

Estimación del impacto generado en el valor comercial de los inmuebles residenciales sometidos a RPH, localizados entre las calles 34 y 42, y las carreras 63 y 65, por la distancia existente entre estos y parques del río.

Propuesta de trabajo de grado desarrollada como requisito de graduación de la Especialización en Valoración Inmobiliaria, La Lonja-ESUMER

Por: Santiago Álvarez Galeano

Asesores: Carolina Quintero Medina – Alexander Zuluaga Correa

Contenido

1	Marco conceptual	5
1.1	Modelo de Precios Hedónicos y Regresión Lineal Multivariada.....	6
2	Planteamiento del problema	8
3	Pregunta de investigación.....	9
4	Objetivo	9
4.1	Objetivos Específicos.....	9
5	Justificación.....	10
6	Área de estudio	10
7	Metodología.....	13
7.1	Análisis Estadístico	14
8	Resultados	14
8.1	Procedimiento estadístico y tratamiento de la información	15
8.2	Análisis exploratorio	17
8.3	Postulación de modelos.....	23
8.4	Cumplimiento de supuestos	25
9	Discusión.....	31
10	Conclusiones	34
	Agradecimientos	36
	Referencias.....	37
	Anexos	40

Listado de Tablas

Tabla 1. Definición de variables cualitativas y cuantitativas para el estudio.	16
Tabla 2. Resumen del comportamiento de los datos de las variables: promedio, máximos y mínimos.	17
Tabla 3. Prueba <i>t</i> de significancia individual para variables cualitativas	22
Tabla 4. Prueba <i>f</i> de significancia para variables cualitativas	23
Tabla 5. Prueba <i>f</i> de significancia para variables cualitativas en presencia de las variables cuantitativas.	23
Tabla 6. Resultado metodología StepWise	24
Tabla 7. Pruebas de multicolinealidad para el modelo 1.	26
Tabla 8. Pruebas de multicolinealidad para el modelo 2.	27
Tabla 9. Prueba de normalidad sobre los errores para el modelo 2.	27
Tabla 10. Pruebas de correlación del valor respecto las variables explicativas	30

Listado de Figuras

Figura 1. Área de estudio.	11
Figura 2. Distribución de las muestras en el área de estudio.	15
Figura 3. Box-Plot de valor del inmueble para la variables cualitativas.	18
Figura 4. Matriz de Correlación entre variables cuantitativas.	20
Figura 5. Gráfico de dispersión.	21
Figura 6. Matriz de correlación entre variables del modelo 1.	25
Figura 7. Matriz de correlación entre variables para el modelo 2.	26
Figura 8. Gráfico Q-Q de prueba de normalidad para los errores en el modelo y diagnóstico de los residuos para el modelo 2.	28
Figura 9. Dispersión de los datos en función del valor para las variables DpieM y Área.	30

Anexos

Anexo 1. Pruebas individuales con estadístico t	40
Anexo 2. Pruebas f con variables factor	42
Anexo 3. Pruebas f en presencia de todas las variables.....	43
Anexo 4. Análisis R Modelo 1	44
Anexo 5. Análisis R Modelo 2	44
Anexo 6. Código R.....	45

1 Marco conceptual

La consolidación de zonas residenciales son altamente influenciadas por las necesidades que resuelve el entorno a las personas que buscan ubicarse en un determinado sector. El acceso a fuentes de transporte, vías de conectividad, sector comercial y equipamientos públicos como iglesias y parques recreativos afectan la forma como se desarrollan los lugares de residencia, por lo tanto, determinan en alguna medida, la dinámica del mercado inmobiliario. Es así como el precio de las propiedades se verá afectado en una proporción por factores intrínsecos, al igual que por los elementos del entorno. (Geltner et al. 2013).

Actualmente, esta bien documentado el efecto que tiene las particularidades propias de un inmueble al igual que sus externalidades sobre el valor que adquiere en el mercado. (Pagourtzi et al. 2003). Análisis espaciales relacionados con el entorno y cercanía a equipamientos son ampliamente conocidos. (Crompton, 2001). También, en otros países se han presentado publicaciones que ilustran el efecto de la distancia sobre el valor de la propiedades circundantes a parques o espacios naturales. En su trabajo, Tekel et. al (2013) evidencian el aumento del valor de la propiedades mas cercanas a un parque natural en Ankara, Turquía. Los hallazgos de Hendon (1971) evidencian el efecto de los equipamientos en los parques sobre el valor de las propiedades circundantes. A si mismo, Votsis (2017), analiza a través de modelos hedónicos el valor que adquieren las propiedades en función de la cercanía a los parques urbanos en la Helsinki, Finlandia.

Por su parte, en Colombia, estos trabajos han sido mejor documentados con relación a terminales de transporte urbano: Rodríguez y Targa (2004) definen que un incremento de cinco minutos en la distancia al sistema de transporte masivo, disminuye entre 6,8% y 9,3% en el precio de inmuebles residenciales. Perdomo (2010) determina, a través de un modelo hedónico espacial, que un predio ubicado a 2.500 metros del sistema de transporte TransMilenio en Bogotá, tiene un valor

promedio (por metro cuadrado) inferior en \$117.500 pesos aproximadamente, comparado con otro ubicado a pocos metros de la estación de transporte TransMilenio. También, Muñoz (2010) presenta el valor de las propiedades en función de la proximidad peatonal de diez y cinco minutos a las estaciones de transporte TransMilenio de Bogotá a través de un modelo hedónico. A su vez, Quintero et. al (2012) analizan el efecto de la distancia al acceso de transporte masivo Megabus sobre el valor de la propiedad en la zona Cuba, Periera.

Las publicaciones acerca del efecto que tiene la distancia sobre el valor de los inmuebles entorno a los parques y zonas verdes de uso público en el municipio de Medellín es aún incipiente. La información suministrada por este trabajo permitirá ilustrar con mayor detalle a las entidades públicas y privadas los beneficios que obtienen los particulares derivados de los programas de obras de infraestructuras, al igual que aportará mayor conocimiento sobre el comportamiento inmobiliario y cómo es afectado por este tipo de obras en la ciudad.

1.1 Modelo de Precios Hedónicos y Regresión Lineal Multivariada

Para determinar la influencia que ejercen las áreas de esparcimiento público y ambientes naturales sobre el valor de inmuebles en zonas residenciales urbanas, el modelo de precios hedónicos es comúnmente usado. Esta metodología es utilizada para la estimación de la influencia de los factores sobre el precio de un producto. El modelo de precios hedónicos permite medir la influencia particular de un atributo dentro de una estructura compuesta de los mismos para un producto como un inmueble residencial. Hedónico hace referencia a “satisfacción” y el término en el marco del sector inmobiliario se refiere a la apreciación, disfrute, satisfacción de los servicios (características) que provee un inmueble (Tekel, 2013). Los modelos de precio hedónico parten de la base de que un producto, en este caso un inmueble, esta determinado por un grupo de atributos que son percibidos por unos compradores

(mercado) por el cual están dispuestos a pagar un valor determinado luego de hacer múltiples comparaciones entre las opciones disponibles (Kaul S, 2006). Dentro de estos atributos existen elementos extrínsecos como la distancia a parques naturales, parques recreativos, zonas comerciales, centros comerciales, estaciones de transporte, campos deportivos, entre otros. Para el análisis del peso que puede tener cierto atributo en el precio se utilizan varias metodologías econométricas, entre ellas la regresión lineal.

En este estudio se define una aproximación de la función lineal para el análisis de los diferentes atributos que pueden estar explicando el valor de un inmueble. En la ecuación de regresión lineal multivariada (RLM) se presentan la variable dependiente, el valor del inmueble y los factores (variables explicativas) que pueden estar explicando el valor de los inmuebles. La RLM trata del estudio de la dependencia de una variable (dependiente) respecto de dos o más variables (explicativas) con el objetivo de estimar el valor promedio poblacional de la primera en términos de los valores conocidos o fijos (en muestras repetidas) de las segundas. (Gujarati y Poter, 2010).

A continuación, se describe la ecuación del modelo teórico.

$$y_i = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} + \varepsilon_i$$

con $\varepsilon_i \sim iid N(0, \sigma^2)$ Los errores se distribuyen idéntica e independientemente normales con media igual a 0 y varianza σ^2 .

Donde:

y_i = Variable respuesta en la observación i

β_0 = Intercepto de la recta de la ecuación lineal

β_j = Coeficiente de la i -ésima observación de la variable x_j

x_{ij} = i -ésima observación de la variable x_j

ε_i = Error del modelo en la i -ésima observación

2 Planteamiento del problema

Los equipamientos del territorio urbano, son factores externos que influyen en el valor comercial que adquieren los inmuebles ubicados en el área de influencia. (Hendon, 1971). En ese sentido, estaciones de transporte, centros comerciales y parques públicos afectan el valor comercial de inmuebles que los circundan. En consecuencia se asume que Parques del Río costado occidental puede estar influenciando el valor comercial que adquieren los predios con uso residencial y sometidos al Régimen de propiedad horizontal en el barrio Conquistadores en alguna proporción que es susceptible de ser estudiada. Por esto, la ubicación de los inmuebles de conquistadores en referencia a Parques del Río, cómo uno de estos factores externos, puede estar impactando el valor que adquieren los mismos en el mercado inmobiliario de la zona.

