos se distribuyen con la misma varianza en cada nivel de la variable independiente. A esta suposición se le conoce como homocedasticidad. Cuando se infringe este supuesto, se dice que la heterocedasticidad está presente en los residuos. En el momento en que esto ocurre, los resultados de la regresión se vuelven poco confiables.

La prueba de Breusch-Pagan se usa para determinar si la heterocedasticidad se encuentra presente en un modelo de regresión o no. Esta prueba utiliza las siguientes hipótesis: la hipótesis nula dice si la homocedasticidad está presente y la alternativa dice que la heterocedasticidad está presente, los residuos no se distribuyen con la misma varianza como ocurre con la homocedasticidad. Si el valor p es más bajo que algún nivel de significancia, normalmente de 0,05, se rechaza la hipótesis

nula y se concluye que la heterocedasticidad está presente en el modelo de regresión (Statologos, 2021).

2.9. STEPWISE Y AKAIKE (AIC)

Stepwise es un método para seleccionar variables. Es la construcción iterativa paso a paso de un modelo de regresión que involucra la selección automática de variables independientes. El objetivo de esto es que, a través de una serie de pruebas (prueba F y prueba t) descubrir un conjunto de variables independientes que impacten significativamente en la variable dependiente (Gonzalez, 2019).

Dentro de este método se encuentran 3 estrategias (Rodrigo, 2016):

- Forward stepwise: el modelo inicial no incluye ningún predictor, solo el parámetro β_0 . A partir de este se producen todos los posibles modelos introduciendo una sola variable de entre las disponibles. La variable que en mayor medida mejore el modelo es la que se selecciona.
- Backward stepwise: el modelo se comienza con todas las variables disponibles contenidas como predictores. Se prueba eliminar una a una cada variable, si se mejora el modelo, queda descartada.
- Stepwise mixto: es la combinación de los dos procedimientos anteriores. Presenta la ventaja de que, si a medida que se agregan predictores, alguno de los ya presentes deja de aportar al modelo, se elimina.

El método stepwise necesita de algún criterio matemático para establecer si el modelo mejora o agrava con cada incorporación o extracción. Uno de ellos es el criterio de información de Akaike (AIC) que evalúa que tan bien se ajusta un modelo a los datos a partir de los cuales se generó. Se usa para comparar distintos modelos posibles y decretar cual se ajusta mejor a los datos (Bevans, 2022).

3. LEVANTAMIENTO DE DATOS Y GENERACIÓN DE BASE DE DATOS

En este capítulo, se presenta el levantamiento y la generación de los datos.

3.1. LEVANTAMIENTO DE DATOS

Como se mencionó en la introducción, el objetivo general de esta memoria es evaluar, con herramientas econométricas, un modelo para entender el valor del suelo entre los años 2017 y 2021 para proyectos inmobiliarios ubicados en la comuna de Providencia. Con el propósito de elaborar este modelo, se ejecutó una serie de levantamiento y procesamiento de datos.

Uno de los requisitos para realizar cualquier tipo de análisis es tener claridad de los datos con los que se trabajará y eliminar cualquiera que no aporte o que no se necesite para cumplir con lo solicitado.

3.1.1. Levantamiento de los permisos

Con la finalidad de obtener la información requerida, se levantaron los datos generales de los permisos de obras nuevas desde 2017 hasta el 2021 de la comuna de Providencia, desde la página del Portal de Transparencia de Providencia. Estos datos son los siguientes: número del permiso, fecha de aprobación y solicitud del permiso, dirección, superficie edificada en m^2 , descripción del acto (vivienda, oficina, comercio, entre otros), también se extrajo el número de pisos, nombre del propietario y el permiso de edificación que se utiliza más adelante.

Al ingresar estos datos a un Excel y ordenarlos, se obtuvo una cantidad de 165 permisos de obras nuevas, de las cuales hay que evaluar cuáles corresponden a proyectos inmobiliarios. Con este fin, se evaluó la superficie edificada en m^2 y quién es el propietario del terreno. Este último puede ser una inmobiliaria, una empresa o una persona, pero en este caso, solo se limitó a inmobiliarias y empresas. Así mismo, se evaluó si el terreno corresponde a una ampliación/modificación o si se está edificando algo nuevo. En el primer caso, la información no es de utilidad para la base de datos debido a que una ampliación/modificación de algún terreno no es significativo para este

trabajo. Otro factor para evaluar si el permiso corresponde a un proyecto inmobiliario es la descripción del acto y el número de pisos del permiso en conjunto. En este caso particular, se incluyó las construcciones de 3 o más pisos de altura, correspondientes a viviendas, oficinas, locales comerciales y hoteles.

Con lo mencionado anteriormente, se determinó que 136 permisos corresponden a un proyecto inmobiliario, es decir, un 82,42% del total.

Tabla 1: Cantidad de proyectos inmobiliarios por año

Año	Cantidad de Permisos de Obras Nuevas	Cantidad de proyectos inmobiliarios	% correspondiente a proyectos inmobiliarios
2017	26	20	76,92
2018	33	28	84,84
2019	45	36	80
2020	28	23	82,14
2021	33	29	87,87
Total	165	136	82,42

Fuente: Elaboración propia

En la Tabla 1, se muestra la cantidad de permisos de obras nuevas por año de la comuna de Providencia y cuáles de esos permisos corresponden a proyectos inmobiliarios, tanto en cantidad como en porcentaje. Esto permite proyectar el volumen total del trabajo realizado.

3.1.2. Roles

Luego de evaluar los permisos, se identificó el rol de avalúo de cada proyecto inmobiliario, que permite identificar una propiedad o bien raíz. Este se obtuvo del permiso de edificación, extraído de los permisos de obras nuevas. Este identificador es único a nivel comunal que, en el caso de la comuna de Providencia, corresponde a 15103.

El número de roles se compone de dos partes: código manzana y de predio, que se separan con un guion. Cada rol puede tener más de un código predio, lo que significa que se compraron anteriormente varias propiedades para construir el proyecto correspondiente. A raíz de esto, un permiso puede tener más de un rol de avalúo, lo que deriva en un aumento significativo de la base de datos, en comparación con la cantidad de permisos.

Entre los años 2017 y 2021 se levantó un total de 540 roles de avalúo asociados con los permisos de los proyectos inmobiliarios. En la Tabla 2, se presenta la cantidad de roles levantados por año, y el aumento significativo del año 2019 se produce porque uno de los permisos contiene la mayor cantidad de roles de avalúo: 72.

Tabla 2: Cantidad de rol por año

Año	Cantidad de rol de avalúo
2017	65
2018	85
2019	184
2020	98
2021	108
Total	540

Fuente: Elaboración propia

Luego de incorporar los roles a un archivo Excel, se obtuvo la foja, número y año de cada uno. De acuerdo con la Ley, esta es la manera en que se singulariza una inscripción determinada en un Registro del Conservador. Lo mencionado anteriormente, se realizó con el fin de extraer las transacciones de los terrenos, ya que la combinación del número de manzana y predio permite obtener esta información.

3.1.3. Transacción

Con el fin de construir el índice del precio del suelo, se requiere conocer la transacción de cada terreno. Esto se realiza con el fin de evaluar la evolución del precio a lo largo de estos 5 años, es decir, entre el 2017 y 2021.

Una vez identificados los roles, se consultó la información que el Servicio de Impuestos Internos tenía con respecto a estas inscripciones, que son antiguas y, por lo tanto, no todas están disponibles. Debido a esto, hay 2 caminos alternativos para encontrar las inscripciones. La primera de ellas es obtener la inscripción a través de la página del Conservador, específicamente por medio de consultas en línea, en el inciso de Índice del Registro de Propiedad, que contiene los documentos de las transacciones de los terrenos.

Con el fin de identificar una transacción, se necesita su foja, número y año. Esta *“es la manera que, de acuerdo con la Ley, se singulariza una inscripción determinada en un Registro del Conservador”* (Conservador de Molina, 2022). Es decir, estas tres variables permiten identificar el documento de la compra del terreno.

Existen 2 maneras identificar este documento en la página del Conservador: la primera es buscar la foja, número y año del rol que se está buscando, que debe coincidir con el propietario, además de la manzana y predio del rol. La segunda forma es con el motor de búsqueda por medio del nombre del propietario, por año de solicitud del permiso y comuna. Esta forma tiene un grado mayor de complicación, porque se debe ingresar en cada documento hasta encontrar el rol que se está buscando.

Si al usar estas dos formas no se obtiene un resultado, la segunda alternativa es buscar una primera venta de los departamentos asociados con ese proyecto, para que luego de encontrarlo, retornar a la inscripción previa correspondiente al terreno. Esta opción es compleja, ya que para identificar un departamento asociado con ese proyecto se necesita la dirección. Por este motivo, esta última se busca en la página de Servicio de Impuestos Internos, específicamente en el inciso Mapas digitales. A continuación, con la dirección identificada, se accede a la página del Conservador

Digital, para obtener la foja, número y año de ese departamento y, posteriormente, buscarla en la página del Conservador. Al acceder al documento de este departamento, en el extremo inferior se encuentra la foja, el número y el año del terreno de la inscripción anterior. Estas tres variables tienen que coincidir con el nombre del propietario, además de su rol.

Del documento obtenido, se extrae el nombre y RUT del comprador que debe coincidir con el del propietario, además del nombre del vendedor de la propiedad, la fecha de aprobación y de escritura. También se extrae el tipo de transacción, que puede ser de tipo compraventa, aporte, fusión, disolución, entre otros, y el monto de la propiedad, que puede estar en UF, USD o pesos chilenos. En este trabajo de investigación, se requiere el precio en UF, por lo tanto, se realizó la conversión con respecto a la UF correspondiente a la fecha de aprobación del permiso.

Uno de los obstáculos interpuesto en el levantamiento de las transacciones es el número de consultas diarias para acceder a la página del Conservador: 15 consultas diarias. Esto generó un problema, ya que el tiempo dedicado a obtener las transacciones de los roles obtenidos de los permisos se extendió más de lo planificado.

Los datos extraídos de la página del Conservador se ordenaron en un Excel, tal como se muestra en el Anexo 1. Al final del levantamiento de estos datos, comprendidos entre los años 2017 y 2021, se extrajeron 616 transacciones en total.

Tabla 3: Cantidad de transacciones por año

Año	Cantidad de transacciones
2017	71
2018	96
2019	206
2020	110
2021	133
Total	616

Fuente: Elaboración propia

En la Tabla 3 se muestra la cantidad de transacciones levantadas por año. En ella se evidencia un aumento en la base de datos, que se origina porque un rol puede tener más de una transacción. El 2019 es el periodo en el cual se levantaron más transacciones debido a que es el año que contiene mayor número de roles.

3.2. GENERACIÓN DE LA BASE DE DATOS

La base de datos se generó a partir de las variables obtenidas de la transacción y del levantamiento de datos del permiso de edificación.

3.2.1. Permisos de edificación

Después de recopilar los datos levantados de la transacción, se obtuvieron más datos sobre la edificación del terreno obtenidos del permiso de edificación. Estos datos serán usados en una regresión lineal para determinar qué variables son significativas para este modelo con respecto a la variable dependiente, que en este caso es el precio unitario del terreno en UF/m².

De cada permiso, se extrajo el Número de Edificación, que corresponde a la cantidad de edificios a construir en el terreno. Además, la superficie edificada en m², la cantidad de pisos, que corresponde a los pisos en altura que tendrá el edificio, la zona de edificación y de uso. También se extrajo el nombre del propietario, su RUT, y si este corresponde a jurídica o natural, en otras palabras, si es una empresa o una persona respectivamente.

Otro dato que se obtiene del permiso de edificación es si el destino de la propiedad es para utilización residencial, equipamiento, actividades productivas o infraestructura. En el primer caso, se debe identificar si es vivienda u hotel. En el segundo, se tiene que determinar si el uso será comercial, oficina o ambos. Otros datos levantados corresponden a las siguientes superficies: edificación útil en m² transformada a porcentaje, edificación común en m² y el total del terreno en m².

Con respecto a las normas urbanísticas aplicadas que se encuentran en el permiso, se extrajo el dato de coeficiente de constructibilidad, el cual fija el máximo de metros cuadrados posibles de construir sobre el terreno, cuanto se puede construir en total, incluyendo todos los pisos. También se obtuvo la altura máxima en pisos, la cual representa la cantidad de pisos máximos que tendrá el edificio, los rasantes medidos en grados y el coeficiente de ocupación de suelo, que es cuánto de la superficie del terreno puede ocupar el primer piso. Además de recabar la densidad, el adosamiento medido en porcentaje, que es la edificación que se ubica contigua a la determinación de los límites de un terreno en su parte no ocupada por construcciones, el antejardín y los números de estacionamientos requeridos.

También se obtuvo información con respecto a las disposiciones especiales a las cuales se acoge el proyecto, es decir, si contiene el D.F.L.N.º2 (Decreto con Fuerza de Ley 2), el conjunto armónico, la Copropiedad Inmobiliaria, Beneficio de fusión, Proyecto Sombras, Conjunto Vivienda económica o si el permiso cuenta con otras disposiciones especiales a las que se acoge el proyecto que no fue mencionada anteriormente. Además, se evalúa si el proyecto cuenta con anteproyecto aprobado o no. Por último, se extrajo el número de unidades totales de viviendas, locales comerciales, oficinas, estacionamientos, educación, si es que corresponden.

Además de obtener los datos del permiso de edificación, a la base de datos se le agregó la cantidad de roles de cada permiso, es decir, la cantidad de transacciones por permiso, el año correspondiente a la transacción y el plazo primera compra, que son los años que han transcurrido entre la primera y última compra del terreno correspondiente al permiso. Finalmente, se agregó el distrito censal y las zonas censales.

Cada comuna se divide en varios distritos censales, en el caso de la comuna de Providencia corresponden a 8 distritos. Cada distrito se divide en varias zonas, en este caso son 16. Para identificar exactamente donde está ubicado el proyecto se creó una variable de localización llamada zona censal, que corresponde a la unión del distrito censal y la zona, esta es importante a la hora de ver el precio de un terreno. Estos datos se pueden observar en el Anexo 2.

3.2.2. Análisis descriptivo de los precios unitarios

Con el fin de obtener el precio unitario en metros cuadrados de cada permiso, se calculó la suma de los precios en UF de cada rol obtenido del documento de inscripción del terreno perteneciente al permiso de edificación. Posteriormente, se dividió con respecto a la superficie total del terreno en metros cuadrados perteneciente al mismo permiso.

Con el propósito de determinar esta variable, se calculó el promedio del precio unitario de cada año. Este consiste en la división de la suma de los precios de cada permiso por la cantidad de estos por año.

Otro concepto para analizar es la desviación estándar que mide la dispersión de una distribución de datos. En la medida en que exista una mayor dispersión de los datos, mayor será su desviación estándar. En este caso se calculó la desviación estándar muestral.

Por último, se analizó el rango de los precios. Este es un valor numérico que indica la diferencia entre el valor máximo y el valor mínimo de la muestra, en este caso, de la muestra de los precios unitarios de cada año.

Lo descrito anteriormente se puede observar en la Tabla 4 presentada a continuación:

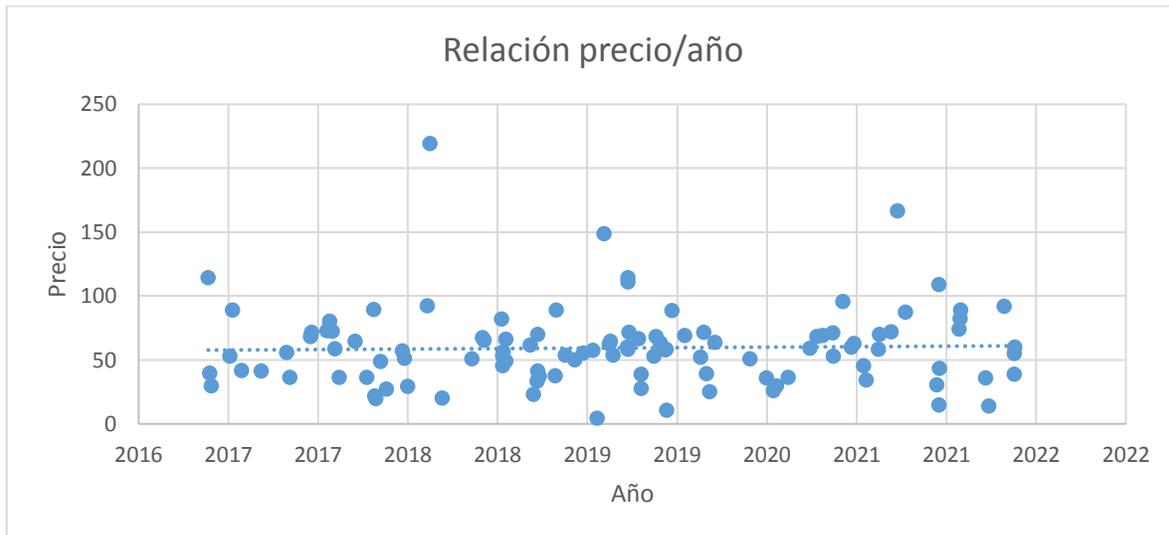
Tabla 4: Análisis precio unitario por año

Año	Promedio	Desviación estándar	Valor Máximo	Valor Mínimo	Rango
2017	60,3238965	22,2696908	114,240401	29,6120817	84,6283189
2018	60,5775556	41,6414707	219,125592	19,4890088	199,636583
2019	59,3792194	28,6705439	148,869774	4,24649	144,623284
2020	53,8226967	18,9506316	95,7611958	25,0273224	70,7338734
2021	70,5478808	41,6572609	166,59313	13,6956738	152,897456

Fuente: Elaboración propia

Para un análisis más profundo se realizó un gráfico de dispersión, que mide la intensidad de la relación entre dos variables: el año del permiso y el precio unitario.

Gráfico 1: Relación precio/Año

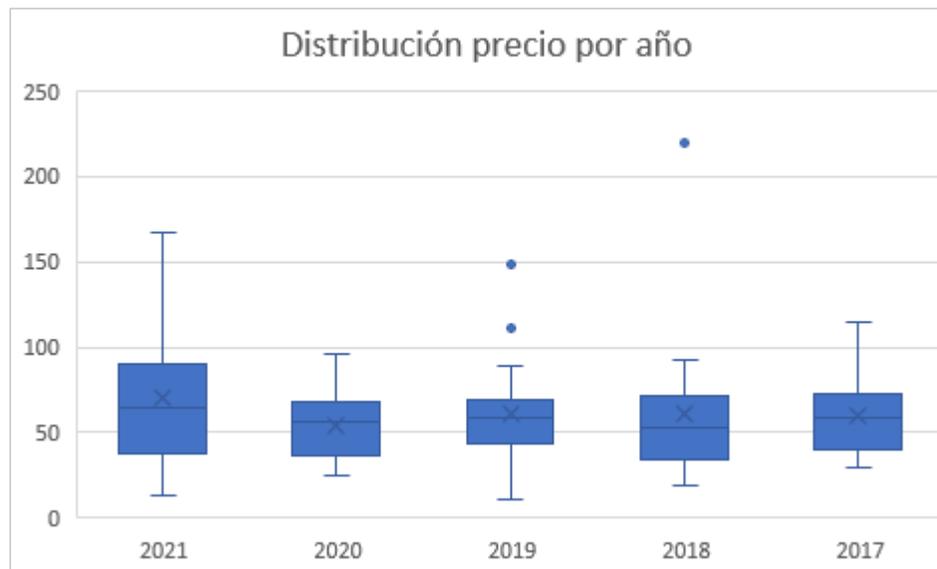


Fuente: Elaboración propia

En el Gráfico 1 el eje X representa el año de aprobación del permiso y el eje Y representa el precio unitario en UF/m² de cada permiso. Se puede apreciar que el precio se centra en los 50 UF y que el precio más alto se encuentra en el año 2018 y el más bajo en el año 2019. Además, se puede apreciar que el precio no varía mucho en el transcurso de los años, pero si hay un leve aumento entre principios del 2017 hasta finales del año 2021, como muestra la línea de tendencia.

También se realizó un gráfico de cajas y bigotes, el cual permite observar y comparar la distribución y tendencia central de los precios durante cada año mediante sus cuartiles.

Gráfico 2: Gráfico de cajas y bigotes precio por año



Fuente: Elaboración propia

Este Gráfico 2 nos demuestra que en el año 2018 la mediana se encuentra en el centro, por lo que los datos no están sesgados en ese año, la distribución es simétrica. En cambio, en los datos de los otros 4 *boxplot* la mediana no se encuentra en el centro, por lo que los datos están sesgados, nos dice que la distribución es asimétrica. El año 2017, 2019 y 2020 representa una simetría negativa debido a que la parte más larga de la caja se encuentra en la parte inferior de la mediana, en cambio, el año 2021 demuestra una asimetría positiva, ya que la parte más larga se encuentra en la parte superior de la mediana.

En los años 2018 y 2019, se encuentran valores atípicos, estos son valores que se encuentran fuera del límite superior y el límite inferior.

4. ANALISIS DE DATOS

Como se mencionó anteriormente, la cantidad de permisos que corresponden a proyectos inmobiliarios son de 136. De esa cantidad solo se usarán 111 permisos, esto se debe a que en los demás permisos no se pudo encontrar el monto de alguna transacción correspondiente a ese permiso o la transacción corresponde a fusión, aporte o disolución.

Al tener la base de datos ya definida, se confirma que hay potenciales riesgos de dimensionalidad debido a la relación entre la cantidad de registros (observaciones) y la cantidad de variables. A raíz de esto, se realizó un trabajo de reducción de dimensionalidad en donde, sin haber corrido el modelo, se decidió que variables eran más importantes para el modelo, según la literatura y la revisión bibliográfica, además de decidir que variables se pueden reducir.

A continuación, se hicieron análisis econométricos preliminares con diferentes modelos como el *random forest* y regresión lineal, el cual ayudó a confirmar que variables excluir de la especificación final del modelo de regresión.

4.1. ANÁLISIS DE COMPORTAMIENTO NORMAL

Con la base de datos ya establecida, se comenzó a realizar el análisis de comportamiento normal de la variable dependiente y de las independientes para obtener una buena regresión lineal. Esto consiste en tratar las variables para validar cuál es la mejor forma de ajuste entre la variable dependiente e independiente, si es mejor dejar la original, transformarla en logaritmo o en cuadrática. El tratamiento de estas no valida si sirve o no, sino que valida la mejor forma para que las variables independientes se ajusten mejor con la dependiente.

De cada variable se consideró el valor original, junto con las transformaciones cuadráticas y logarítmicas. Esta última se usa a menudo para variables sesgadas a la derecha, donde la media es mayor que la mediana, ya que hace que la distribución de las variables transformadas parezca más simétrica.

Se comenzó evaluando la variable dependiente, que en este caso es el precio unitario del terreno en UF/m² llamado en la base de datos valor unitario. Que dicha variable sea dependiente significa que todas las conclusiones serán a base de esta.

A esta variable se le creó un nuevo atributo y se le asignó el valor de transformación, tal como se muestra en la Ilustración 1.

valor_unitario	valor_unitario_log	valor_unitario_quad
<dbl>	<dbl>	<dbl>
59.8	4.09	3570.
92.0	4.52	8458.
87.5	4.47	7648.
82.4	4.41	6792.
38.7	3.66	1499.
13.7	2.62	188.

Ilustración 1: Valor de las transformaciones

Fuente: Elaboración propia

A continuación, se examinó un resumen estadístico del valor unitario y sus transformaciones.

valor_unitario	valor_unitario_log	valor_unitario_quad
Min. : 10.75	Min. : 2.375	Min. : 115.5
1st Qu.: 39.40	1st Qu.: 3.674	1st Qu.: 1552.7
Median : 58.12	Median : 4.062	Median : 3377.5
Mean : 61.57	Mean : 3.998	Mean : 4796.9
3rd Qu.: 71.68	3rd Qu.: 4.272	3rd Qu.: 5138.1
Max. : 219.13	Max. : 5.390	Max. : 48016.0

Ilustración 2: Resumen estadístico

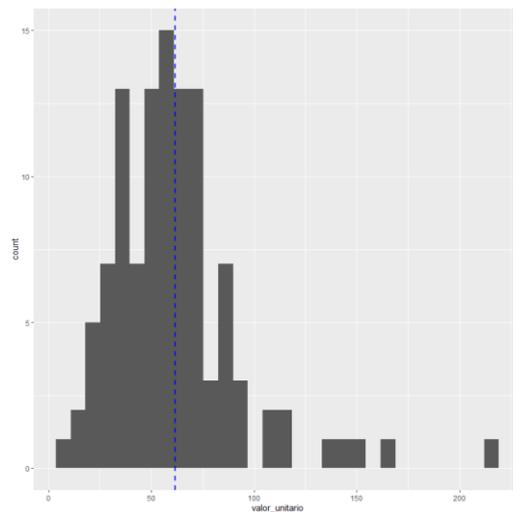
Fuente: Elaboración propia

La Ilustración 2 nos muestra el valor mínimo y máximo, el primer y tercer cuartil, además de la media y mediana tanto del valor original como de sus transformaciones. Se observa que la variable transformada a logaritmo corrige la distribución sesgada de la no transformada, ya que los valores mínimo y máximo tienen una mediana de 4,062 que se acerca a la mitad de los dos extremos, en

cambio, el valor original tiene una mediana de 58,12 la cual está sesgada hacia la derecha, y la cuadrática tiene una mediana de 3.377 que está aún más sesgada hacia la derecha.

A continuación, se realizó un histograma para observar los datos de la variable sin transformar y ver la tendencia de esta.

Gráfico 3: Histograma del valor unitario

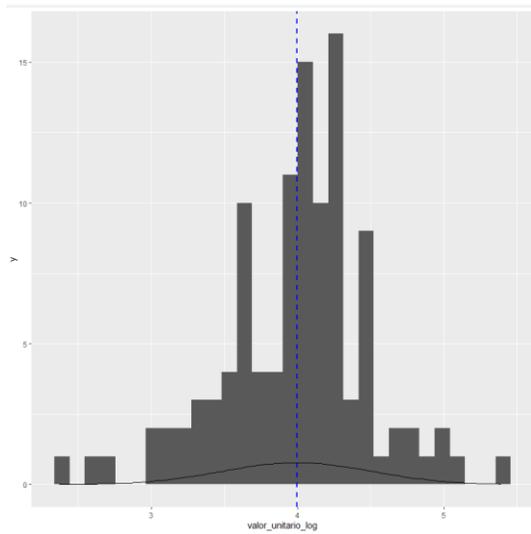


Fuente: Elaboración propia

El Gráfico 3 nos muestra que la variable original está sesgada hacia la derecha con la mayoría de los valores cerca del valor mínimo, junto con una cola que se extiende hasta el valor máximo, no hay simetría.

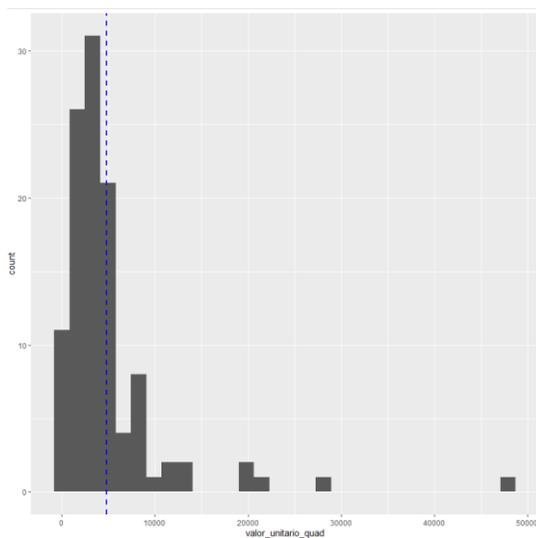
A raíz de lo mencionado anteriormente, se realizó un histograma para la variable transformada a logarítmico y para la transformada a cuadrática. Esto con el fin de ver si tiene una tendencia a distribución normal.

Gráfico 4: Histograma valor unitario transformada a logaritmo



Fuente: Elaboración propia

Gráfico 5: Histograma valor unitario transformada a cuadrática



Fuente: Elaboración propia

Como se observa en los histogramas, el Gráfico 4 con la variable unitaria transformada a logaritmo proporciona una distribución que se puede considerar normal y, como tal, esta variable transformada se puede utilizar para la regresión lineal. En cambio, el Gráfico 5 con la variable

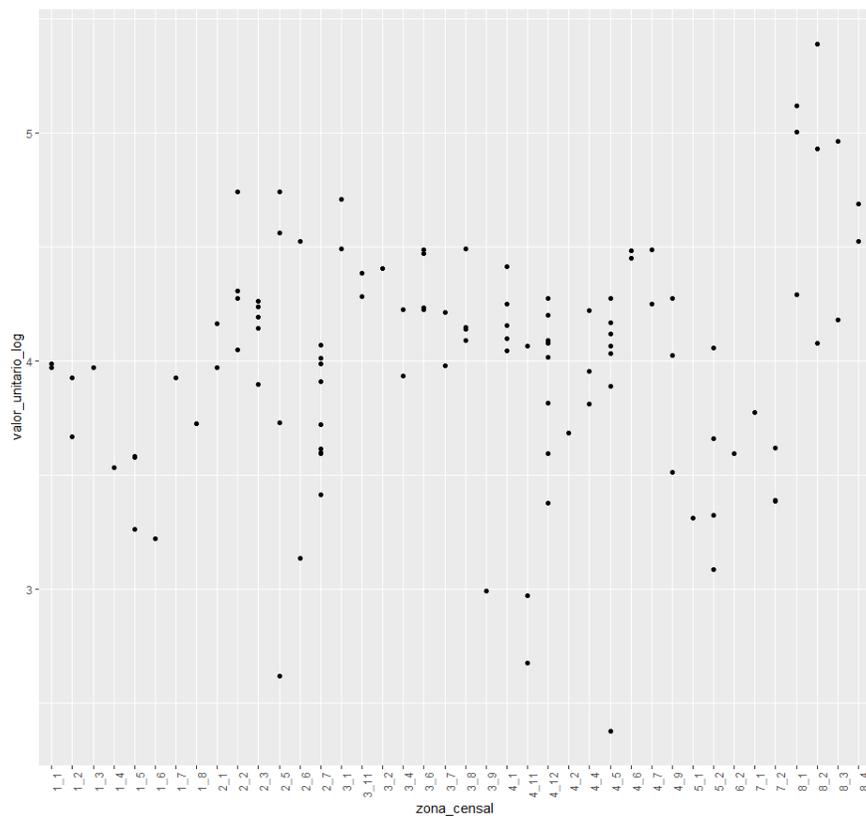
transformada a cuadrática demuestra que los datos están cerca del valor mínimo, es decir, están estrechamente sesgadas hacia la derecha.