Analizar el efecto de la distancia de apartamentos en el área directa de influencia del barrio Conquistadores con relación a Parques del Río costado Occidental, puede aportar información importante para la comprensión de estos fenómenos de valoración local y amplían el conocimiento del impacto sobre el valor de inmuebles cercanos a zonas de equipamientos públicos e infraestructuras urbanas

3 Pregunta de investigación

¿Existe algún efecto en el valor comercial de los apartamentos sometidos a Régimen de Propiedad Horizontal localizados entre las calles 34 y 42, y las carreras 63 y 65 en el barrio Conquistadores, generado por la distancia a pie de los mismos a Parques del Río, costado occidental?

4 Objetivo

Establecer el nivel de correlación entre la distancia respecto a Parques del Río costado occidental, y el valor comercial de los inmuebles sometidos a Régimen de Propiedad Horizontal con uso residencial localizados en el barrio Conquistadores de Medellín, entre las calles 34 y 42, y las carreras 63 y 65.

4.1 Objetivos Específicos

- Establecer la correlación que existe entre la distancia a Parques del Río, costado occidental y el valor comercial de los inmuebles sometidos a Régimen de Propiedad Horizontal con uso residencial localizados en el barrio Conquistadores, entre la calles 34 y 42, y las carreras 63 y 65 en la ciudad de Medellín.
- Determinar el modelo de regresión lineal que explica el comportamiento del valor en función de la distancia a pie.

5 Justificación

El conocimiento acerca de los efectos que genera los equipamientos de esparcimiento social urbano, como parques públicos, sobre el valor comercial de los inmuebles en la ciudad de Medellín es mínimo. Por tanto, ampliar la ilustración sobre factores que pueden estar influenciando el valor de los inmuebles en zonas residenciales representa un elemento de importancia que aporta en el conocimiento integral de la valoración y desarrollo de proyectos inmobiliarios. Adicionalmente, permite mayor ilustración para entidades públicas sobre los impactos derivados de los programas de obras de infraestructura, a su vez que aumenta la comprensión para los profesionales que se forman en esta temática.

Parques del Río, etapa 1A, representa la primera iniciativa de equipamientos lineales en el eje central de la ciudad en el trayecto del cuerpo de agua principal que atraviesa Medellín. Por lo tanto, la comprensión del impacto sobre el valor comercial de las residencias en su área de influencia es de gran importancia para el futuro de proyectos similares a nivel local y regional.

6 Área de estudio

La zona de estudio se ubica entre la carrera 63c (quebrada La Picacha) y la calle 34, de norte a sur, y la carrera 63 y la carrera 65 de oriente a occidente, en el barrio conquistadores del municipio de Medellín. El punto de referencia concéntrico esta determinado por las coordenadas $75^{\circ}34'26.762''W - 6^{\circ}14'30.659''N$. con una extensión de 36,022 hectáreas (Acuerdo 48, 2014) (Figura 1). Se determina por el polígono de tratamiento urbanístico Z4_CN1_12 (Acuerdo 48 de 2014) ubicado en la zona centro occidente de Medellín, con tratamiento de Consolidación de nivel 1 que “corresponde a sectores del suelo urbano que presentan una morfología predial y

urbana, ocupación y densificación acorde con el modelo de ocupación definido en el Plan y con una dotación de espacio público, equipamientos e infraestructura vial y de servicios públicos adecuada”. El polígono de tratamiento se ajusta a los máximos aprovechamientos de un índice de construcción (I.C) de 3.4, densidad de 350 (Viv/ha) sin límite de altura. Complementariamente, la normatividad actual permite aprovechamientos adicionales por ventas de derechos de índice de construcción (I.C) de 0.6 y densidad de 50 (Viv/ha).

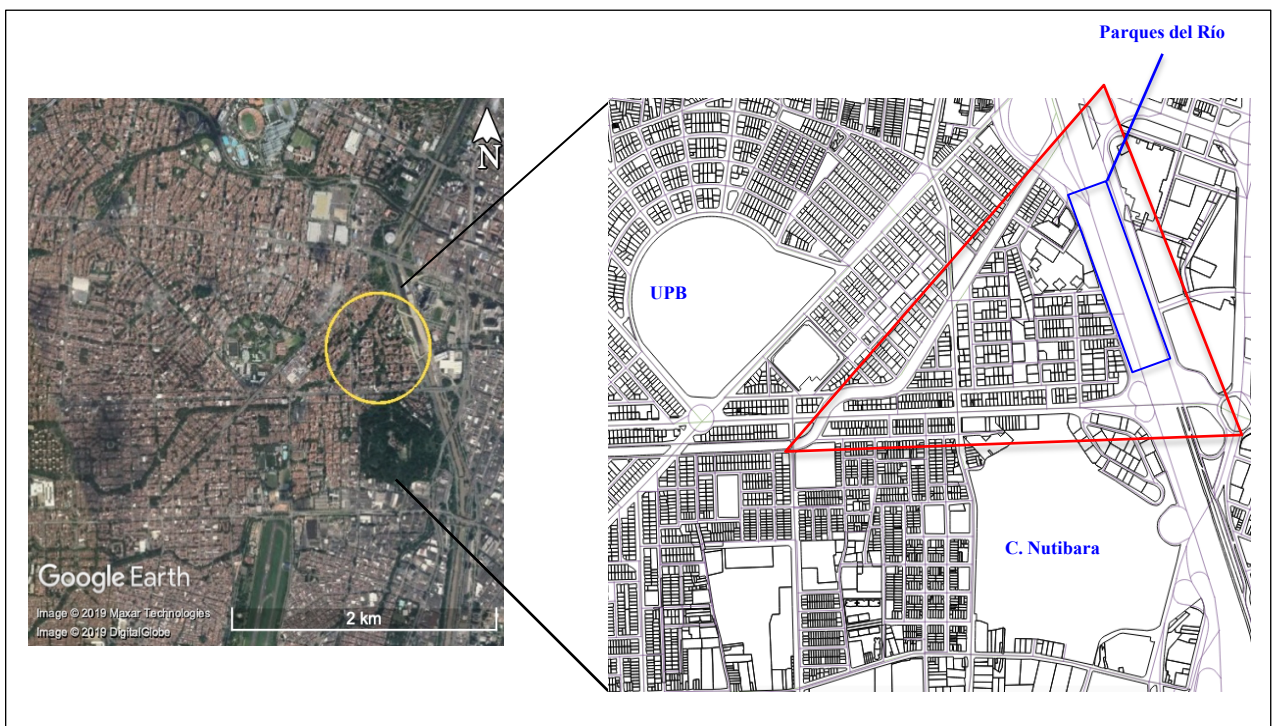


Figura 1. Área de estudio.

Ubicación general del barrio conquistadores respecto a la comuna 11 y localización del área de muestreo y el área de influencia de Parques del Río.

La zona se caracteriza con equipamientos como Parques de Río, Cerro Nutibara, La universidad Bolivariana, el Centro Comercial Unicentro y La iglesia del Verbo Divino. Dentro de las vías de acceso, se destacan: calle 33, la carrera 66B (Avenida

Bolivariana) y las Carreras 65 y 63 (Autopista sur). Los Barrios circundantes son Laureles, San Joaquín, Naranjal y Fátima.

Desde su diseño, Conquistadores se caracterizó históricamente por ser un barrio residencial de casas unifamiliares y amplias zonas verdes desde los años 50s. El Barrio está delimitado por bordes marcados, sobre condiciones de topografía plana. Presenta un plano urbano de vías discontinuas hacia el interior, con ejes de zonas naturales de uso público, con registro visual corto y restricciones de altura en sus edificaciones determinadas principalmente por el cono de aproximación al aeropuerto Olaya Herrera (Gómez y Wolf, 2008). Los ejes principales están conformados por la canalización La Picacha que se define como un trazado central conformado por una franja verde arborizada. Vías de acceso y regulación vehicular como la calle 34 y las carreras 64b y 63b (figura 1). Los otros dos ejes son áreas verdes que fluyen alternados entre zonas residenciales, parques públicos que estructuran la definición del barrio.

El Cerro Nutibara, el río Medellín, la quebrada La Picacha y el trazado de los parques lineales centrales y Parques del Río definen la forma natural del barrio.

En la actualidad, Conquistadores se identifica por la mixtura de una traza urbana donde predomina el uso residencial compuesto por casas unifamiliares, locales comerciales y edificaciones en propiedad horizontal, principalmente en estrato 5 y 6. Las edificaciones presentan en su mayoría sistemas de construcción estructural aporticado, con buenos acabados en las construcciones recientes. La zona cuenta con un mercado de inmuebles dinámico para la venta y alquiler de casas, locales, y apartamentos nuevos y usados. Así mismo, al ser una zona consolidada, de buenas condiciones urbanísticas, es una plaza atractiva para el desarrollo constante de nuevos proyectos constructivos.

Respecto a Parques del Río, en la ciudad de Medellín, se describe como una iniciativa integral de transformación urbana que une al municipio a través de las dos márgenes del río en el sector central de la Ciudad. Esta obra, de connotación lineal, inicia su construcción en el año 2014 y finaliza en el 2016. Contempla en su diseño el mejoramiento de la infraestructura, el paisajismo y la vegetación, con el objetivo de convertirse en un corredor de movilidad metropolitana, principal eje ambiental y de espacio público para favorecer el encuentro de los ciudadanos. Actualmente se cuenta con la etapa 1A finalizada que abarca la zona entre la calle 44 y la calle 33 en el costado occidental de la margen del río Medellín. Sus áreas de influencia son principalmente la comuna 11 del municipio de Medellín, en especial el Barrio Conquistadores.

7 Metodología

Se realizó un muestreo de ofertas de apartamentos en venta, con recorridos de campo y consulta de plataformas de ofertas de inmuebles en el barrio conquistadores delimitado por el área de influencia en el polígono residencial ubicado entre: la carrera 63c (quebrada La Picacha) y la calle 34, de norte a sur, y la Carrera 65 y la carrera 63. El muestreo se delimitó a un momento puntual: noviembre de 2018.

A continuación, se realizó la identificación de la ubicación satelital de las ofertas con el equipo de georreferenciación Garmin eTREX 20x, confirmado con la capa de tratamientos urbanísticos georreferenciados del municipio de Medellín en el programa QGIS. La estimación de distancias se determinó con ayuda del programa QGIS y Google Earth Pro.

Se seleccionaron los inmuebles que estaban en condiciones de equivalencia residencial, por lo que se delimitó el área de estudio a la zona que no estuviera

influenciada por vías comerciales y áreas de alto ruido vehicular (calle 34, carrera 65, carrera 63c y la entrada y salida a Parques del Río).

7.1 Análisis Estadístico

Para el tratamiento de la información se utiliza el programa Microsoft Excel. Posteriormente, se usó el software de programación y estadística R para realizar el análisis exploratorio, análisis de correlación, definición de modelos de regresión lineal y prueba de errores y supuestos. Para cada regresión se probaron los supuestos de homogeneidad y normalidad basados en los análisis de los errores. Para determinar si existía una relación entre el valor de los inmuebles y la distancia a Parques del Río, se aplicó el modelo de precios hedónicos a través de la definición de variables *dummy*, la regresión lineal multivariada (RLM), el coeficiente de correlación de Pearson, pruebas de significancia estadística *t* de Student y *f* de Fisher. Todas las pruebas se consideran significativas para un nivel $p < 0.05$.

8 Resultados

Un total de 38 muestras fueron obtenidas durante el muestreo (figura 2). 17 variables fueron documentadas de las cuales diez son variables cuantitativas y siete son cualitativas. La definición de variables se presenta en la tabla 1. Se utilizaron variables *dummy** para incluir en el modelo características como la presencia de ascensor, cuarto útil, habitación del servicio y registro de paisaje. En la tabla 2 se visualiza el resumen del comportamiento de la muestra de datos para las variables analizadas. Una vez definidas las variables se realizó una depuración y recodificación

* Variable *dummy* indica la presencia o ausencia de una cualidad o atributo. Estas son variables que adquieren valor de 1 en una submuestra y 0 en el resto de la muestra.

de la información para evitar errores en el análisis con el software R de análisis estadístico.



Figura 2. Distribución de las muestras en el área de estudio. Los puntos en rojo presentan la ubicación de los inmuebles que fueron muestreados. En el costado derecho se visualiza la ubicación de Parques del Río con la línea media (en rojo) de referencia determinada para la estimación de las distancias.

8.1 Procedimiento estadístico y tratamiento de la información

El proceso del análisis de información está separado en varias etapas. Inicialmente se realiza una definición y descripción de las variables utilizadas. Posteriormente se realiza el análisis exploratorio que permite inferir el uso de algunas variables, continuando con la definición de modelos con las variables seleccionadas y finalización con la prueba de supuestos del modelo que mejor explica el comportamiento del valor de los inmuebles muestreados. En la tabla 1 y 2 se presentan la descripción y definición de las variables.

Tabla 1. Definición de variables cualitativas y cuantitativas para el estudio.

Variable	Definición	Clasificación	Unidad	Observaciones
Valor de oferta	Valor	Cuantitativa	pesos (\$)	
Área Construida del inmueble	Área	Cuantitativa	m ²	
Costo de administración	Administración	Cuantitativa	pesos (\$)	
Nivel desde la calle	Altura	Cuantitativa	-	Definición del nivel desde la calle en su unidad.
Edad de construcción del inmueble	Edad	Cuantitativa	años	
Número de Balcones	Balcón	Cuantitativa	-	De 0 (sin balcones) a 3 (con tres Balcones)
Número de Parqueaderos vehiculares	Parqueadero	Cuantitativa	-	De 0 (sin parqueadero) a 2 (con dos parqueaderos)
Distancia de recorrido a pie media a Parques del Río	D. PieM	Cuantitativa	m	Definición de una línea media a lo largo del parque
Distancia de recorrido a pie puntual a Parques del Río	D. PieP	Cuantitativa	m	Definición de un punto central en el parque
Distancia en línea recta media a Parques del Río	DistanciaM	Cuantitativa	m	Definición de una línea media a lo largo del parque
Distancia en línea recta puntual a Parques del Río	DistanciaP	Cuantitativa	m	Definición de un punto central en el parque
Presencia de Cuarto Útil	C.U.	Cualitativa		0 en ausencia, 1 en presencia
Habitación del servicio	H. Servicio	Cualitativa		0 en ausencia, 1 en presencia
Presencia de ascensor en el edificio	Ascensor	Cualitativa		0 en ausencia, 1 en presencia
Presencia de vigilancia en portería	Portería	Cualitativa		0 en ausencia, 1 portería diurna y 2 las 24horas del día.
Registro de paisaje general	Paisaje	Cualitativa		El registro al paisaje se definió en aquellos inmuebles que registran su visual a parques públicos, paisaje de la ciudad y del Área Metropolitana. 0 en ausencia, 1 en presencia
Registro de visual a Parques del Río	Paisaje P.	Cualitativa		El registro a Parques se definió en aquellos inmuebles que su visual registra únicamente a parques del Río. 0 en ausencia, 1 en presencia

Tabla 2. Resumen del comportamiento de los datos de las variables: promedio, máximos y mínimos.

Variable	Promedio	Máximo	Mínimo
Valor	444605263	590000000	320000000
Valor/m ²	3709172	8250000	1822917
Área	130	204	74
Adm.	403925	708080	100000
Altura	5	13	1
Edad	22	40	1
Balcón	1	3	0
Parq.	1	2	0
D. PicM	274	498	32
D. PicP	364	661	85
DistanciaM	233	456	34
DistanciaP	281	460	92
C.U.	1	1	0
H. Servicio	1	1	0
Ascensor	1	1	0
Portería	1	2	0
Paisaje	0	1	0
Paisaje P.	0	1	0

8.2 Análisis exploratorio

Se presenta el resultado del análisis descriptivo y exploratorio entre variables cualitativas y cuantitativas con el fin de determinar cuáles de estas podrían estar incidiendo en el valor. Este proceso se condujo a través de matrices de correlación, gráficos Box-Plot y de dispersión presentadas en la figura 3.

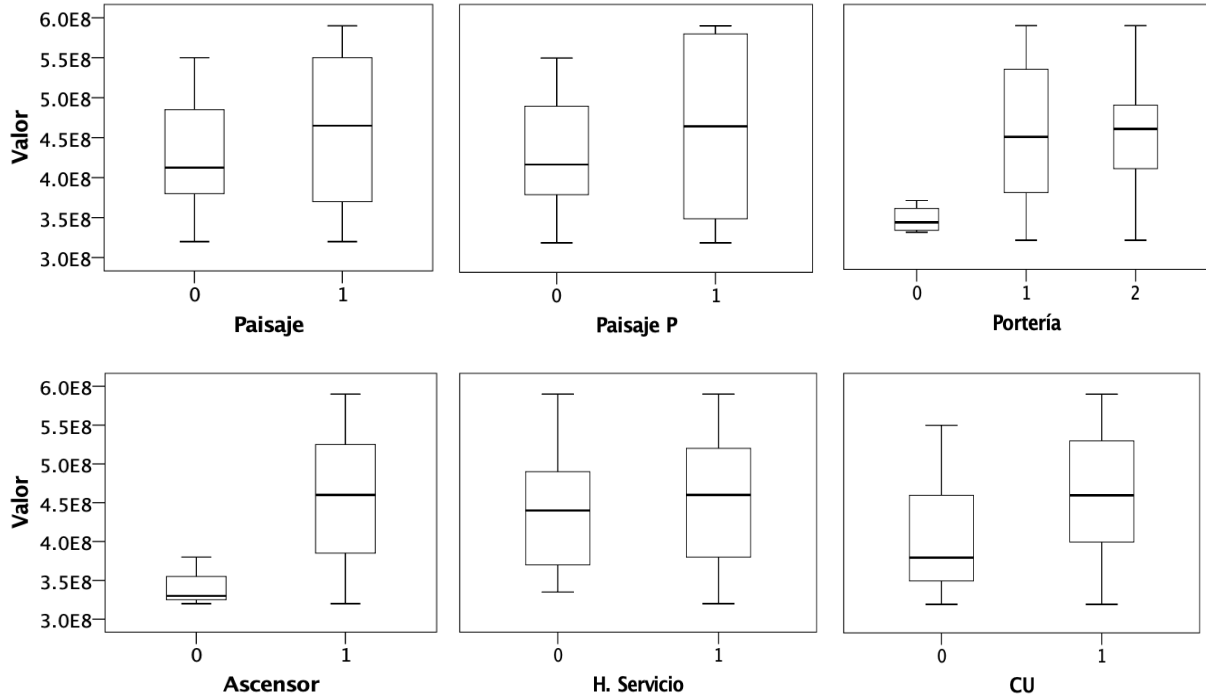


Figura 3. Box-Plot de valor del inmueble para la variables cualitativas. La línea central representa la mediana, la caja central indica la distribución del 50% de los datos y la barras (bigotes) indican la distribución del 25% de los datos en la parte inferior y el 25% de los datos en la parte superior.