Tras lo descrito anteriormente, se llegó a la conclusión de que la variable dependiente que mejor se ajusta al modelo con una distribución normal es la transformada a logaritmo, por lo cual esta es la que se usará para el modelo de regresión lineal.

Luego de evaluar la variable dependiente para definir cuál es la transformación que mejor ajusta al modelo, se hizo la transformación a cuadrática y logarítmica de las independientes que pueden ser tratadas, tales como: la superficie total del terreno, la edificada y la útil en porcentaje, el año de transacción, el número de roles, la altura máxima, el coeficiente de constructibilidad y el coeficiente de ocupación de suelo.

Para las variables categóricas, las transformaciones no se pueden completar. Sin embargo, se puede emplear un gráfico de dos vías para examinar una variable independiente, en este caso la zona censal, frente al resultado, el valor unitario en logaritmo.

Gráfico 6: Variable categórica



Fuente: Elaboración propia

El Gráfico 6 muestra el número de observaciones y la distribución del precio del terreno en m² transformado a logaritmo para cada clase de categoría en la zona censal de variables categóricas. Por ejemplo, la zona censal 2_7 parece contener varias observaciones, pero la zona censal 1_3 solo contiene una observación.

4.2. PAQUETE CARET

Al terminar de transformar las variables se utilizó el paquete Caret de R (**C**lassification **A**nd **R**egression **T**raining). Este paquete proporciona herramientas para informar sobre la relevancia e importancia de las variables, además de seleccionar las características más importantes para el modelo. En este caso se empleó para estimar la importancia de las variables a partir de los datos mediante la construcción de un modelo.

Con el fin de evaluar la importancia de las variables independientes con respecto a la dependiente en logaritmo, se usó la variable original y las variables tratadas a logaritmo y a cuadrática. Este paquete mediante *varImp*, que se usa para estimar la importancia de la variable, confirma cuáles son las variables tratadas y no tratadas de mayor importancia. Los representa a través de un gráfico que nos muestra qué variables dejar, además de decirnos que variables no son de importancia para el modelo, tal como se muestra en el Anexo 3. Este proceso es más racional y objetivo, se enfoca en todas las variables a la vez. Además, verifica la correlación de las variables.

El Anexo 3 nos muestra que las variables de mayor importancia para este modelo son el plazo entre la primera y última compra, el coeficiente de constructibilidad original y tratada, la altura máxima original y tratada, el antejardín, la cantidad de viviendas, la superficie útil en porcentaje original y tratada, la densidad, si el proyecto se acoge al D.F.L.N.°2, la superficie edificada en m² original y tratada, la cantidad de estacionamientos requeridos por permiso, la superficie total del terreno original y tratada, el número de roles original y tratada, la cantidad de locales comerciales, el año de transacción original y tratada, el número de oficinas, el coeficiente de ocupación de suelo original y tratada y el beneficio de fusión.

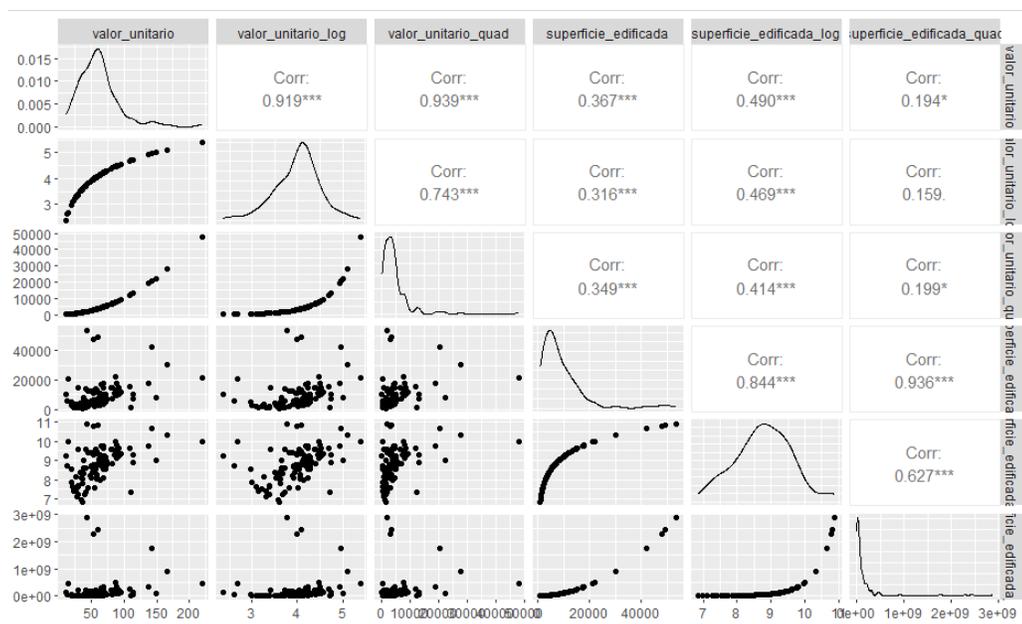
Debido a que hay tantos valores unitarios de cada zona y no todas las zonas censales son relevantes para el modelo, se crearon *dummies* con las más relevantes, en este caso corresponden a la zona censal 2_7, la zona censal 4_11, zona censal 2_5, zona censal 5_2 y la zona censal 3_6. Esto se realizó después de evaluar la importancia de las variables incluidas las zonas a través del paquete *Caret*. Al efectuar esto, se elimina la variable categórica zona censal, ya que generaba problemas de dimensionalidad y la mayoría de ellas no aportaban al modelo. Esto ayuda a que mejore la significancia de este.

4.3. MATRIZ DE AUTOCORRELACIÓN

Con el fin de verificar si las variables elegidas que fueron tratadas tienen correlación con la variable dependiente, se realizó una matriz de correlación. Esta matriz muestra la distribución de frecuencia de cada variable (visualizada como un gráfico de líneas), su diagrama de dispersión en relación con otras variables y su correlación con otras variables (representada por el valor *corr*).

Esto se muestra para las variables independientes y dependiente sin transformar y transformadas. En este caso se comprobó con la variable independiente superficie de edificación transformada a logaritmo, a cuadrática y no transformada, y la dependiente transformada a logaritmo, a cuadrática y en su valor original.

Gráfico 7: Matriz de correlación



Fuente: Elaboración propia

Como era de esperarse, el Gráfico 7 nos muestra que existe una alta correlación entre las variables originales y las transformadas en logaritmo y cuadrática. Por ejemplo, el valor unitario tiene 0,919 de correlación con su variable transformada a logaritmo, esto debido a que los valores de uno se derivan del otro.

El objetivo de esta matriz es examinar los gráficos entre la variable valor unitario transformada a logaritmo y la variable independiente, superficie edificada, en su valor original y transformada. Como demuestra el Gráfico 7, la variable original y transformada a logaritmo tienen una alta correlación con la variable dependiente en logaritmo, esto nos dice que la superficie edificada explica el precio unitario del terreno en UF/m², en cambio, con la superficie edificada transformada a cuadrática existe una baja correlación, aunque este igual explica, en parte, el precio

unitario. Como resultado, los valores de correlación se utilizan simplemente, en este caso, para comparar las relaciones entre diferentes formas funcionales de cada variable.

Se observa el tratamiento que mejor ajuste para el modelo. En este caso, como se dijo anteriormente, puede ser la original y la variable tratada a logaritmo. Esto mismo se hizo para todas las variables que fueron tratadas. Para el coeficiente de constructibilidad, la transformada y la original tienen una alta correlación con el valor unitario en logaritmo, por lo que esta variable explica el precio unitario, tal como se observa en el Anexo 4.

Luego se verificó con la variable superficie del terreno, y tal como se observa en el Anexo 5, existe mayor correlación con la variable transformada a logaritmo, pero la sin transformar, a pesar de tener una correlación baja de 0,067, se ha demostrado al menos en parte, que la superficie del terreno explica el precio unitario. En cambio, la correlación entre la variable dependiente y la independiente transformada a cuadrática tienen una baja correlación negativa, esto nos dice que los valores de una variable tienden a incrementarse mientras que la de la otra variable descienden. Estas comparaciones se hicieron con todas las variables independientes que se trataron.

5. ANÁLISIS REGRESIÓN LINEAL MÚLTIPLE

A partir de las variables mencionadas anteriormente, se creó la regresión lineal para el modelo con valor unitario en UF/m² transformado a logaritmo como variable dependiente. En la regresión se fue evaluando cuál variable tratada es mejor para el modelo, si la logarítmica, cuadrática o la original. Esto se hizo con las variables tratadas que generan importancia al modelo, es decir, con las que generó el Anexo 3 y con las que se tiene correlación con la dependiente, como se demostró anteriormente.

Después de evaluarlas, las variables que mejor se ajustan al modelo son el plazo entre la primera y última compra, el coeficiente de constructibilidad transformado a logaritmo, la altura máxima, el antejardín, la cantidad de viviendas, la superficie útil en porcentaje transformada a logaritmo, la densidad, si el proyecto se acoge al D.F.L.N.º2, la superficie edificada en m², la cantidad de estacionamientos requeridos por permiso, la superficie total del terreno, el número de roles, la cantidad de locales comerciales, el año de transacción, el número de oficinas, el coeficiente de ocupación de suelo transformado a logaritmo, el beneficio de fusión, la zona censal 2_7, la zona censal 4_11, la zona censal 2_5, la zona censal 5_2 y la zona censal 3_6. A partir de estas variables se creó una regresión lineal tal como se muestra en la Ilustración 3:

```

Residuals:
    Min       1Q   Median       3Q      Max
-0.60267 -0.14558  0.00878  0.14408  0.68022

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.549e+02  4.145e+01 -3.737 0.000330 ***
coef_de_constructibilidad_log  4.380e-01  1.475e-01  2.969 0.003851 **
plazo_primera_compra  1.908e-01  2.078e-02  9.183 1.72e-14 ***
altura_maxima  1.958e-02  1.360e-02  1.440 0.153501
superficie_edificada  2.303e-05  2.149e-05  1.072 0.286637
viviendas -1.861e-04  1.048e-02 -0.018 0.985866
n_rolas  1.723e-02  1.013e-02  1.700 0.092649 .
antejardin  1.468e-02  1.778e-02  0.825 0.411344
s_total_terreno -1.661e-04  1.131e-04 -1.468 0.145640
s_util_porcentaje_log  1.412e-01  7.843e-02  1.801 0.075168 .
densidad  4.345e-03  9.630e-03  0.451 0.652960
estacionamientos -2.063e-03  1.463e-03 -1.411 0.161859
DFLN2 -4.003e-01  1.518e-01 -2.637 0.009901 **
año_transaccion  7.797e-02  2.058e-02  3.789 0.000276 ***
locales_comerciales -4.284e-03  1.096e-02 -0.391 0.696724
n_oficina -1.168e-03  8.864e-04 -1.317 0.191147
zona_censal_2_7 -2.690e-01  1.002e-01 -2.683 0.008706 **
zona_censal_4_11 -5.397e-01  1.607e-01 -3.359 0.001156 **
coef_ocupacion_suelo_log -1.016e-01  8.536e-02 -1.190 0.237222
zona_censal_2_5  4.295e-01  1.602e-01  2.681 0.008775 **
beneficio_de_fusion  8.046e-02  9.116e-02  0.883 0.379872
zona_censal_5_2 -9.490e-02  1.309e-01 -0.725 0.470430
zona_censal_3_6 -1.250e-01  1.611e-01 -0.776 0.439682
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2621 on 88 degrees of freedom
Multiple R-squared:  0.7883,    Adjusted R-squared:  0.7353
F-statistic: 14.89 on 22 and 88 DF,  p-value: < 2.2e-16

```

Ilustración 3: Resultado de la regresión lineal con las variables importantes

Fuente: Elaboración propia

Un concepto importante que hay que comprender es la interpretación de los *p-values*, que son el nivel de significancia que se establece para determinar qué tan pequeño debe ser el valor p para rechazar la hipótesis nula. En este análisis, como se muestra en el resultado de R en la Ilustración 3, un valor de p bajo (por debajo de 0,05) se considera significativo para el modelo y se indica con una estrella. Los valores más bajos se asocian con una mayor importancia, donde las puntuaciones inferiores a 0,01 se indican con dos estrellas y las puntuaciones inferiores a 0,001 se indican con tres estrellas y se consideran muy significativas.