Se observa que los datos no presentan diferenciación en el valor del inmueble respecto al Paisaje, por lo que es razonable pensar que este factor no tiene un efecto significativo en el valor, o su efecto esta siendo enmascarado por interacción con otras covariables, lo que se determinará en análisis posteriores. Igualmente, se observa que los valores no presentan diferenciación entre las muestras con o sin registro de paisaje a Parques del Río, por lo que se puede inferir que este factor no tiene un efecto significativo en el valor.

Por su parte, la distribución de los datos indica una desviación mayor para los inmuebles con portería diurna o 24 horas. Por su parte, en la zona central de los datos de apartamentos sin portería se observa que no existe una desviación considerable.

Respecto a la Portería, se evidencia un marcado desnivel entre los valores de los inmuebles sin portería y los demás, lo que permite sugerir que existe un efecto de esta variable sobre el valor.

Así mismo, en el gráfico se observa una distribución de los datos con diferencia considerable entre aquellos inmuebles con presencia de ascensor y los que no. La desviación de la tendencia central es amplia entre los factores, por lo que se puede inferir que la presencia y ausencia del ascensor podría estar influenciando el valor.

El comportamiento de los datos para la variable H. Servicio presenta una leve diferencia entre los inmuebles que tienen o no habitación del servicio. La tendencia central presenta un valor superior para inmuebles con habitación del servicio. Por lo tanto, es posible que estos factores estén explicando el valor de los inmuebles de la muestra.

Finalmente, el gráfico para la variable C.U. muestra una distribución de los datos con una diferencia considerable entre aquellos inmuebles con presencia de cuarto útil y los que no. La desviación de la tendencia central es amplia entre los factores, por lo que se puede inferir que la presencia y ausencia del cuarto útil puede estar influenciando el valor.

Del anterior análisis exploratorio de las variables cualitativas se concluye que existen diferencias marcadas en comportamiento de los datos para las variables: C. Útil, Ascensor y Portería.



Figura 4. Matriz de Correlación entre variables cuantitativas.

La medida de correlación determina una relación lineal entre 2 variables: -1 es relación lineal inversa, 1 relación lineal directa. Para mayor comprensión el grado de correlación entre -1 y 1 se representa con el color rojo oscuro y azul oscuro respectivamente al igual que el tamaño de la esfera.

En la matriz se puede observar que existe una correlación positiva y leve entre la administración, el área, parqueadero y el valor. Las distancias a pie y lineales tanto puntual como media, presentan una correlación positiva y moderada similar respecto al valor, resaltando la correlación de 0.3 para la DpieM respecto a las otra distancias que presentan valores inferiores. A su vez, las distancias entre ellas presentan

relaciones fuertes y positivas, indicando correlaciones entre variables explicativas de estimación de distancia.

En función del análisis anterior se eligieron las variables (cuantitativas) que no presentaban problemas de dependencia entre ellas. En el siguiente gráfico se presenta el comportamiento de las variables seleccionadas.

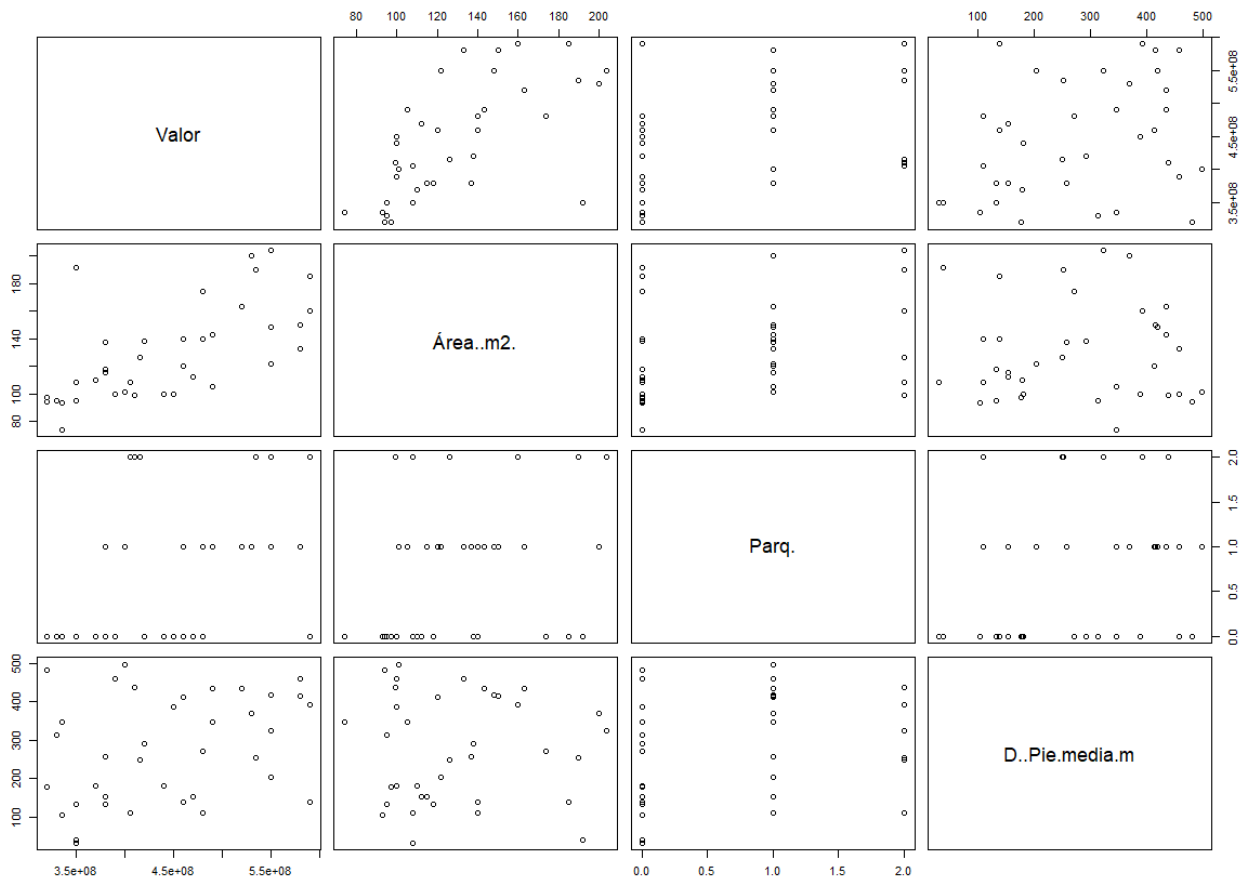


Figura 5.Gráfico de dispersión.

Indica el comportamiento de los datos para las variables cuantitativas explicativas elegidas que no presentaron algún problema de correlación entre ellas.

Con la información de los análisis exploratorios y la selección de algunas de las variables se continuo con los análisis exploratorios para determinar la significancia estadística.

Para la selección de los modelos, inicialmente se realizaron pruebas con el estadístico t para el análisis de significancia individual de cada una de estas variables, probando la variable respuesta en función de cada variable cualitativa de forma independiente. Posteriormente se realizan pruebas con el estadístico f para el análisis de significancia de la variable respuesta en función de cada variable cualitativa en presencia de las demás variables. Finalmente se realizan pruebas con el estadístico f para el análisis de significancia de la variable respuesta en función de las variables cualitativas en presencia de las variables cuantitativas previamente seleccionadas. A continuación se presentan los resultados de la prueba (Tabla 3, 4 y 5).

Tabla 3. Prueba t de significancia individual para variables cualitativas

Pruebas individuales	
Variable	Resultado
Paisaje	No significativa
Portería	Significativa
Ascensor	Significativa
Habitación del servicio	No significativa
Cuarto útil	No significativa

Aunque el análisis de la tabla anterior presenta resultados de significancia estadística, este no resulta concluyente pues la presencia de las demás variables puede generar un efecto de asociación o enmascaramiento en el comportamiento de los datos. En el anexo 1 se presentan los resultados para cada prueba realizada.

Tabla 4. Prueba f de significancia para variables cualitativas

En presencia de las demás			
Variable	Estadístico (F)	Valor p	Conclusión
Paisaje	3.0641	0.02318	No significativa
Portería	2.188	0.09348	Significativa
Ascensor	2.1394	0.08689	Significativa
Habitación del servicio	3.3143	0.01637	No significativa
Cuarto útil	2.4993	0.05170	Significativa

Tabla 5. Prueba f de significancia para variables cualitativas en presencia de las variables cuantitativas.

En presencia de variables cualitativas			
Variable	Estadístico F	Valor p	Conclusión
Paisaje	6.7739	0.01484	No significativa
Portería	1.5177	0.2373	Significativa
Ascensor	0.4572	0.5047	Significativa
Habitación del servicio	2.1324	0.1558	Significativa
Cuarto útil	0.9531	0.3376	Significativa

De el proceso anterior, se concluye que el mayor efecto individual sobre el valor del inmueble parece ejercerlo las variables: Portería y Ascensor, por lo que son las variables candidatas a ser incluidas en la postulación de modelos. Igualmente, aquellas variables cualitativas seleccionadas serán confirmadas en los modelos al ser corridos en ausencia de las mismas. Por ejemplo: si un modelo considerara una variable en particular como: el paisaje, se probará sin está para confirmar los resultados.

8.3 Postulación de modelos

Para la postulación de modelos, se consideraron tres análisis estadísticos de RLM a través de la metodología StepWise utilizando como criterio de decisión el Criterio de Información de Akaike (AIC), y la selección de subgrupos de variables de la función estadística *regsubsets()* del paquete *leaps*, ambos realizados con el software R. Este

último procedimiento permite fijar el número de variables deseadas en el modelo y genera todas las combinaciones de variables posibles para comparar el AIC de los modelos e informar que variables quedaron en los mejores ajustes estadísticos, siendo las que aparezcan en más modelos las que parecen tener mayor certeza de pertenecer al modelo final.

Primero, se realizó el procedimiento de StepWise para todas las variables iniciales pese a los posibles impedimentos de cumplimientos de supuestos de independencia entre las co-variables con el objetivo de no perder información de pronóstico o descartar previamente variables que podrían hacer parte de un posible modelo explicativo. Una vez filtradas las variables por el proceso StepWise, se probaron los supuestos de independencia lineal. Segundo, se realizó el procedimiento de StepWise para el grupo de las variables que ya se había verificado en el análisis exploratorio: Portería, Ascensor, Área y D. PieM, las cuales no tenían problemas de dependencia. Tercero, se realizó el procedimiento *regsubset()* para la selección de variables óptimas para un modelo de regresión. Este generó como resultado el mismo modelo 2 presentado en la tabla. El proceso y código utilizado se puede ver en el anexo 1. A continuación se presentan los modelos definidos (tabla 6).

Tabla 6. Resultado metodología StepWise

Modelos StepWise			
Modelo	R² ajustado	Estadístico <i>f</i>	Valor p
Modelo 1	0.7995	15.76	8,1x10 ⁻⁹
Modelo 2	0.6697	13.5	1,78*10 ⁻⁷

El modelo 1 considera que las variables: Altura, Área, Edad, Parqueadero, D. PieM, DistanciaM están explicando el valor de los inmuebles con un R² ajustado de 0.7995. El modelo 2 considera que las variables: Paisaje, Portería, H. Servicio, Área, y D. PieM, están explicando el valor de los inmuebles con un R² de 0.6697.

Ambos modelos presentan un R^2 ajustado estadísticamente alto. Las pruebas f de significancia permiten confirmar que ambos modelos son estadísticamente significativos.

8.4 Cumplimiento de supuestos

Una vez definidos los modelos, se continuó con el análisis de cumplimiento de los supuestos para cada modelo. El proceso considera las pruebas de multicolinealidad de las variables y normalidad en los errores con media cero, varianza constante e independencia de los datos. En estadística la multicolinealidad es una condición donde se presenta una correlación entre variables explicativas de un modelo.

La Independencia de las co-variables se realiza a través del análisis de multicolinealidad con un grupo de pruebas para cada modelo presentados en la tabla 7.

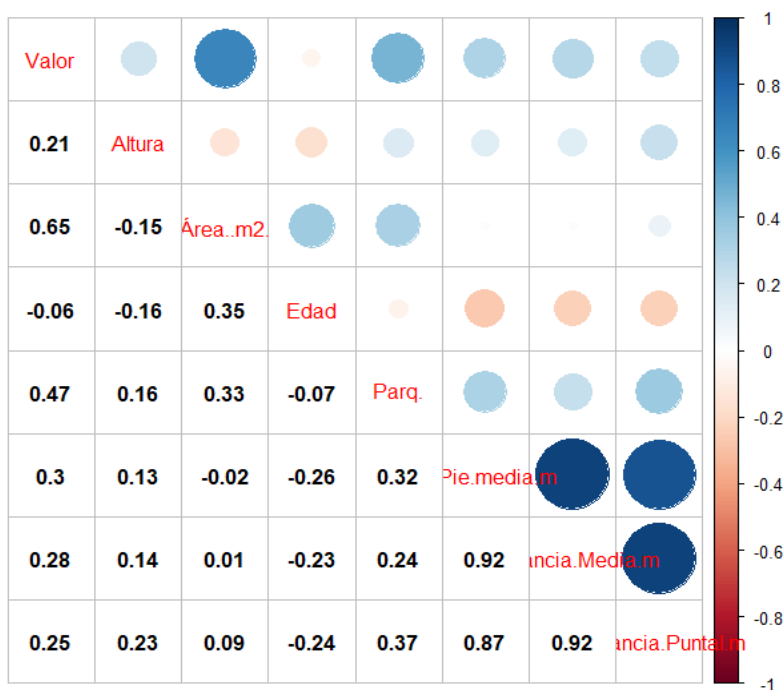


Figura 6. Matriz de correlación entre variables del modelo 1.

En la matriz de correlación (figura 6) indica una medida de correlación importante entre las variables Área y Valor. Adicionalmente se observa un problema de multicolinealidad entre las variables relativas a la distancia y una correlación positiva y media entre el parqueadero y la distancia, al igual que el parqueadero y el área. La información se confirma con el resultado de los test de multicolinealidad. Por esta razón el modelo 1 no es considerado para posteriores análisis de supuestos estadísticos y se continúa el proceso con el modelo 2.

Tabla 7. Pruebas de multicolinealidad para el modelo 1.

Overall Multicollinearity Diagnostics		
Test	MC Results	Detection
Determinant /X'X/:	0.0094	1
Farrar Chi-Square:	159.9423	1
Red Indicator:	0.3987	0
Sum of Lambda Inverse	34.799	0
Theil's Method:	-0.9736	0
Condition Number:	29.2156	0

1--- Collinerity is detected by the test
0--- Collinearity is not detected by the test

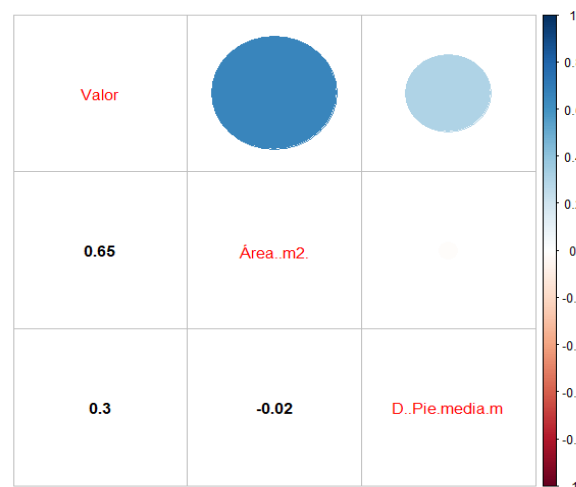


Figura 7. Matriz de correlación entre variables para el modelo 2.

En la matriz de correlación se observa independencia entre las variables (tabla 7). Tanto la variable Área como la variable D.PieM indican una correlación positiva con el valor (figura 7). La información confirma el resultado de las pruebas de multicolinealidad (Tabla 8). Por lo cual se continúa a probar los supuestos de los errores con el modelo 2.

Tabla 8. Pruebas de multicolinealidad para el modelo 2.

Overall Multicollinearity Diagnostics		
Test	MC Results	Detection
Determinant /X'X/:	0.9998	0
Farrar Chi-Square:	0.0078	0
Red Indicator:	0.0154	0
Sum of Lambda Inverse	2.0005	0
Theil's Method:	-0.5218	0
Condition Number:	9.8264	0

1--- Collinearity is detected by the test
0--- Collinearity is not detected by the test

A continuación, se presentan los resultados donde se prueban los supuestos de media cero, varianza constante e independencia sobre los errores en el modelo 2. El resultado del análisis de Shapiro-wilk con $p > 0.05$ acepta la hipótesis alternativa que establece que el conjunto de datos de los errores se distribuye normalmente (tabla 9).

Tabla 9. Prueba de normalidad sobre los errores para el modelo 2.