Al observar la regresión con todas las variables relevantes del modelo, se muestra que muchas de las variables independientes no son significativas, lo que es un indicativo de que podrían no contribuir al modelo. A raíz de esto se realiza la selección de los mejores predictores basándose en

las variables de esta regresión. Se emplea la estrategia de stepwise mixto, que es un método para seleccionar variables, y el valor matemático empleado para determinar la calidad del modelo va a ser Akaike (AIC), que es una métrica que se utiliza para comparar el ajuste de varios modelos de regresión. El mejor modelo que resultó luego de emplear el proceso de selección se muestra en la Ilustración 4:

```

Coefficients:
  (Intercept)  coef_de_constructibilidad_log  plazo_primera_compra
-1.691e+02    4.374e-01                        1.937e-01
altura_maxima  n_roles                                s_total_terreno
 2.695e-02    1.862e-02                        -1.343e-04
s_util_porcentaje_log  densidad  estacionamientos
 2.060e-01    4.480e-03                        -1.185e-03
DFLN2        año_transaccion  n_oficina
-4.449e-01   8.477e-02                        -1.271e-03
zona_censal_2_7  zona_censal_4_11  coef_ocupacion_suelo_log
-2.403e-01   -5.098e-01        -1.484e-01
zona_censal_2_5
 4.592e-01
  
```

Ilustración 4: Variables proceso de selección

Fuente: Elaboración propia

Al realizar la regresión lineal con estas variables ha dado como resultado:

```

Residuals:
  Min       1Q   Median       3Q      Max
-0.59231 -0.14886  0.01118  0.16074  0.69655

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.691e+02   3.696e+01  -4.575 1.44e-05 ***
coef_de_constructibilidad_log  4.374e-01   1.202e-01   3.639 0.000445 ***
plazo_primera_compra          1.937e-01   2.028e-02   9.555 1.48e-15 ***
altura_maxima                 2.695e-02   1.100e-02   2.451 0.016062 *
n_roles                       1.862e-02   9.142e-03   2.037 0.044431 *
s_total_terreno              -1.343e-04   9.649e-05  -1.392 0.167304
s_util_porcentaje_log        2.060e-01   6.777e-02   3.039 0.003067 **
densidad                     4.480e-03   2.384e-03   1.879 0.063321 .
estacionamientos            -1.185e-03   8.867e-04  -1.336 0.184660
DFLN2                       -4.449e-01   1.251e-01  -3.558 0.000585 ***
año_transaccion             8.477e-02   1.838e-02   4.612 1.24e-05 ***
n_oficina                   -1.271e-03   8.550e-04  -1.486 0.140521
zona_censal_2_7             -2.403e-01   9.423e-02  -2.551 0.012351 *
zona_censal_4_11           -5.098e-01   1.566e-01  -3.256 0.001567 **
coef_ocupacion_suelo_log    -1.484e-01   6.732e-02  -2.205 0.029897 *
zona_censal_2_5            4.592e-01   1.526e-01   3.009 0.003351 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2583 on 95 degrees of freedom
Multiple R-squared:  0.778,    Adjusted R-squared:  0.743
F-statistic: 22.2 on 15 and 95 DF,  p-value: < 2.2e-16
  
```

Ilustración 5: Resultado de la regresión lineal con las variables predictoras

Fuente: Elaboración propia

En comparación con la regresión anterior, en la Ilustración 5 se puede observar una mayor cantidad de variables significativas para el modelo. Además de un mayor R- cuadrado ajustado y un menor error estándar residual que será explicado más adelante. Sin embargo, sigue habiendo variables que no contribuyen en el modelo, como la superficie total del terreno, el estacionamiento y el número de oficinas. Estas variables no aportan al modelo, por lo que se irá probando la eliminación de cada una por separado para ver si afecta a las demás variables.

La primera en ser eliminada es la superficie total del terreno, la nueva regresión se puede ver en el Anexo 6. Como se puede observar, al eliminar esta variable, aumentó la significancia de la variable estacionamiento, pero disminuyó la significancia de la variable densidad y la cantidad de roles.

A continuación, se probó eliminando solo la variable estacionamiento, qué era otra variable sin significancia. Al eliminarla, se muestra que todas las demás variables son significativas para el modelo, por lo cual esta es la regresión final, tal como se muestra en la Ilustración 6.

```

Residuals:
    Min       1Q   Median       3Q      Max
-0.62988 -0.14040  0.01179  0.14593  0.72322

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    -1.738e+02  3.695e+01  -4.703  8.60e-06 ***
coef_de_constructibilidad_log  4.109e-01  1.190e-01   3.451  0.000831 ***
plazo_primera_compra  1.942e-01  2.036e-02   9.540  1.44e-15 ***
altura_maxima    2.357e-02  1.074e-02   2.194  0.030678 *
n_roles          1.950e-02  9.156e-03   2.130  0.035730 *
s_total_terreno  -2.270e-04  6.729e-05  -3.374  0.001069 **
s_util_porcentaje_log  1.927e-01  6.732e-02   2.862  0.005161 **
densidad         6.185e-03  2.022e-03   3.058  0.002887 **
DFLN2            -4.816e-01  1.225e-01  -3.932  0.000159 ***
año_transaccion  8.718e-02  1.837e-02   4.747  7.21e-06 ***
n_oficina        -1.629e-03  8.151e-04  -1.999  0.048439 *
zona_censal_2_7  -2.397e-01  9.461e-02  -2.533  0.012921 *
zona_censal_4_11 -5.141e-01  1.572e-01  -3.271  0.001489 **
coef_ocupacion_suelo_log -1.521e-01  6.754e-02  -2.251  0.026642 *
zona_censal_2_5   4.774e-01  1.526e-01   3.128  0.002329 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2593 on 96 degrees of freedom
Multiple R-squared:  0.7738,    Adjusted R-squared:  0.7409
F-statistic: 23.46 on 14 and 96 DF,  p-value: < 2.2e-16

```

Ilustración 6: Resultado final de la regresión lineal múltiple

Fuente: Elaboración propia

5.1. EVALUACIÓN NORMALIDAD Y HOMOSCEDASTICIDAD DE LA REGRESIÓN

Luego de realizar la regresión lineal, se evaluó la normalidad de los residuos usando la prueba de hipótesis de Shapiro. Esta plantea la hipótesis nula que dice que una muestra proviene de una distribución normal y la hipótesis alternativa que dice que la distribución no es normal. Se eligió un nivel de significancia del 0,05.

```
shapiro-wilk normality test
data:  modelo$residuals
W = 0.98961, p-value = 0.5586
```

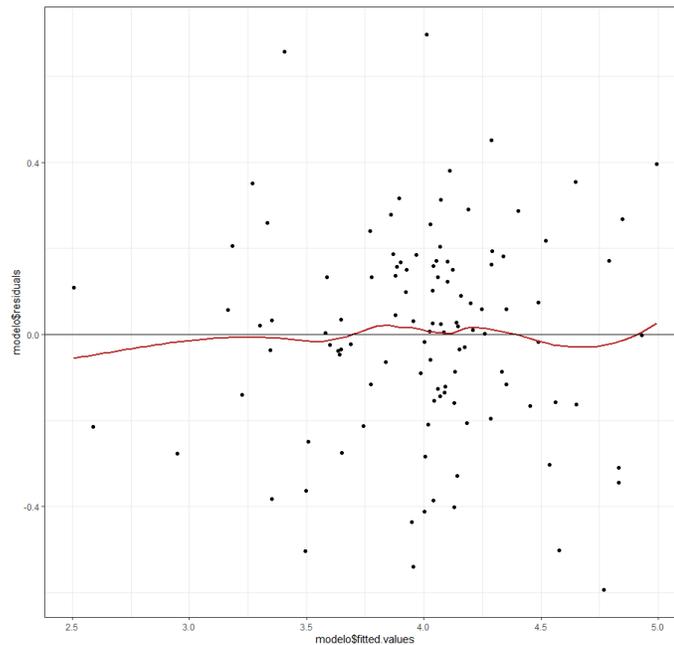
Ilustración 7: Prueba de Shapiro

Fuente: Elaboración propia

Como se observa en la Ilustración 7, se obtiene un valor p de 0,5586, por lo que no existe suficiente evidencia estadística para rechazar la hipótesis nula con un nivel de confianza del 95%, por lo que se asume normalidad de los residuos.

Luego de evaluar la normalidad, se analizó la variabilidad de los residuos, si son homoscedásticos o heteroscedásticos. Al realizar un gráfico de los residuos con los valores ajustados por el modelo, este debe mostrar una distribución de forma aleatoria en torno a cero, manteniendo aproximadamente la misma variabilidad a lo largo del eje x. Si se observa que no están idénticamente distribuidas, significa que la variabilidad es dependiente del valor ajustado, por lo tanto, no hay homoscedasticidad, no tienen la misma varianza (Rodrigo, 2016).

Gráfico 8: Residuos vs Valor Ajustado



Fuente: Elaboración propia

Como se observa en el Gráfico 8, el eje x corresponde a los valores ajustados y el eje y a los residuos correspondientes. Se muestra que la dispersión de los residuos no es constante, no mantienen la misma variabilidad, la varianza de los errores del modelo no es la misma para todas las observaciones, por lo que la regresión es heteroscedástica. Esto puede afectar en la eficiencia de la estimación del modelo, los errores estándar del resumen del modelo no son confiables.

Para verificar que el modelo es heteroscedástico se realiza la prueba de White, en donde BP representa el estadístico de prueba, df son los grados de libertad y *p-value* es el valor p. La prueba de White utiliza las siguientes hipótesis: la nula que dice que la homoscedasticidad está presente y la alternativa que nos dice que hay heteroscedasticidad.

```
studentized Breusch-Pagan test
data: modelo
BP = 41.199, df = 14, p-value = 0.0001655
```

Ilustración 8: Prueba de White

Fuente: Elaboración propia

Como se observa en la Ilustración 8 el valor p corresponde a 0,0001655, por lo que se rechaza la hipótesis nula con un intervalo de confianza del 95%. Hay evidencia para decir que en este modelo de regresión está presente la heteroscedasticidad.

El problema de la heteroscedasticidad es que los errores estándar se subestiman, y al estar en el cociente del estadístico de prueba t, esto implica que el modelo podría estar arrojando valores significativos cuando no lo son. Una forma de explicar este problema es utilizar errores estándar robustos que proporcionan una medida más precisa del verdadero error estándar de un coeficiente de regresión.

```

Residuals:
  Min       1Q   Median       3Q      Max
-0.62988 -0.14040  0.01179  0.14593  0.72322

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    -1.738e+02  4.268e+01  -4.071  9.63e-05 ***
coef_de_constructibilidad_log  4.109e-01  1.823e-01   2.253  0.026517 *
plazo_primera_compra    1.942e-01  1.883e-02  10.312 < 2e-16 ***
altura_maxima    2.357e-02  1.339e-02   1.759  0.081700 .
n_roles    1.950e-02  7.531e-03   2.590  0.011101 *
s_total_terreno   -2.270e-04  6.120e-05  -3.710  0.000347 ***
s_util_porcentaje_log  1.927e-01  7.306e-02   2.637  0.009745 **
densidad    6.185e-03  1.883e-03   3.285  0.001426 **
DFLN2    -4.816e-01  1.054e-01  -4.571  1.45e-05 ***
año_transaccion    8.718e-02  2.108e-02   4.135  7.60e-05 ***
n_oficina    -1.629e-03  1.007e-03  -1.617  0.109077
zona_censal_2_7   -2.397e-01  8.618e-02  -2.781  0.006525 **
zona_censal_4_11  -5.141e-01  3.008e-01  -1.709  0.090626 .
coef_ocupacion_suelo_log  -1.521e-01  6.716e-02  -2.264  0.025812 *
zona_censal_2_5    4.774e-01  1.738e-01   2.746  0.007198 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2593 on 96 degrees of freedom
Multiple R-squared:  0.7738,    Adjusted R-squared:  0.7409
F-statistic: 32.81 on 14 and 96 DF,  p-value: < 2.2e-16

```

Ilustración 9: Regresión con errores estándares robustos

Fuente: Elaboración propia

Cuando se utiliza errores estándares robustos, las estimaciones de los coeficientes no cambian, son iguales para ambas regresiones. Como se observa en la Ilustración 9 el error estándar de las variables predictoras varió con respecto a la regresión original, en algunas aumentó y en otras disminuyó. A raíz de esto varió el valor t de cada coeficiente, esto se debe a que el valor t es el cociente entre el coeficiente estimado y los errores estándar. Cuanto mayor sea el error estándar,

menor será el estadístico de prueba. Los valores p también cambiaron y esto se debe a que valores t más pequeños están asociados a con valores p más grandes.

5.2. EVALUACIÓN DEL MODELO

Hay varias estadísticas que examinan el ajuste del modelo e indican su calidad en términos de precisión, entre ellas se encuentra el error estándar residual. Este mide la desviación estándar de los residuos en un modelo de regresión, se puede considerar como el error promedio del modelo y se representa por ϵ . Mientras menor sea este valor, mejor se podrá ajustar el modelo a los datos. Como se observa en la Ilustración 9, el error estándar residual corresponde a un 0,2593, que indica que, en promedio, la predicción del precio para cualquier valor de las variables independiente tendría un error aproximado de 0,2593 UF/m².

También se puede observar un resumen estadístico que muestra la distribución de los residuos (error). El residuo es la diferencia entre el valor observado y el valor que predice el modelo de regresión. Esta tabla muestra el valor medio, que es el punto medio de los datos, que en este caso corresponde a 0,01179. También muestra el residuo mínimo que es -0,62988 y el residuo máximo que es 0,72322. Con estos dos valores se calcula el rango de los valores residuales, que es de 1,3531, que es la diferencia entre el valor mínimo y el máximo.

Otra estadística es el R-cuadrado que representa el porcentaje de variación en la variable dependiente explicada por la variación en la variable independiente. Mientras mayor sea el R cuadrado, mejor se ajusta el modelo a los datos, en este caso corresponde a un 0,7738. El valor R-cuadrado ajustado se interpreta de la misma forma, pero considera cuántas variables independientes se incluyen en el modelo. Esto es importante, ya que agregar más variables hará que aumente el valor normal de R-cuadrado, sin embargo, estas variables pueden no ser estadísticamente significativas, por lo que se corre el riesgo de aumentar artificialmente este valor generando sobre ajuste en el modelo. Sin embargo, el valor R cuadrado ajustado penaliza la adición de variables insignificantes y, por lo tanto, puede usarse para optimizar la selección de características. En la Ilustración 9 se observa que el R cuadrado ajustado es de 0,7409. Esto nos indica que las variables independientes explican en un 74,09% el precio unitario en UF/m².

Otra estadística que se observa es el estadístico F, que es un método para probar la significancia conjunta de múltiples regresores, mide la importancia general del modelo. Tiene un valor p que indica la significancia estadística, al igual que el descrito para los coeficientes, mide la probabilidad de que la hipótesis nula sea verdadera. En el caso del estadístico F, la hipótesis nula es que un modelo sin características predecirá el resultado tan bien como el modelo que se examina. El valor p que se observa en la Ilustración 9 corresponde a $2,2e-16$, el cual es menor al nivel de significancia típico 0,05, por lo que el modelo es estadísticamente significativo. Se rechaza la hipótesis nula, lo que respalda que el modelo predice mejor que un modelo sin características (variables).

5.3. EVALUACIÓN DE LOS COEFICIENTES

Para examinar la importancia del modelo, se debe comprender dos medidas estadísticas: el valor t (*t-value*) y el valor p (*p-value*). El valor t es la relación entre el valor estimado y su error estándar e indica que tan grande es la estimación en relación con su error, es decir, los valores más grandes representan una variabilidad más pequeña, mientras mayor sea su valor, mayor será la evidencia en contra de la hipótesis nula. En cambio, el valor p es un valor de probabilidad calculado, cada valor t está asociado con un valor p. Este nos indica la probabilidad de obtener resultados por lo menos tan extremos como el de los datos de la muestra, asumiendo que la hipótesis nula es verdadera (MIT SCHOOL OF ARCHITECTURE AND PLANNING, 2020).