Shapiro-wilk normality test
Data: resmod2
W = 0.98452, p-value = 0.8671

El gráfico de Normalidad Q-Q indica el ajuste normal de los errores. Este método analiza de forma gráfica las diferencias entre la distribución de probabilidad de una muestra aleatoria y una distribución normal para su comparación. Como la distribución de la variable esta muy ajustada a la línea (en rojo) que representa la distribución de comparación el estadístico aprueba la hipótesis de similitud (figura 8). De igual forma la distribución de los datos observada en el gráfico de diagnósticos de residuos indica que no existe un patrón de comportamiento en los datos y que estos están distribuidos aleatoriamente entre los cuantiles 2 y -2.

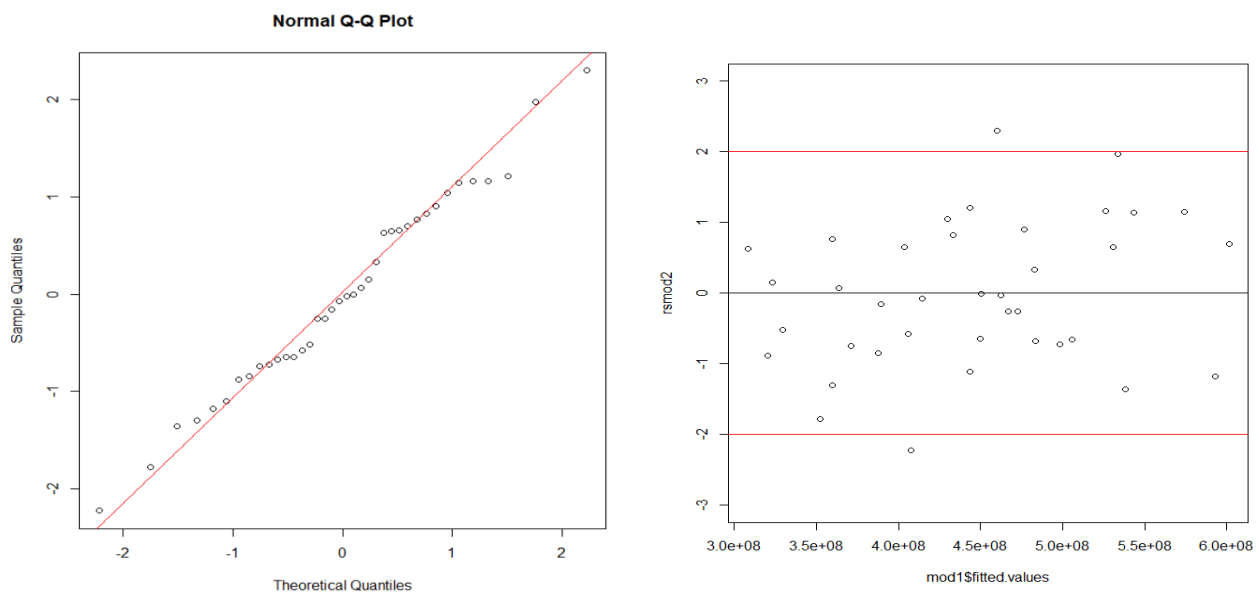


Figura 8. Gráfico Q-Q de prueba de normalidad para los errores en el modelo y diagnóstico de los residuos para el modelo 2.

A continuación, se presenta el modelo definido después del proceso con R^2 ajustado y valor $p < 0.05$.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon_i$$

Donde:

y = Valor respuesta de los inmuebles

β_0 = Intercepto de la recta de la ecuación lineal

β_1 = Coeficiente de la observación de la variable Área

x_1 = observación de la variable Área

β_2 = Coeficiente de la observación de la variable D.pieM

x_2 = observación de la variable D.pieM

β_3 = Coeficiente de la observación de la variable Paisaje

x_3 = observación de la variable Paisaje

β_4 = Coeficiente de la observación de la variable Portería

x_4 = observación de la variable Portería

β_5 = Coeficiente de la observación de la variable H. Servicio

x_5 = observación de la variable H. Servicio

ε_i = Error del modelo en la i -ésima observación

Por lo tanto tenemos:

$$\mathbf{Valor} = 108'620,933 + 1'861,400*\mathbf{Área} + 220,233*\mathbf{D.pieM} + (0, 1)*\mathbf{Paisaje} + (0, 1, 2)*\mathbf{Portería} - (0, 1)*\mathbf{H.servicio}$$

Donde Paisaje asume: 0 = **sin paisaje** y 1 = 51'212,189 **con Paisaje**

Donde Portería asume: 0 = **sin portería**, 1 = 15'495,422 con **portería diurna** y 2 = 69'190,907 con **portería 24h**

Donde H.servicio: 0 = **sin H.servicio** y 1 = - 48'908,016 con **H.servicio**

A continuación, se realizó el análisis de correlación de Pearson para probar la significancia estadística de la correlación independiente entre el valor y las variables explicativas: distancia a pie media. La prueba realizada indica que la distancia y el valor no presentan una correlación estadísticamente significativa, con $p > 0.05$. Por su

parte, el área y el valor presentan una correlación positiva con $R^2 = 0.0651$ con $p < 0.05$ (Tabla 10).

Tabla 10. Pruebas de correlación del valor respecto las variables explicativas

Correlación Pearson		
D. Pie media/m	Pearson Correlation	0.304
	Sig. (2-tailed)	0.064
	N	38
Área (m2)	Pearson Correlation	.651**
	Sig. (2-tailed)	0.000
	N	38

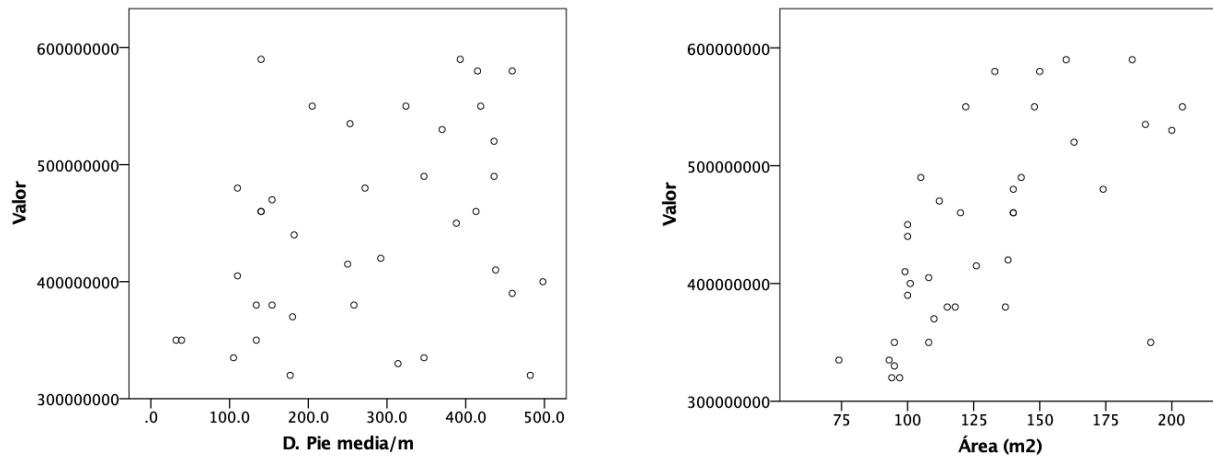


Figura 9. Dispersión de los datos en función del valor para las variables Distancia a pie media y Área.

El modelo final se aplicó sobre 2 inmuebles del muestreo inicial para revisar el valor comercial obtenido por el modelo y contrastar la diferencia con los valores de oferta real de la muestra.

Muestra 1 (Valor de campo) = \$530,000,000 pesos
 Área = 200m²
 D.pieM = 370m
 Paisaje = Si
 Portería = Diurna

H. Servicio = Si

$$\text{Valor} = 108'620,933 + 1'861,400 * 200\text{m}^2 + 220,233 * 370\text{m} + 51'212,189 + 15'495,422 - 48'908.016$$

$$\text{Valor} = 580,186,738 \text{ pesos}$$

El valor obtenido por el modelo en la muestra 1 representa el 9,47% mayor del valor de oferta real.

Muestra 2 (Valor de campo) = \$380,000,000 pesos

$$\text{Área} = 115\text{m}^2$$

$$\text{D.pieM} = 154\text{m}$$

Paisaje = No

Portería = 24h

H. Servicio = Si

$$\text{Valor} = 108'620,933 + 1'861,400 * 115\text{m}^2 + 220,233 * 154\text{m} + 69'190,907 - 48'908.016$$

$$\text{Valor} = 376,880,706 \text{ pesos}$$

El valor obtenido por el modelo en la muestra 2 representa el 0,82% inferior del valor de oferta real.

Finalmente se realiza el cálculo del tamaño mínimo de muestra para resultados estadísticos. En el área delimitada se determinó una población de 2052 inmuebles sometidos a Régimen de Propiedad Horizontal. Las 38 muestras analizadas indican un nivel de confianza del 91% y un margen de error del 5% en una distribución normal.

9 Discusión

Numerosos factores pueden estar influenciando el valor de los inmuebles en el sector de conquistadores. Aspectos de la condición física del inmueble (intrínsecos) como estado de los acabados, área, edad, presencia de cuarto útil, habitación del servicio, balcón, número de parqueaderos, registro visual y exposición al poniente pueden ser

importantes en este proceso. Otros factores relacionados con la condición de la edificación como: disponibilidad de servicios recreativos, estado de conservación de zonas comunes, valor de administración, y también, condiciones extrínsecas como el acceso servicios de transporte, educativos, parques recreativos y zonas naturales, centros comerciales, y servicios religiosos y espirituales son de igual relevancia.