Un concepto importante que hay que comprender al interpretar el valor p es el nivel de significancia, que se establece para determinar qué tan pequeño debe ser el valor p para rechazar la hipótesis nula. En este análisis, como se muestra en el resultado de R en la Ilustración 5, un valor de p bajo (por debajo de 0,05) se considera significativo para el modelo y se indica con una estrella. Los valores más bajos se asocian con una mayor importancia, donde las puntuaciones inferiores a 0,01 se indican con dos estrellas y las puntuaciones inferiores a 0,001 se indican con tres estrellas y se consideran muy significativas.

En este análisis, el intercepto tiene un valor p pequeño ($9,63e-05$) y es muy significativo, como indican tres estrellas. La importancia de la estimación alfa tiene sentido, ya que este valor de

relación de referencia entre el precio unitario en UF/m² y las variables independientes mencionadas anteriormente juegan un papel importante en el modelo.

En la Ilustración 9 se puede observar que cada coeficiente indica los valores estimados, el error estándar asociado a esos valores, su valor t y la probabilidad de que la hipótesis nula sea cierta (valor p) basado en el valor t. Además, contiene estimaciones para los coeficientes alfa (constante) y beta para cada variable independiente.

La estimación alfa, conocida como constante, es el valor de la variable dependiente cuando la variable independiente es cero. No siempre el valor de intercepción tiene sentido en el contexto de un análisis y, por lo tanto, no siempre se considera significativo y posteriormente interpretado (MIT SCHOOL OF ARCHITECTURE AND PLANNING, 2020).

La constante muestra el valor promedio de la variable dependiente cuando las variables predictivas son cero. En esta regresión, la aplicación promedio es de -173,8 cuando el coeficiente de las demás variables independientes es de cero. Esto no tiene sentido, ya que, por ejemplo, el año de transacción no puede ser cero, pero se necesita el número -173,8 para formar una ecuación de regresión.

Con respecto a la interpretación de la estimación beta de una variable numérica, esta indica el cambio en el valor (y) con un cambio de unidad en la característica (x). El signo positivo o negativo indica la relación en términos de si los valores se mueven juntos (los valores de resultado (y) y característica (x) aumentan o disminuyen juntos) o de manera diferente (un valor aumenta mientras que el otro disminuye y viceversa) (MIT SCHOOL OF ARCHITECTURE AND PLANNING, 2020).

Las transformaciones logarítmicas influyen en la interpretación de la regresión. Un modelo de regresión lineal examina los cambios de unidades entre las variables x e y interpretándose como un cambio de una sola unidad en x asociado con un cambio constante en y. Sin embargo, en términos de transformaciones logarítmicas, un cambio en una variable transformada logarítmicamente representará un cambio porcentual. Esto se describe para tres tipos diferentes de

modelos que se pueden observar en la Tabla 5. Dos de estos tienen una variable transformada a logaritmo y otra variable en forma lineal, mientras que el otro tiene sus dos variables transformadas a logaritmo.

Tabla 5: Interpretación estimación beta

Modelo	Variable dependiente	Variable independiente	Interpretación	Ejemplo
lin-log	y	log(x)	Un cambio de punto porcentual en x da como resultado un cambio unitario en y igual a beta dividido por 100.	Si $\beta = 3000$, entonces un aumento del 1% en x daría como resultado un aumento en y de 30 unidades
log-lin	log(y)	x	Un cambio de unidad en x da como resultado un cambio de punto porcentual en y igual a beta multiplicado por 100.	Si $\beta = 0,3$, entonces un aumento de una unidad en X daría como resultado un aumento del 30 % en y.
log-log	log(y)	log(x)	Un cambio de punto porcentual de x da como resultado un cambio de punto porcentual de y.	Si $\beta = 300$, el aumento del 1% en x daría como resultado un aumento de 3% en y.

Fuente: MIT SCHOOL OF ARCHITECTURE AND PLANNING, 2020

La Tabla 5 contiene información sobre la interpretación de los tres tipos de modelos: "lineal - logaritmo", "logaritmo - lineal" y "logaritmo - logaritmo", cada uno con un ejemplo. En este caso, se utiliza el log-log y el log-lin. En el log-log, el coeficiente va a ser la elasticidad.

Es importante tener en cuenta que la magnitud de cada coeficiente depende de las unidades en las que se mida la variable dependiente. Para determinar el impacto de las variables en el modelo, se usan los coeficientes.

En los resultados que se muestran en la Ilustración 9, las variables más relevantes y significativas para el modelo son el plazo entre la primera y última compra, la superficie total del terreno, la variable binaria D.F.L.N.^o2 y el año en que se realiza la transacción. Estas variables tienen un valor p de 0,001, por lo que se rechaza la hipótesis nula, el valor beta es diferente de 0.

La estimación beta de la relación entre el precio unitario en logaritmo y el plazo de la primera compra, que es la cantidad de años que transcurre entre la primera hasta la última compra, es de $1,942e-01$ (o 0,1942). Esto indica que, si aumenta un año más la transacción entre la primera compra del terreno hasta la última, entonces el precio unitario en UF/m² aumentará en un 19,42%.

La estimación beta de la relación entre el precio unitario en logaritmo y la superficie total del terreno es de $-2,270e-04$ (o -0,0002270). Esto nos dice que, al aumentar en un metro cuadrado la superficie del terreno, el precio unitario en UF/m² disminuye en un 0,02270%. Esto tiene sentido, ya que los terrenos con mayor superficie tienden a tener un valor unitario menor.

La estimación beta de la relación entre el precio unitario en logaritmo y la variable binaria D.F.L.N.^o2, que indica si el proyecto se acoge a esa normativa o no, es de $-4,816e-01$ (o -0,4816). Esto indica que, los permisos que acogen esta normativa disminuyen en promedio 48,16% su precio unitario en UF/m² en comparación a los que no acogen este permiso, manteniendo las demás variables constantes.

La estimación beta de la relación entre el precio unitario en logaritmo y el año de transacción, que es el año de la última transacción que se hizo para ese permiso, es de $8,718e-02$ (o 0,08718). Esto indica que, al aumentar un año la transacción, el precio unitario en UF/m² aumentará en un 8,718%.

A continuación, con un valor p de 0,01, se encuentran las variables superficie útil edificada en porcentaje, densidad, la zona censal 2_7 y la 2_5. Estas variables siguen siendo significativas e importantes para el modelo, pero con una menor intensidad. También se rechaza la hipótesis nula.

La estimación beta de la relación entre el precio unitario en logaritmo y la superficie útil del terreno en porcentaje transformada a logaritmo es de $1,927e-01$ (o 0,1927). Esto quiere decir que, si la superficie útil aumenta un 1%, entonces el precio unitario en UF/m² aumenta en 0,1927%. Esto tiene sentido, ya que mientras más se pueda construir, más podrá vender la inmobiliaria, por lo tanto, debería estar disponible a pagar más por el terreno.

La estimación beta de la relación entre el precio unitario en logaritmo y la densidad permitida es de $6,185e-03$ (o 0,006185). Esto indica que, al aumentar la densidad en un VIV, el precio unitario en UF/m² aumentará en un 0,6185 %. Esto tiene sentido, ya que, con densidades permitidas más altas, se puede dividir la superficie a construir permitida en más unidades. Las unidades más pequeñas son más preferidas en la comuna y se venden con mayor velocidad. Además, tienen un valor unitario de venta más alto que los departamentos más grandes.

La estimación beta de la relación entre el precio unitario en logaritmo y la variable binaria zona censal 2_7 correspondiente al sector cerca del parque Inés de Suárez es de $-2,397e-01$ (o -0,2397). Esto indica que, los proyectos que se encuentren en esa zona disminuyen en promedio 23,97% su precio unitario en UF/m² en comparación a los que no se encuentran en esa zona, manteniendo las demás variables constantes. Este es un sector residencial.

La estimación beta de la relación entre el precio unitario en logaritmo y la variable binaria zona censal 2_5 correspondiente al sector cercano a Tobalaba es de $4,774e-01$ (o 0,4774). Esto indica que, los proyectos que se encuentren en esa zona aumentan en promedio 47,74% su precio unitario en UF/m² en comparación a los que no se encuentran en esa zona, manteniendo las demás variables constantes. Este es un sector comercial.

Luego se encuentran las variables con un valor p de 0,05, estas siguen siendo significativas e importantes para el modelo, pero con una menor importancia que las demás. Entre estas variables

se encuentra el coeficiente de constructibilidad, el número de roles del permiso y el coeficiente de ocupación de suelo. Estas variables rechazan la hipótesis nula, el valor beta es diferente de 0.

La estimación beta de la relación entre el precio unitario transformado a logaritmo y el coeficiente de constructibilidad también transformado a logaritmo es de $4,109e-01$ (o 0,4109). Esto quiere decir que, si el coeficiente de constructibilidad aumenta en un 1%, entonces el precio unitario en UF/m² aumenta en 0,4109%. Tiene sentido de que el coeficiente sea positivo, ya que si aumenta el espacio máximo que se puede construir en el terreno, se espera que aumente el precio de ese terreno. Si aumenta el valor unitario se da por la condición de que se permite una mayor superficie construida que puede ser orientada a un nicho de mayor condición económica

La estimación beta de la relación entre el precio unitario en logaritmo y el número de roles de cada permiso es de $1,950e-02$ (o 0,01950). Esto nos indica que, al aumentar en un rol, el precio unitario en UF/m² aumentará en un 1,950%. Esto tiene sentido, ya que, al aumentar un rol, se aumenta una propiedad, porque el rol de avalúo es lo que identifica una propiedad, lo que conlleva a aumentar el precio unitario.

La estimación beta de la relación entre el precio unitario transformado a logaritmo y el coeficiente de ocupación de suelo también transformado a logaritmo es de $-1,521e-01$ (o -0,1521). Esto quiere decir que, si el coeficiente de ocupación de suelo aumenta un 1%, entonces el precio unitario en UF/m² disminuye en 0,1521%.

Finalmente, entre las variables que se acercan a ser significativas para el modelo, pero no lo son con un valor p de 0,1 se encuentran la altura máxima permitida y la zona censal 4_11. Y la variable no significativa es el número de oficinas. Estas variables no rechazan la hipótesis nula, por lo que el valor beta puede ser 0.

La estimación beta de la relación entre el precio unitario en logaritmo y la altura máxima, que es la altura máxima permitida representada en pisos, es de $2,357e-02$ (o 0,02357). Esto indica que, si se aumenta un piso de altura, es decir, si construyo un piso más, entonces el precio unitario en

UF/m² aumentará en un 2,357%. Tiene sentido que el valor beta sea positivo, ya que mientras más pisos se pueda construir, mayor será el precio del terreno.

La estimación beta de la relación entre el precio unitario en logaritmo y la variable binaria zona censal 4_11 correspondiente al sector cercano a Andrés Bello es de $-5,141e-01$ (o $-0,5141$). Esto indica que, los proyectos que se encuentren en esa zona disminuyen en promedio 51,41% su precio unitario en UF/m² en comparación a los que no se encuentran en esa zona, manteniendo las demás variables constantes.

La estimación beta de la relación entre el precio unitario en logaritmo y el número de oficinas es de $-1,629e-03$ (o $-0,001629$). Esto indica que, al aumentar una oficina, es decir, construir una oficina más, el precio unitario en UF/m² disminuirá en un 0,1629%.

A continuación, se presenta un ejemplo que compara un proyecto con un alto valor unitario con respecto a uno de bajo valor. Al correr el modelo, el valor unitario mayor corresponde a 4,520723 UF/m², en cambio, el de menor precio corresponde a 2,949801 UF/m².

Tabla 6: Comparación proyecto con alto y bajo valor unitario

Variable	Valor unitario mayor	Valor unitario menor
Plazo primera compra	La diferencia entre la primera y última compra corresponde a dos años.	La diferencia entre la primera y última compra corresponde a cuatro años.
Superficie total del terreno	Corresponde a 1.309,98.	Corresponde a 5.141,37.
D.F.L.N.º2	El proyecto no se acoge a esta normativa.	El proyecto se acoge a esta normativa.
Año de transacción	Corresponde al año 2017.	Corresponde al año 2020.
Superficie útil del terreno	Corresponde al 69%.	Corresponde al 71%
Densidad	No tiene densidad.	Corresponde a 134 VIV.
Zona censal 2_7	No corresponde a esta zona.	No corresponde a esta zona.
Zona censal 2_5	Corresponde a esta zona.	No corresponde a esta zona.
Coefficiente de constructibilidad	Corresponde a 3,64.	Corresponde a 2,60.
Número de roles	Corresponde a 4.	Corresponde a 2.
Coefficiente de ocupación de suelo	Corresponde a 0,6.	Corresponde a 0,4.
Altura máxima	Corresponde a 10 pisos.	Corresponde a 9 pisos.
Zona censal 4_11	No corresponde a esta zona.	Corresponde a esta zona.
Número de oficinas	Corresponde a 18 oficinas.	Corresponde a 8 oficinas.

Fuente: Elaboración propia

Como se observa en la Tabla 6, el proyecto con valor unitario más alto es más caro debido a que tiene una superficie del terreno menor al otro proyecto, y como se dijo anteriormente, los terrenos con mayor superficie tienden a tener un precio unitario menor. Asimismo, este proyecto se encuentra en la zona de Tobalaba, que corresponde a un sector comercial, en cambio, el otro proyecto se encuentra en la zona cerca de la avenida Andrés Bello, que es un sector cercano al río Mapocho, por lo que tienden a tener un precio más bajo.

También el proyecto no se acoge a la normativa D.F.L.N.º2, en cambio, el otro proyecto sí, y como se dijo anteriormente, el precio unitario será menor en los proyectos que acogen esta normativa. Además, el coeficiente de constructibilidad es mayor en este proyecto, y como se mencionó anteriormente, al permitir una mayor superficie construida, se espera que aumente el precio del terreno. Por último, la altura máxima permitida de edificios y el número de oficinas es mayor en este proyecto, por ende, aumenta el precio del terreno, en comparación con el otro proyecto que es más bajo.

6. CONCLUSIONES Y RECOMENDACIONES

En esta memoria, se ha investigado qué variables son las más relevantes o importantes para explicar el precio del suelo para proyectos inmobiliarios entre los años 2017 al 2021. Es de relevancia la investigación del precio del suelo, ya que no hay un registro objetivo del costo de este. Esto se debe a que no es un producto homogéneo, por lo tanto, es complejo comparar suelos con diferentes tamaños, formas y condiciones normativas.

Por medio de la regresión lineal múltiple, se evaluó la relación que existe entre las variables que resultaron relevantes para el modelo con respecto al precio unitario en UF/m². Como resultado se obtuvo que el valor p del estadístico F es menor a 0,05, que nos dice que el modelo es estadísticamente significativo, por lo que rechaza la hipótesis nula que dice que las variables no tienen efecto, es decir, todos los coeficientes son iguales a 0. También se obtuvo que el modelo explica un 74,09% el precio unitario en UF/mt². Es un buen ajuste del modelo, ya que, mientras mayor es este valor, mayor se ajusta el modelo a los datos.

Las variables más importantes y significantes del modelo corresponden al plazo entre la primera y última compra, la superficie total del terreno, la variable binaria D.F.L.N^o2 y el año en que se realiza la transacción. De estas cuatro variables, solamente tres tienen un coeficiente positivo.

Una de esas variables es el plazo entre la primera y última compra que, al aumentar un año el plazo entre la primera y última transacción del terreno, el precio unitario en UF/m² también aumentará. Esto tiene sentido, ya que, si se demora un año más en efectuar la última compra, el terreno en ese año aumentará su valor, por ende, aumentará el valor del precio unitario. Esto también pasa con el año de transacción, que es otra variable considerable con coeficiente positivo. Al aumentar un año la transacción del terreno, el valor unitario aumentará también.

La otra variable relevante con coeficiente positivo es la variable binaria D.F.L.N.^o2, que indica si el proyecto se acoge a esa normativa o no. Esta variable indica que los permisos que se acogen a esta normativa disminuyen su precio unitario en UF/m² en comparación a los que no acogen el permiso. Esto se debe a que principalmente son proyectos enfocados a la primera vivienda para la

clase media y media emergente, con beneficio tributario y hasta una superficie construida menor a 140 m². Estos proyectos por lo general se ubican en las zonas no céntricas de las comunas.

Con respecto a la variable correspondiente a la superficie del terreno, esta tiene un coeficiente negativo, es decir, al aumentar un m² del terreno, el precio unitario disminuye. Esto tiene sentido debido a que los terrenos más grandes, tienden a tener un valor unitario menor. La relación no es lineal, es como comparar la compra al por mayor y al detalle, comprar al detalle es más caro.

En la práctica, esto está sustentado en que un terreno muy grande va a implicar un proyecto más grande, y un mayor plazo de venta. Dado esto, el plazo de ventas también sube, por lo tanto, se amplía el plazo en que la inmobiliaria recupera el dinero invertido en comprar el terreno y construir el edificio. Para soportar este mayor plazo, el proyecto debe endeudarse más (en comparación con uno que vende menos unidades). Dado esto, podrá pagar un valor unitario menor por cada m² de terreno, debido a que las utilidades finales del proyecto traídas al momento de la evaluación serán menores.

Providencia tiene muchas zonas censales y son solo tres las que finalmente son importantes para el modelo. Para evaluar la importancia de todas las zonas censales y ver cuáles son las más relevantes, se utilizó el paquete Caret, que evalúa que variables son de importancia para el modelo, luego se evaluó cuáles de esas zonas considerables son las que mejor predicen el modelo. Estas zonas corresponden a la zona censal 2_5 que corresponde al sector cercano a Tobalaba, la zona censal 2_7 correspondiente al sector cercano al Parque Inés de Suárez y la zona censal 4_11 correspondiente al sector cercano a Andrés Bello, tal como se observa en la Ilustración 10. De estas zonas, solo dos son significativas, la zona 2_5 y la zona 2_7.

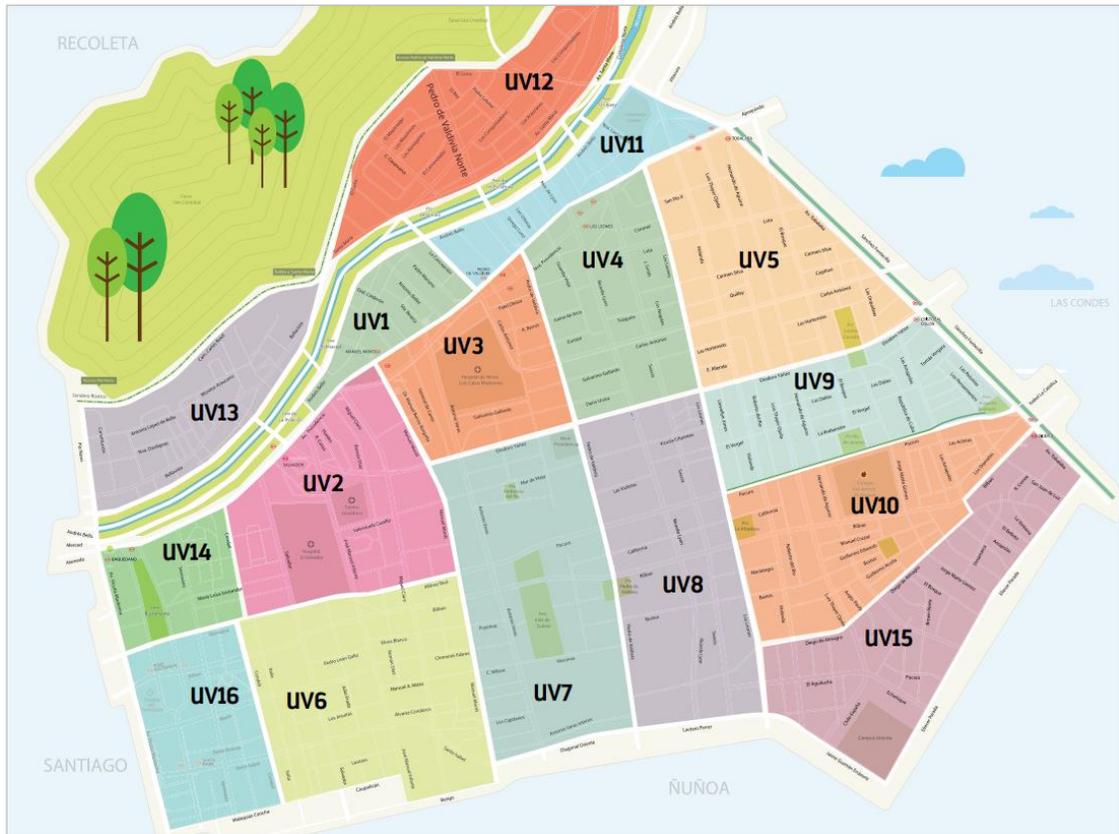


Ilustración 10: Plano Providencia

Fuente: Elaboración propia

Con respecto a la zona censal 2_7, sector ubicado cerca del parque Inés de Suárez, tiene un coeficiente negativo, lo que nos dice que, si el terreno se encuentra en ese territorio, el precio unitario disminuye en comparación a los que no se encuentran ahí. Esto se puede deber a que en esa provincia no se encuentran estaciones de metros cercanas, por lo que la movilidad es más reducida. También se puede deber a que se encuentra más cerca del sector Barrio Italia, en comparación con las demás zonas.

Con respecto a la zona censal 2_5, sector ubicado cerca de Tobalaba, tiene un coeficiente positivo, lo que nos dice que, si el terreno se encuentra en ese territorio, el precio unitario aumenta en comparación a los que no se encuentran ahí. Esto se puede deber a que es una provincia que se encuentra ubicada cerca de dos estaciones de metro, como es el metro Tobalaba y el metro Cristóbal Colón.

El propósito de este proyecto fue encontrar los factores más predican el precio del suelo y a partir de estas variables, analizar el impacto que provocan en el precio unitario en UF/m². Más de la mitad de las variables importantes tienen un impacto positivo en el precio, es decir, a medida que aumentan una unidad, el precio unitario también aumenta.

Con respecto a futuras líneas de investigación, a base de esta memoria, se recomienda enfocarse en la creación de un índice que mida la evolución del precio del suelo, en particular, para proyectos inmobiliarios. La creación de este índice ayudaría a comprender el precio de la vivienda, ya que este es un factor fundamental en el proceso de venta de esta.

7. BIBLIOGRAFÍA

- Agustin Gatto, R. D. E. C., s.f. *Cómo construir un índice de precios a medida de cada empresa*, s.l.: s.n.
- Anon., s.f. s.l.:s.n.
- ArcGis, s.f. *ArcGis Insights*. [En línea]
Available at: <https://doc.arcgis.com/es/insights/latest/analyze/regression-analysis.htm>
- Arévalo, R., 2018. *Catálogo arquitectura*. [En línea]
Available at: <https://www.catalogoarquitectura.cl/cl/oguc/requisitos-de-distanciamientos-y-rasantes-para-edificaciones-establecidos-por-la-oguc>
- Banco Central de Chile, 2014. *Índice de Precios de Viviendas en Chile: Metodología y Resultados*, s.l.: Estudios Económicos Estadísticos N°107.
- Cepchile, 2019. *Centro de estudios públicos*. [En línea]
Available at:
https://www.cepchile.cl/cep/site/docs/20190910/20190910163323/pder518_clarrain_srazmilic.pdf
- CNN CHILE, 2019. *CNN CHILE*. [En línea]
Available at: https://www.cnnchile.com/economia/presidente-del-cchc-advierte-crisis-social-por-dificil-acceso-a-viviendas-se-esta-convirtiendo-en-un-bien-inalcanzable_20190510/
- Conservador de Molina, 2022. *Conservador de Molina*. [En línea]
Available at:
<https://www.conservadormolina.cl/preguntas.php#:~:text=%C2%BFQu%C3%A9%20es%20Foja%2C%20N%C3%BAmero%20y,practica%20una%20inscripci%C3%B3n%20de%20dominio.>
- Dagnino, J., 2014. *Regresión lineal*, s.l.: Revista chilena anestesia.
- Data Science Team, 2020. *DATA SCIENCE*. [En línea]
Available at: <https://datascience.eu/es/matematica-y-estadistica/definicion-del-valor-p/>
- Datademia, 2022. *Datademia*. [En línea]
Available at: <https://datademia.es/blog/que-es-r>
- Diaz, G. C., 2014. *PRECIO DEL SUELO Y METODOLOGÍAS DE AVALUACIÓN.*, Santiago: s.n.
- Editor de blogs de Minitab, 2016. *Minitab*. [En línea]
Available at: <https://blog.minitab.com/en/adventures-in-statistics-2/understanding-t-tests-t-values-and-t-distributions>
- Franco, F. d. l. C., 2017. *SlidePlayer*. [En línea]
Available at: <https://slideplayer.es/slide/10616738/>
- Generación M, 2021. *elmostrador*. [En línea]
Available at: <https://www.elmostrador.cl/generacion-m/2021/10/12/cuatro-factores-que-influyen-en-el-alza-historica-del-precio-de-las-viviendas/>
- Gerencia de Información Estadística, 2019. *Actualización metodológica del Índice de Precios de Vivienda que elabora el Banco Central de Chile*, s.l.: Departamento de microdatos.
- González, P., s.f. *Billin.net*. [En línea]
Available at: <https://www.billin.net/glosario/definicion-indice-de-precios-ipc/>
- Gutiérrez, M., 2015. *Economía y Negocios*. [En línea]
Available at: <http://www.economiaynegocios.cl/noticias/noticias.asp?id=136308>
- Holz, M., 2022. *Evolución del Índice de precios de materiales de la construcción: actualización a febrero de 2022*, s.l.: Biblioteca del Congreso Nacional de Chile.

IBM , 2023. *IBM*. [En línea]
 Available at: <https://www.ibm.com/docs/es/cognos-analytics/11.2.0?topic=terms-adjusted-r-squared>

IBM, s.f. *IBM*. [En línea]
 Available at: <https://www.ibm.com/cl-es/analytics/learn/linear-regression>

INEI, 2013. *Índice de Precios de Materiales de Construcción*, Lima: Dirección Técnica de Indicadores Económicos.

Instituto Nacional de Estadística, 2009. *Índice de Precios de Vivienda*, Madrid: s.n.

James H. Stock, M. M. W., 2012. *Introducción a la Econometría*. 3° edición ed. s.l.:Pearson Educación, S.A..

JMP Statistical Discovery, 2022. *JMP Statistical Discovery*. [En línea]
 Available at: https://www.jmp.com/es_co/statistics-knowledge-portal/exploratory-data-analysis/box-plot.html

JMP, 2023. *JMP*. [En línea]
 Available at: https://www.jmp.com/es_co/statistics-knowledge-portal/t-test.html#:~:text=Una%20prueba%20t%20puede%20usarse,de%20muestras%20dependientes%20o%20pareada.

Koporovic, R. G., 2013. *Develop Gestión Comercial*. [En línea]
 Available at: <https://develop.cl/como-saber-el-precio-de-un-terreno-inmobiliario/>

León, E., 2021. *Baoss Analytics Everywhere*. [En línea]
 Available at: <https://www.baoss.es/te-contamos-que-es-el-lenguaje-r/>
 [Último acceso: 5 noviembre 2022].

López, J. F., 2023. *economipedia*. [En línea]
 Available at: <https://economipedia.com/definiciones/r-cuadrado-coeficiente-determinacion.html>

Maximilianos, 2021. *Rankia*. [En línea]
 Available at: <https://www.rankia.cl/blog/mejores-creditos-hipotecarios/5022509-que-proyecto-inmobiliario>

Minitab Blog Editor, 2019. *Minitab*. [En línea]
 Available at: <https://blog.minitab.com/es/analisis-de-regresion-como-puedo-interpretar-el-r-cuadrado-y-evaluar-la-bondad-de-ajuste>

MIT SCHOOL OF ARCHITECTURE AND PLANNING, 2020. *How to do a regression analysis*, s.l.: s.n.

MIT SCHOOL OF ARCHITECTURE AND PLANNING, 2020. *Interpreting a regression analysis*, s.l.: s.n.

Montes, D., 2018. *Proyectos Gestión Conocimiento*. [En línea]
 Available at: <https://www.pgconocimiento.com/diagrama-boxplot/>

Osorio, V., 2021. *Diario Financiero*. [En línea]
 Available at: <https://www.df.cl/empresas/construccion/precio-del-suelo-en-el-gran-santiago-aumenta-un-218-en-la-ultima-decada>

Ostornol, J. T., 2022. *Diario Financiero*. *Chile terminó entre los 25 países del mundo con mayor alza en el precio de la vivienda*, 21 marzo.

Oxford Languages, 2022. *Oxford University Press*. s.l.:Oxford Languages.

Parrado, C. F., 2009. *Evolución de los precios de viviendas en Chile*, s.l.: Revista Economía Chilena.

Peña, K., 2018. *Diario Financiero*. [En línea]
 Available at: <https://www.df.cl/empresas/infraestructura-inmobiliaria/nunoa-y-providencia-registran-la-mayor-alza-en-el-valor-de-viviendas-en>

Pérez, J. L., 2013. *LA MATEMÁTICA Y LA ESTADÍSTICA: UNA ORQUESTA HECHA INSTRUMENTO*. [En línea]
 Available at: <https://jlllopisperez.com/2013/04/30/test-de-wald/#:~:text=El%20Test%20de%20Wald%20es,generalista%2C%20aplicable%20en%20muchos%20%20C3%A1mbitos.>

portal inmobiliario, s.f. *portalinmobiliario.com*. [En línea]
 Available at: <https://www.portalinmobiliario.com/h/blog/que-factores-determinan-el-precio-del-suelo-urbano/>

Quintana, O. G., 2018. *Factores que explican el valor del suelo urbano*, s.l.: s.n.