En este estudio se realiza un acercamiento a algunos de los factores que pueden estar influyendo en un proceso económico en conquistadores, y a través de los resultados aquí obtenidos, se intentó comprender como algunas de estas variables están impactando sobre el valor comercial de los inmuebles en la zona. Por esta razón, este estudio planteó estimar el efecto que tiene la distancia de los inmuebles respecto a Parques del Río con el valor de los inmuebles en el sector de Conquistadores a través del análisis de correlación y regresión lineal.

Los análisis exploratorios indican que el área, el costo de administración, la cantidad de parqueaderos y la distancia a Parques del Río están influenciado de forma directa el valor de los inmuebles en área de estudio. Por su parte, el área y la distancia a pie analizadas en conjunto son los factores que mejor se ajustan a el modelo (modelo 2) de RLM para explicar el valor de los inmuebles.

El modelo indica que los parámetros: área del inmueble, celda de parqueo, vigilancia de la edificación, habitación del servicio, registro de paisaje y distancia a pie de parques del Río esta influenciando el valor que puede tomar un inmueble en el sector. De este análisis, se confirma que las variables tradicionales como el área de los inmuebles, la presencia de celda de parqueo y la vigilancia de la edificación están determinando el valor del inmueble. A su vez la distancia a Parques del Río surge como un parámetro que esta hace parte de las variables explicativas del valor.

Con una R^2 ajustado de 0.6697 se puede estimar que por cada metro de distancia a pie que un inmueble se aleje del parque este incrementará \$220,233 pesos. A su vez, determina que los inmuebles incrementan \$1'861,400 pesos por cada m^2 de área construida. En este sentido, los datos permiten inferir que, de forma leve, la distancia tiene un efecto sobre el valor de los inmuebles en el área de estudio y a medida que un apartamento se aleja de Parques del Río este tendrá un valor superior.

Este fenómeno contrasta con los hallazgos presentados por Tekel y otro (2013) y Trojanek et al. (2018). En ambos casos, infieren que la cercanía al parque impacta positivamente el valor de los inmuebles. Se esperaría que en el área de estudio se viviera el mismo fenómeno basado en el impacto positivo que tiene la obra. Si la obra agrega calidad social, mejoramiento arquitectónico y urbanístico, aumento de zonas verdes, y calidad paisajística, se esperaría que la demanda por inmuebles con estos servicios extrínsecos es mas alto y por lo tanto su valor sea superior. Es posible que este fenómeno que observamos se presente de forma inversa con Parques del Río debido a una percepción de inseguridad, suciedad y altos niveles de ruido que han tenido algunos parques públicos de esparcimiento y recreación de la ciudad (Moncado, 2019). Adicionalmente, la presencia de parques arborizados en el trazado central de conquistadores pueda estar influenciando el valor en función de la ubicación, y quizá la influencia del trazado central de estas áreas verdes puedan estar generando un efecto.

Es de reconocer que el estudio se realizó con valores de oferta sin considerar los precios de venta. Esto se debe a que la información de la transacción de los inmuebles es de carácter privado y se desconoce el valor real de la compraventa. Por lo tanto, esto puede estar influenciado el comportamiento de los datos y por ende el resultado del análisis.

Son múltiples factores los que pueden estar explicando el resultado obtenido que debe corroborarse con estudios posteriores incluyendo datos que complementen la muestra actual.

En otro sentido el análisis de correlación independiente de Pearson evidencia que no existe una relación estadísticamente significativa entre el valor de los inmuebles y la distancia a Parques del Río. En general y teniendo en cuenta los análisis de correlación de Pearson para el análisis independiente, indicadores de R^2 ajustados del modelo presentado y el ajustado número de muestras obtenido de la población de inmuebles usados en la zona de estudio, podemos inferir que la distancia no presenta un impacto sobre el valor de los inmuebles del polígono estudiado.

En otro aspecto, Mercado (2017) presenta el impacto positivo de Parques del Río para la comunidad de Conquistadores y la ciudad en general. Lo cual indica que la presencia del parque mejoró la calidad de vida de la población que habita en la zona. En este sentido la influencia del parque es susceptible de ser medida a escala temporal y quizá el impacto sobre el valor de las propiedades en el área de influencia de este tipo de parques de esparcimiento y recreación se observe mejor con estudios de escala temporal en contraste con la distribución espacial. Por lo tanto, mayores estudios son requeridos para profundizar en el conocimiento de esta temática, y así permitir ilustrar mejor el impacto que pueda estar teniendo los parques de esparcimiento y recreación urbanas en el municipio de Medellín.

10 Conclusiones

El análisis de Regresión Lineal Multivariada indica que el modelo con las variables: Área, Distancia a Pie Media, Paisaje, Portería y Habitación del Servicio explican con un 66.97% el valor de los inmuebles en el área de estudio. Lo cual permite inferir

que, con este análisis, la distancia a pie de parques del Río incide de forma inversa sobre el valor comercial de los inmuebles.

Por su parte la correlación de Pearson indica que la distancia a pie a Parques del Río no presenta una correlación estadísticamente significativa con el valor comercial de los inmuebles en el área de estudio. Por lo tanto, según esta prueba, la influencia de la distancia sobre el valor comercial inmuebles no es perceptible.

Teniendo en cuenta el R^2 ajustado del modelo, el resultado de correlación de Pearson y la consideración del mínimo muestral, se requiere de posteriores análisis con un número mayor de muestras para determinar el efecto que puede tener Parques del Río en función de la distancia sobre el valor comercial de los inmuebles en el sector estudiado.

Respecto a los hallazgos de este estudio, se esperaba que este parque de esparcimiento y recreación de uso público influenciara positivamente el valor comercial de los inmuebles debido a que son desarrollos urbanísticos que agregan calidad de vida al área de influencia, en este caso conquistadores. Este fenómeno se evidencia en los estudios realizados por otros investigadores donde el valor comercial de viviendas se incrementa a medida que se acerca a parques urbanos.

Los resultados del presente trabajo confirman los hallazgos de estudios en la ciudad de Medellín, específicamente en el Barrio conquistadores donde se observa una correlación positiva y fuerte entre el valor de los inmuebles y el área de construcción. Igualmente, el registro de paisaje y la presencia de vigilancia son factores que agregan valor comercial.

El modelo de precios hedónicos es un análisis que aporta información fundamental a la comprensión del análisis del valor de inmuebles y la interacción con múltiples

variables explicativas. Por lo tanto esta es una herramienta pertinente para determinar cómo los equipamientos están influenciando atributos, entre ellas el valor comercial que dan cuenta de un comportamiento social en la demanda y oferta de inmuebles en zona determinada.

Esta investigación provee información acerca de la influencia que generan, en los valores comerciales, las áreas de esparcimiento (Parques del Río), en la Ciudad de Medellín.

11 Recomendaciones

El avance en estudios considerando un N muestral superior, al igual que mayor número de variables y conociendo los precios de venta son necesarios para lograr conclusiones más robustas acerca del valor comercial de los inmuebles en áreas de esparcimiento en la ciudad. Programas de monitoreo y de investigación de mercado deben estar enfocados en el planteamiento de estos proyectos de transformación urbana.

Agradecimientos

Este estudio se realizó con el apoyo del Observatorio Inmobiliario de la Subsecretaría de Catastro de Medellín, el acompañamiento del Profesor Jorge Agudelo Torres, docente de la especialización en Valoración de ESUMER, y el Estadístico Daniel Betancur de la Universidad Nacional, sede Medellín.

Referencias

Alcaldía de Medellín (2014) “Acuerdo n 48° de 2014, Por medio del cual se adopta la revisión y ajuste de largo plazo del Plan de Ordenamiento Territorial del Municipio de Medellín y se dictan otras disposiciones complementarias” Gaceta oficial 4267 de 2014.

Crompton John L. (2001) “The Impact of Parks on Property Values: A Review of the Empirical Evidence, Journal of Leisure Research”, Vol. 33 Issue: 1, pp. 1-31.

Geltner David M. , Clayton Jim , Miller Norman G. , Eichholtz Piet (2013), “Commercial Real Estate: Analysis & Investments” third edition, OnCourse Learning Publishing.

Gómez Beatriz, Wolf Amaya Gilda (2008).” Barrio Conquistadores”, Escuela de Urbanismo, Universidad Nacional de Colombia Sede Medellín, Vol 1.

Gujarati Damodar N. y Porter Dawn C. (2010), “Econometría” Quinta edición, McGraw Hill.

Hendon William S. (1971), “The Park as a Determinant of Property Values” The American Journal of Economics and Sociology, Vol. 30, No. 3, pp. 289-300.

Kaul, S. (2006), “Hedonism and Culture: Impact on Shopper Behaviour”, Indian Institute of Management Publication, Ahmedabad, India, 10:4-5

Mercado David Alejandro (2017), “Los parques que unen la ciudad en los costados

del río Medellín” El Tiempo.