Rodrigo, J. A., 2016. *R Pubs - Regresión Lineal Múltiple en R*. [En línea]
 Available at: https://rpubs.com/Joaquin_AR/226291

Sanchez, A. D., 2017. *Economipedia.com*. [En línea]
 Available at: <https://economipedia.com/definiciones/indice-de-precios.html>

Sañudo, A., 2020. *ABCDMXYZ*. [En línea]
 Available at: <https://www.abcdm.xyz/verbetes/valor-del-suelo/#:~:text=El%20valor%20del%20suelo%20se,hacen%20m%20%20C3%A1s%20o%20menos%20atractivo.>
 [Último acceso: 03 10 2022].

Serrano, S. C., 2018. *Arquitecto*. [En línea]
 Available at: <https://scsarquitecto.cl/normas-urbanisticas-que-son/#:~:text=%E2%80%9CNormas%20urban%C3%ADsticas%E2%80%9D%3A%20todas%20aquellas,superficie%20de%20subdivisi%C3%B3n%20predial%20m%C3%ADnima%2C>

Suman, S., s.f. *Economics Discussion*. [En línea]
 Available at: <https://www.economicsdiscussion.net/price/price-index/price-index-meaning-uses-and-importance/12722>
 [Último acceso: 03 11 2022].

The Wall Street Journal, 2022. El boom de la vivienda se desvanece en todo el mundo mientras suben las tasas de interés. *La Tercera*, 19 julio.

Tinsa, 2021. *Tinsa Chile*. [En línea]
 Available at: <https://www.tinsa.cl/sala-de-prensa/notas-de-prensa/precio-vivienda-sube-y-tamano-se-reduciria/#:~:text=El%20precio%20promedio%20de%20la,hoy%20en%2020%20%20C8%20UF>

UNIR, 2022. *UNIR LA UNIVERSIDAD EN INTERNET*. [En línea]
 Available at: <https://www.unir.net/ingenieria/revista/lenguaje-r-big-data/#:~:text=R%20es%20un%20entorno%20de,a%20instrucciones%20en%20lenguaje%20m%C3%A1quina.>

Universidad de Harvard, 2021. *Como hacer referencias usando la notación Harvard (esto es solo un ejemplo)*. La Edición ed. La Ciudad: La Editorial.

Watson, J. H. S. y. M. M., 2012. *Introducción a la Econometría 3.ª edición*. En: Madrid: s.n., p. 600.

8. ANEXOS

8.1. ANEXO 1

cod mz	cod pr	fojas	num	año	Comprador	Rut comprad	Vendedor	Fecha	Precio UF	Precio \$	Precio Dólar	Fecha escrit	tipo transaccion
11535	52	91402	129206	2018	INMOBILIARIA MPC	76917378-1	EMPRESA CONS	04-12-2018	34925			28-09-2018	COMPRAVENTA
11535	52	55038	79309	2018	EMPRESA CONSTRUCC	92770000-k	NUÑEZ CANCIN	25-07-2018	29000			30-05-2018	COMPRAVENTA
11535	53	91402	129207	2018	INMOBILIARIA MPC	76917378-1	EMPRESA CONS	04-12-2018	43793			28-09-2018	COMPRAVENTA
11535	53	54277	78188	2018	EMPRESA CONSTRUCC	92770000-k	QUEZADA ALIST	20-07-2018	36000			30-05-2018	COMPRAVENTA
11535	54	91403	129208	2018	INMOBILIARIA MPC	76917378-1	EMPRESA CONS	04-12-2018	13703			28-09-2018	COMPRAVENTA
11535	54	55709	80290	2018	EMPRESA CONSTRUCC	92770000-k	MACURAN NOC	27-07-2018	11400			30-05-2018	COMPRAVENTA
11535	55	91403	129209	2018	INMOBILIARIA MPC	76917378-1	EMPRESA CONS	04-12-2018	5616			28-09-2018	COMPRAVENTA
11535	55	54285	78198	2018	EMPRESA CONSTRUCC	92770000-k	GALLARDO JAC	20-07-2018	4580			30-05-2018	COMPRAVENTA
11535	55	11174	15793	2018	GALLARDO JACOMET	8829289-8	UGARTE AGELL	12-02-2018		120.000.000		25-01-2018 / 31	COMPRAVENTA
11535	56	91404	129210	2018	INMOBILIARIA MPC	76917378-1	EMPRESA CONS	04-12-2018	17551			28-09-2018	COMPRAVENTA
11535	56	55063	79346	2018	EMPRESA CONSTRUCC	92770000k	VINEZ VINEZ SA	25-07-2018	14427,94			30-05-2018	COMPRAVENTA
11535	57	91404	129211	2018	INMOBILIARIA MPC	76917378-1	EMPRESA CONS	04-12-2018	17342			28-09-2018	COMPRAVENTA
11535	57	54277	78187	2018	EMPRESA CONSTRUCC	92770000-k	SCHEEL MAHN	20-07-2018	14427,94			30-05-2018	COMPRAVENTA
11535	58	91405	129212	2018	INMOBILIARIA MPC	76917378-1	EMPRESA CONS	04-12-2018	17341			28-09-2018	COMPRAVENTA
11535	58	54942	79177	2018	EMPRESA CONSTRUCC	92770000-k	WALTERS GAST	24-07-2018	14427,94			30-05-2018	COMPRAVENTA
11535	59	65335	94999	2019	INMOBILIARIA MPC	76917378-1	SOTO SALINAS	22-08-2019	14427,94			14-06-2019	COMPRAVENTA
11535	60	65863	95775	2019	INMOBILIARIA MPC	76917378-1	PEREZ JOSE LUI	23-08-2019	15582,18			14-06-2019	COMPRAVENTA

Ilustración 11: Captura de pantalla de los datos de transacción

Fuente: Elaboración propia

8.2. ANEXO 2

Año	Numero pi	valor_uni	sumatoria	año_trans	plazo_priv	n_roles	n_edificac	superficie	propietari	rut	residencia	hotel	equipam	comercio	oficina	cafeteria	educacion	actividad	infraestruc	s_edificac	s_util_por	util_vivier	s_edificac	s_totale
2021	33	59,752659	192582,82	2019	-1	10	1	11.586,99	INMOBILIARI	76917378-1	1	0	0	0	0	0	0	0	0	8.240,26	0,71	98,10	3.346,73	3.223,00
2021	30	91,9686138	210976	2019	0	3	1	10.167,45	INMOBILIARI	77014739-9	1	0	1	1	1	0	0	0	0	6.681,71	0,66	106,06	3.485,74	1.633,04
2021	11	87,4551808	142887,773	2017	-2	8	1	22.185,04	INMOBILIARI	77219250-9	0	0	1	1	1	1	0	0	0	16.546,20	0,75		5.638,84	1.633,84
2021	22	82,4125835	148013	2019	0	5	1	6.991,55	INMOBILIARI	77005646-2	1	0	0	0	0	0	0	0	0	4.832,44	0,69	105,05	2.159,11	1.796,00
2021	32	38,7181805	101938	2021	0	5	1	5.570,28	INMOBILIARI	77290472-k	1	0	0	0	0	0	0	0	0	4.308,86	0,77	119,69	1.261,42	2.632,82
2021	29	13,9956738	16500	2021	-12	2	1	6.116,71	INMOBILIARI	76036255-7	0	0	1	0	0	0	0	0	0	3.626,49	0,59		2.488,22	1.204,76
2021	3	34,1326735	47905	2020	0	3	1	3.239,48	INMOBILIARI	77130255-9	1	0	0	0	0	0	0	0	0	2.509,22	0,77	78,41	730,26	1.401,40
2021	28	35,6840043	12937,3211	2021	-2	2	1	1.060,52	INMOBILIARI	76993989-k	1	0	1	1	1	1	0	0	0	696,25	0,66	696,25	364,27	642,79
2021	26	137,977482	355868,762	2020	0	5	1	17.655,06	SEGUROS VID	99301000-6	0	0	1	1	1	1	0	0	0	13.936,04	0,79		3.719,02	2.579,18
2021	31	55,1229448	39898,5386	2020	0	2	1	1.975,07	INMOBILIARI	76297527-0	1	0	0	0	0	0	0	0	0	1.614,18	0,82	84,96	360,27	723,81
2021	25	88,8269696	185482,26	2019	0	2	1	10.805,30	PENTA VIDA	96812960-0	0	0	1	1	1	1	0	0	0	8.037,43	0,74		2.767,87	2.088,13
2021	21	74,1804858	146036,155	2020	-1	5	1	7.533,87	INMOBILIARI	77029064-3	1	0	0	0	0	0	0	0	0	5.344,66	0,71	98,98	2.189,21	1.968,66
2021	19	43,4971286	679855,77	2020	0	3	2	53.730,46	INÉS MATTE U	77113007-0	1	0	1	1	0	0	0	0	0	32.145,52	0,60	70,96	21.584,94	15.629,90
2021	18	108,658432	214382	2020	-1	4	1	15.375,84	INMOBILIARI	77041837-2	1	0	1	1	1	1	0	0	0	10.098,88	0,66	198,02	5.276,96	1.972,99
2021	17	14,4966684	74532,7361	2020	-1	2	2	21.129,23	INMOBILIARI	76116221-7	1	0	1	0	1	1	0	0	0	15.078,95	0,71	114,23	6.050,28	5.141,37
2021	16	30,3877406	87877,3953	2020	-1	4	1	14.724,35	ITAHUE RAICE	77006221-7	1	0	1	1	1	1	0	0	0	9.758,30	0,66	130,11	4.966,05	2.891,87
2021	10	166,59313	529186,41	2020	0	3	1	30.152,33	BICE VIDA CO	96656410-5	0	0	1	1	1	1	0	0	0	21.920,56	0,73		8.231,71	3.176,52
2021	9	71,7703924	101563	2020	0	4	1	5.237,04	INMOBILIARI	77606525-3	1	0	0	0	0	0	0	0	0	3.660,99	0,70	98,95	1.576,05	1.415,11
2021	8	145,039195	440000	2021	0	1	1	42.042,01	COMPANIA D	96571890-7	1	0	1	1	1	1	0	0	0	31.352,08	0,75		10.689,93	3.076,08
2021	7	70,0007508	130522	2019	0	5	1	8.424,32	INMOBILIARI	76960195-7	1	0	1	0	1	0	0	0	0	5.145,24	0,61	107,15	3.281,08	1.864,58
2021	6	58,4374344	77397,46	2020	0	2	1	3.908,57	INMOBILIARI	76903056-3	1	0	0	0	0	0	0	0	0	3.145,68	0,80	92,52	767,89	1.324,45
2021	2	45,1940559	33606,3	2020	0	2	1	1.968,45	INMOBILIARI	77139186-9	1	0	0	0	0	0	0	0	0	1.536,81	0,78	127,90	433,64	743,60
2020	15	29,4855709	23500	2019	0	1	1	1.423,43	INMOBILIARI	77042671-5	1	0	0	0	0	0	0	0	0	1.160,59	0,82	128,95	262,84	797,00
2020	22	70,9641428	169187,032	2020	0	6	1	10.287,92	INMOBILIARI	77114689-9	1	0	1	0	1	1	0	0	0	7.066,15	0,69	117,77	3.221,78	2.384,12
2020	21	69,1488752	105901,502	2020	0	3	1	5.092,34	INMOBILIARI	77081606-8	1	0	0	0	0	0	0	0	0	4.031,18	0,79	95,98	1.061,16	1.531,50
2020	4	71,7009546	110935	2019	0	3	1	5.751,09	INVERSIONES	96974830-4	1	0	0	0	0	0	0	0	0	4.104,60	0,71	102,62	1.646,49	1.547,19
2020	2	52,0501311	79200	2019	0	4	1	4.412,13	INMOBILIARI	76899795-0	1	0	1	1	1	0	0	0	0	3.591,75	0,81	92,10	820,38	1.521,61
2020	17	36,3384474	39571,479	2019	0	2	1	3.145,29	INMOBILIARI	77047713-1	1	0	0	0	0	0	0	0	0	2.578,55	0,82	92,09	566,74	1.088,97
2020	6	25,0273224	22900	2019	0	1	1	1.870,25	INMOBILIARI	77038447-8	1	0	0	0	0	0	0	0	0	1.524,73	0,82	117,29	345,52	915
2020	5	39,0348235	73622,8	2019	0	3	1	4.229,40	INMOBILIARI	76984337-k	1	0	0	0	0	0	0	0	0	3.038,51	0,72	72,35	1.190,89	1.886,08
2020	28	62,841521	106198,4	2018	0	2	1	10.818,23	INMOBILIARI	76754111-2	1	0	1	1	1	1	0	0	0	6.865,00	0,63	163,45	3.953,23	1.689,94
2020	25	60,0864779	667443	2017	0	28	4	49.358,14	INMOBILIARI	76701870-3	1	0	1	1	1	1	1	0	0	36.662,40	0,74	123,86	12.695,74	11.108,04
2020	24	95,7611958	222364,2	2020	-1	4	1	11.526,21	INMOBILIARI	77054097-6	1	0	1	1	1	1	0	0	0	8.323,54	0,72	132,12	3.202,67	2.322,07
2020	23	53,0038797	161872,434	2020	-1	7	1	11.450,51	INMOBILIARI	76092183-1	1	0	1	0	0	0	0	0	0	8.138,82	0,71	103,02	3.311,69	3.050,20
2020	20	68,084499	81090	2019	0	2	1	5.355,29	INMOBILIARI	77020019-9	1	0	1	1	1	1	0	0	0	3.741,51	0,70	120,69	1.613,78	1.191,02
2020	19	59	106014,84	2019	0	8	1	13.184,20	LA COLMENA	77029531-9	0	0	1	1	1	1	0	0	0					

coef_de_ci	altura_ma	rasantes	coef_ocup	densidad	adosamie	antejardir	estaciona	DFLN2	conjunto	copropied	beneficio	proyeccia	conj_viv	otros	anteproye	viviendas	locales_ct	numero_h	n_oficina	estaciona	estaciona	distrito_ci	zona_cens	
2,08	7 70°		0,2	84	no	40%	5	93	1	0	1	1	1	0	0	1	84	0	0	0	121	47	4_4_12	
2,34	12 70°		0,28	63,00			5	74	1	0	1	1	1	0	0	1	63	3	0	0	78	42	2_2_6	
9,1	22 70°		1,0	0	EC 3P		0	249	0	0	1	1	1	0	0	1	0	4	0	214	249	126	3_3_6	
2,21	12 70°		0,20	49	0		8,48	79	1	0	1	1	1	0	0	1	46	0	0	0	58	29	4_4_1	
1,36	3 70°		0,35	36	40%		5	40	1	0	1	0	0	1	0	1	36	0	0	0	53	40	5_5_2	
2,63	7 70°		0,60	0	40%		6,32	85	0	0	1	1	1	0	0	1	0	0	0	92	85	17	2_2_5	
1,44	4 70°		0,35	32	ART 6.1.8		3	35	1	0	1	0	0	1	0	1	32	0	0	0	35	14	1_1_4	
1,43	3 70°		0,60	15	OGUC		3	1	0	0	1	0	0	1	1	1	1	19	0	0	9	1	11_5	
5,33	14 70°		0,60	0	40%		5	174	0	0	1	1	1	0	0	0	0	8	0	13	125	155	8_8_2	
2,08	7 70°		0,20	19	0		5	21	1	0	1	1	1	1	0	0	1	19	0	0	0	24	9	2_2_7
3,37	10 70°		0,53	0	0		3	122	0	0	1	1	1	0	0	0	0	10	0	171	99	62	4_4_7	
2,21	12 70°		0,2	54	0		7,52	60	1	0	1	1	0	0	0	0	1	54	0	0	0	67	32	2_2_2
2,34	5 70°		0,6	453	0		3	570	1	0	1	1	1	0	1	1	453	3	0	0	572	333	7_7_1	
4,77	17 70°		0,60	51	40%		5	119	1	0	1	1	1	0	1	1	51	4	0	136	148	82	8_8_4	
2,60	9 70°		0,40	134	0		4,6	146	1	0	1	1	1	0	0	1	132	0	0	8	213	120	4_4_11	
3,12	10 70°		0,6	75	40%		4	114	1	1	1	0	1	0	0	0	75	6	0	17	114	60	2_2_7	
5,42	19 70°		0,60	0	40%		7,2	425	0	0	1	1	1	0	1	1	0	6	0	21	298	127	8_8_1	
2,08	7 70°		0,2	37	0		5,84	61	1	0	1	1	1	0	0	1	0	37	0	0	54	37	4_4_12	
9,10	19 70°		1,00	0	0		0	484	0	0	1	0	1	0	0	1	0	8	0	113	484	153	8_8_3	
2,08	7 70°		0,20	48	0		5	80	1	0	1	1	1	0	0	1	48	0	0	16	90	32	4_4_1	
2,08	7 70°		0,20	34	no		6	37	1	0	1	1	1	0	0	1	34	0	0	0	37	19	2_2_7	
2,08	7 70°		0,2	19	0		5,86	22	1	0	1	1	1	0	0	0	12	0	0	0	13	16	4_4_4	
1,10	3 70°		0,36	11	0		5	17	1	0	1	0	0	1	0	1	9	0	0	0	17	0	7_7_2	
2,08	7 70°		0,20	62	0		6	116	1	0	1	1	1	0	0	1	60	0	0	24	111	39	2_2_3	
2,21	12 70°		0,20	42	0		6,86	69	1	0	1	1	1	0	0	1	42	0	0	0	47	50	2_2_3	
2,08	7 70°		0,20	40	0		6,025	67	1	0	1	1	1	0	0	1	40	0	0	0	62	23	4_4_9	
2,08	7 70°		0,2	40	0		6,06	69	1	0	1	1	1	0	0	1	39	1	0	0	47	26	4_4_4	
2,08	7 70°		0,20	28	0		5,95	31	1	0	1	1	1	0	0	1	28	0	0	0	31	17	4_4_12	
0,7	3 70°		0,4	13	0		5	23	1	0	1	0	0	1	0	1	13	0	0	0	18	9	1_1_6	
1,31	4 70°		0,28	42	0		5	62	1	0	1	1	1	1	0	0	42	0	0	0	47	16	1_1_2	
2,81	12 70°		0,60	46	40%		6,82	96	1	0	1	1	1	0	0	1	42	3	0	78	148	38	2_2_3	
2,158	10 70°		0,20	299	0		10,40	477	1	0	1	1	1	0	0	1	296	10	0	144	534	193	4_4_1	
2,21	12 70°		0,60	64	40%		7,87	66	1	0	1	1	1	0	0	1	63	1	0	56	110	52	2_2_5	
2,08	7 70°		0,20	79	0		6,12	87	1	0	1	1	1	0	0	1	79	0	0	0	107	43	1_1_3	
3,64	10 70°		0,60	31	ART. 2.6.3		4	53	1	0	1	1	1	0	0	1	31	5	0	27	58	27	4_4_4	
4,37	12 70°		0,60	0	40%		8,24	149	0	0	1	1	1	0	0	0	0	7	0	164	149	52	8_8_2	
1,36	3 70°		0,40	23	0		5	38	1	0	1	1	0	1	0	1	20	0	0	0	28	16	1_1_5	
1,43	3 70°		0,6	36	0		3	74	1	0	1	0	0	1	1	1	32	7	0	13	35	38	1_1_5	
2,06	5 70°		0,6	0	40%		5	67	0	0	1	1	1	0	0	1	0	7	0	64	113	0	1_1_2	
2,18	10 70°		0,20	47	0		8,23	52	1	0	1	1	1	0	0	1	47	0	0	0	71	38	4_4_1	

Ilustración 12: Captura de pantalla base de datos

Fuente: Elaboración propia

8.3. ANEXO 3

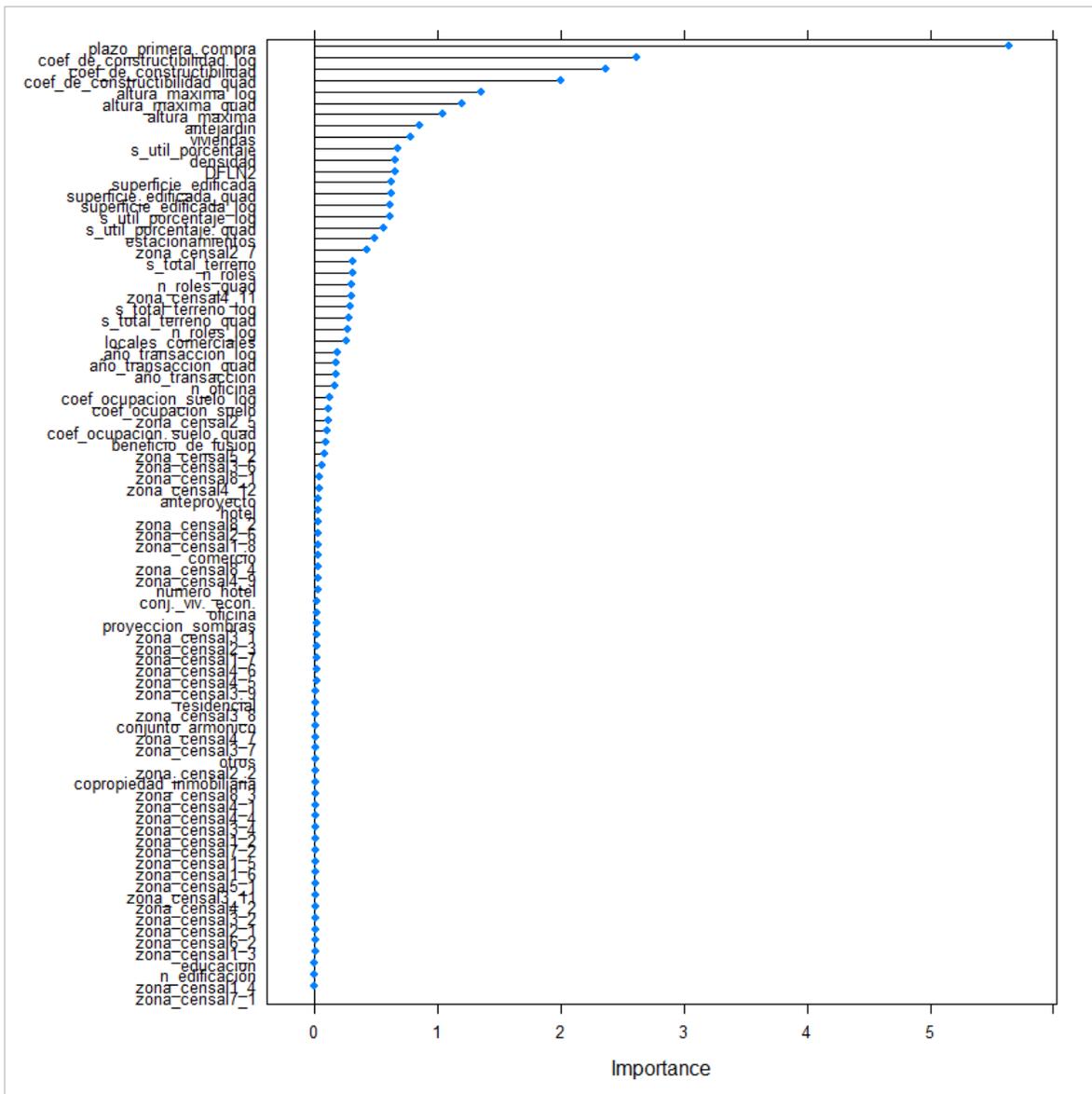


Ilustración 13: Gráfico importancia de las variables

Fuente: Elaboración propia

8.4. ANEXO 4

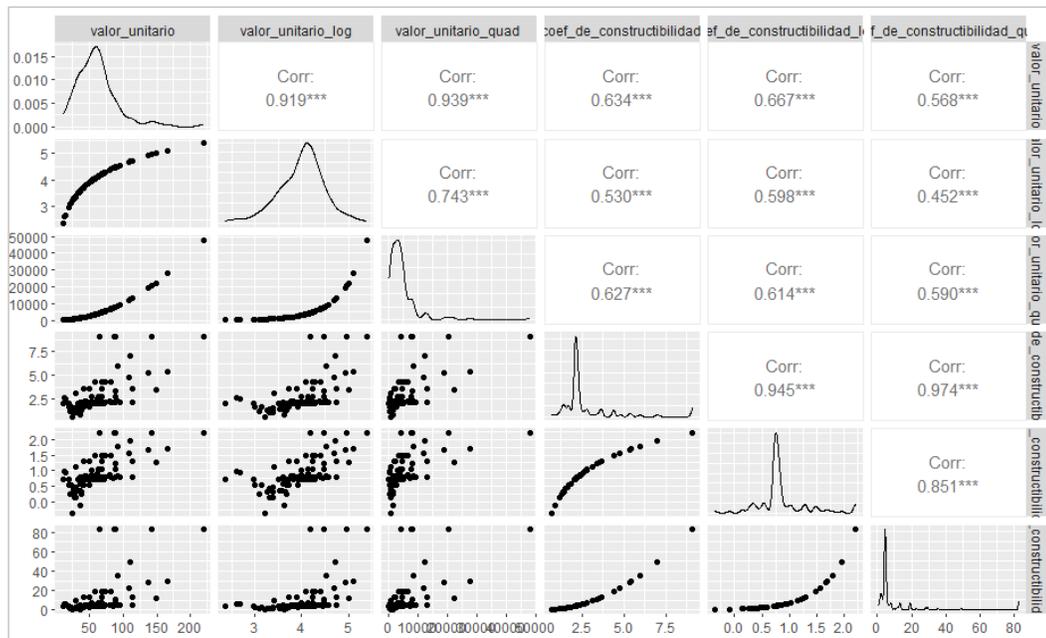


Ilustración 14: Matriz de correlación valor unitario vs coeficiente de constructibilidad

Fuente: Elaboración propia

8.5. ANEXO 5

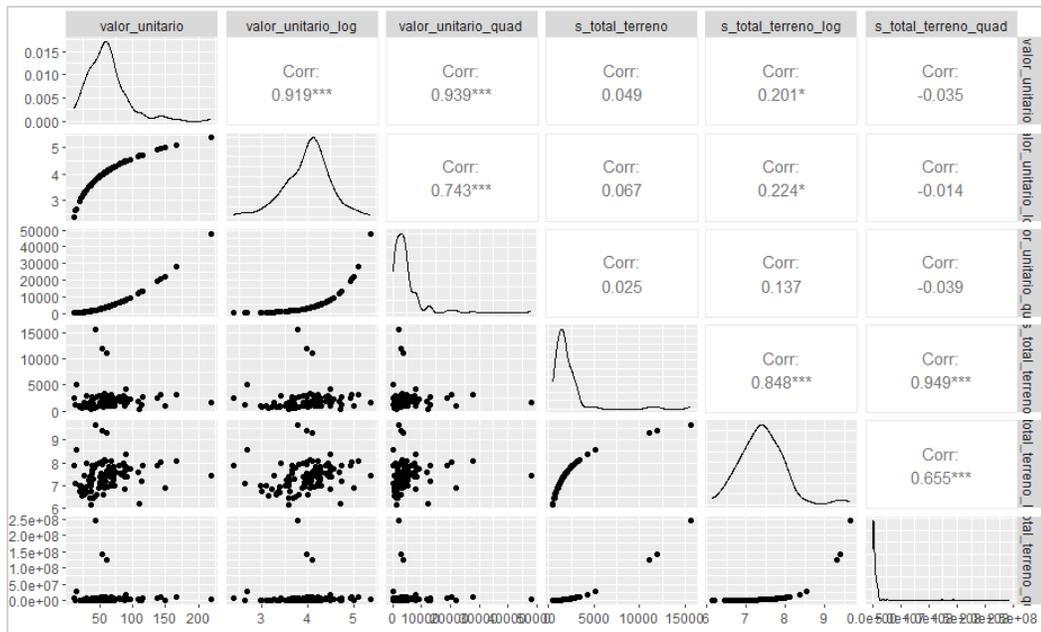


Ilustración 15: Matriz de correlación valor unitario vs superficie del terreno

Fuente: Elaboración propia

8.6. ANEXO 6

```

Residuals:
  Min       1Q   Median       3Q      Max
-0.57429 -0.15079  0.01016  0.16232  0.67504

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.597e+02  3.652e+01 -4.373 3.11e-05 ***
coef_de_constructibilidad_log  4.754e-01  1.176e-01  4.041 0.000107 ***
plazo_primera_compra          1.904e-01  2.023e-02  9.411 2.74e-15 ***
altura_maxima                 3.083e-02  1.069e-02  2.885 0.004836 **
n_roles                       1.707e-02  9.118e-03  1.872 0.064181 .
s_util_porcentaje_log         1.854e-01  6.647e-02  2.790 0.006363 **
densidad                     1.350e-03  7.942e-04  1.699 0.092494 .
estacionamientos            -2.073e-03  6.188e-04 -3.350 0.001158 **
DFLN2                       -3.575e-01  1.087e-01 -3.290 0.001400 **
año_transaccion              8.010e-02  1.816e-02  4.411 2.69e-05 ***
n_oficina                    -1.041e-03  8.430e-04 -1.235 0.219898
zona_censal_2_7              -2.340e-01  9.457e-02 -2.474 0.015099 *
zona_censal_4_11             -5.112e-01  1.573e-01 -3.249 0.001595 **
coef_ocupacion_suelo_log     -1.523e-01  6.759e-02 -2.254 0.026494 *
zona_censal_2_5              4.462e-01  1.531e-01  2.915 0.004421 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2595 on 96 degrees of freedom
Multiple R-squared:  0.7735,    Adjusted R-squared:  0.7405
F-statistic: 23.42 on 14 and 96 DF,  p-value: < 2.2e-16

```

Ilustración 16: Regresión lineal sin superficie total del terreno

Fuente: Elaboración propia