Moncado Perez David Alejandro (2019), Seguridad en el Parque del Poblado no mejoró con el Código de Policía, El Tiempo.

Munoz-Razkin Ramon (2010) “Walking accessibility to bus rapid transit: Does it affect property values? The case of Bogotá, Colombia”. *Transport Policy*, Vol. 17 pp. 72-84.

Pagourtzi Elli, Assimakopoulos Vassilis, Hatzichristos Thomas, French Nick, (2003) "Real estate appraisal: a review of valuation methods", *Journal of Property Investment & Finance*, Vol. 21 Issue: 4, pp.383-401.

Perdomo Jorge Andrés (2010), “Una propuesta metodológica para estimar los cambios sobre el valor de la propiedad: estudio de caso para Bogotá aplicando Propensity Score Matching y Precios Hedónicos Espaciales” *Lecturas de Economía*, Vol. 73 pp. 49-65.

Quintero Katherin, Rivera Leidy J., Marín Nury (2012) “Efecto de la distancia al acceso de transporte masivo Megabús sobre el valor de la propiedad^[1] en la zona sub-centro del barrio Cuba” *Gráficas Disciplinarias de la UCP*, Vol. 19 pp. 73-78.

Rodríguez Daniel A., y Targa Felipe (2004) “Value of Accessibility to Bogotá’s Bus Rapid Transit System” *Transport Review*, Vol. 24 pp. 587-610.

Tekel Ayse, Akbarishahabi Leila (2013) “Determination of Open-green Space's Effect on Around House Prices by Means of Hedonic Price Model; in Example of Ankara/Botanik Park” *Gazi University Journal of Science*, Vol. 6 Issue: 2 pp. 347-360.

Trojanek Radoslaw, Gluszek Michal, Tanas Justyna, (2018), “The effect of urban green spaces on house prices in Warsaw” *International Journal of Strategic Property Management*, Vol. 22 Issue 5: pp. 358–371

Votsis Athanasios (2017) “Planning for green infrastructure: The spatial effects of parks, forests, and fields on Helsinki's apartment prices” *Ecological Economics*, Vol. 132, pp. 279-289.

Anexos

En los anexos se visualiza los resultados de las diferentes pruebas realizadas con el programa R para análisis estadístico. El anexo 1 al 5 presenta los resultados arrojados para las pruebas individuales, pruebas múltiples y definición de modelos. El anexo 6 contiene el código R de todo el análisis realizado.

Anexo 1. Pruebas individuales con estadístico *t*

```
Call:
lm(formula = Valor ~ Paisaje, data = datosf)

Residuals:
    Min       1Q   Median       3Q      Max
-140588235 -54607843  -588235   59142157 158333333

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  431666667   18400339   23.460  <2e-16 ***
PaisajeCon_paisaje  28921569   27510163    1.051    0.3
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 84320000 on 36 degrees of freedom
Multiple R-squared:  0.02979, Adjusted R-squared:  0.002836
F-statistic: 1.105 on 1 and 36 DF, p-value: 0.3001

Call:
lm(formula = Valor ~ Portería, data = datosf)

Residuals:
    Min       1Q   Median       3Q      Max
-140238095 -53988095    73260   52261905 140384615

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  346250000   39611113   8.741 2.53e-10 ***
PorteríaDiurna 103365385   45297041   2.282  0.0287 *
Portería24horas 113988095   43219267   2.637  0.0124 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 79220000 on 35 degrees of freedom
Multiple R-squared:  0.1674, Adjusted R-squared:  0.1198
F-statistic: 3.518 on 2 and 35 DF, p-value: 0.04055
```



```

Call:
lm(formula = Valor ~ Ascensor, data = datosf)

Residuals:
    Min       1Q   Median       3Q      Max
-133285714 -60785714  1714286  59214286 136714286

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  343333333   46189406   7.433 8.95e-09 ***
AscensorTiene 109952381   48128259   2.285  0.0283 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8e+07 on 36 degrees of freedom
Multiple R-squared:  0.1266, Adjusted R-squared:  0.1024
F-statistic: 5.219 on 1 and 36 DF, p-value: 0.02834

```

```

Call:
lm(formula = Valor ~ H..Servicio, data = datosf)

Residuals:
    Min       1Q   Median       3Q      Max
-128200000 -68073077  7053846  66926923 152307692

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  437692308   23700174   18.47 <2e-16 ***
H..ServicioTiene 10507692   29219536    0.36  0.721
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 85450000 on 36 degrees of freedom
Multiple R-squared:  0.003579, Adjusted R-squared: -0.0241
F-statistic: 0.1293 on 1 and 36 DF, p-value: 0.7212

```

```

Call:
lm(formula = Valor ~ CU, data = datosf)

Residuals:
    Min       1Q   Median       3Q      Max
-138448276 -57198276 -9224138  67887931 150000000

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 400000000   27236684   14.686 <2e-16 ***
CUTiene     58448276   31177912    1.875  0.069 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 81710000 on 36 degrees of freedom
Multiple R-squared:  0.08894, Adjusted R-squared:  0.06363
F-statistic: 3.514 on 1 and 36 DF, p-value: 0.06897

```

Anexo 2. Pruebas *f* con variables factor

Analysis of Variance Table

Model 1: Valor ~ Paisaje + Portería + Ascensor + H..Servicio + CU

Model 2: Valor ~ Paisaje

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	31	1.7130e+17				
2	36	2.5596e+17	-5	-8.4658e+16	3.0641	0.02318 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(modf0, modf2) # Significativa

Analysis of Variance Table

Model 1: Valor ~ Paisaje + Portería + Ascensor + H..Servicio + CU

Model 2: Valor ~ Portería

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	31	1.7130e+17				
2	35	2.1967e+17	-4	-4.8363e+16	2.188	0.09348 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(modf0, modf3) # Significativa

Analysis of Variance Table

Model 1: Valor ~ Paisaje + Portería + Ascensor + H..Servicio + CU

Model 2: Valor ~ Ascensor

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	31	1.7130e+17				
2	36	2.3041e+17	-5	-5.9111e+16	2.1394	0.08689 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(modf0, modf4) # No significativa

Analysis of Variance Table

Model 1: Valor ~ Paisaje + Portería + Ascensor + H..Servicio + CU

Model 2: Valor ~ H..Servicio

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	31	1.7130e+17				
2	36	2.6287e+17	-5	-9.1572e+16	3.3143	0.01637 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> anova(modf0, modf5) # Significativa

Analysis of Variance Table

Model 1: Valor ~ Paisaje + Portería + Ascensor + H..Servicio + CU

Model 2: Valor ~ CU

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	31	1.7130e+17				
2	36	2.4036e+17	-5	-6.9053e+16	2.4993	0.0517 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Anexo 3. Pruebas *f* en presencia de todas las variables

Analysis of Variance Table

```
Model 1: Valor ~ Paisaje + Portería + Ascensor + H..Servicio + CU + Adm. +
  Área..m2. + Parq. + D..Pie.media.m
Model 2: Valor ~ Portería + Ascensor + H..Servicio + CU + Adm. + Área..m2. +
  Parq. + D..Pie.media.m
  Res.Df      RSS Df    Sum of Sq      F Pr(>F)
1      27 6.6796e+16
2      28 8.3555e+16 -1 -1.6758e+16 6.7739 0.01484 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> anova(modd2, modd4) # Significativa
```

Analysis of Variance Table

```
Model 1: Valor ~ Paisaje + Portería + Ascensor + H..Servicio + CU + Adm. +
  Área..m2. + Parq. + D..Pie.media.m
Model 2: Valor ~ Paisaje + Ascensor + H..Servicio + CU + Adm. + Área..m2. +
  Parq. + D..Pie.media.m
  Res.Df      RSS Df    Sum of Sq      F Pr(>F)
1      27 6.6796e+16
2      29 7.4306e+16 -2 -7.5094e+15 1.5177 0.2373
> anova(modd2, modd5) # Signifivativa
```

Analysis of Variance Table

```
Model 1: Valor ~ Paisaje + Portería + Ascensor + H..Servicio + CU + Adm. +
  Área..m2. + Parq. + D..Pie.media.m
Model 2: Valor ~ Paisaje + Portería + Ascensor + CU + Adm. + Área..m2. +
  Parq. + D..Pie.media.m
  Res.Df      RSS Df    Sum of Sq      F Pr(>F)
1      27 6.6796e+16
2      28 6.7928e+16 -1 -1.1311e+15 0.4572 0.5047
> anova(modd2, modd6) # Significativa
```

Analysis of Variance Table

```
Model 1: Valor ~ Paisaje + Portería + Ascensor + H..Servicio + CU + Adm. +
  Área..m2. + Parq. + D..Pie.media.m
Model 2: Valor ~ Paisaje + Portería + Ascensor + CU + Adm. + Área..m2. +
  Parq. + D..Pie.media.m
  Res.Df      RSS Df    Sum of Sq      F Pr(>F)
1      27 6.6796e+16
2      28 7.2072e+16 -1 -5.2754e+15 2.1324 0.1558
> anova(modd2, modd7) # Significativa
```

Analysis of Variance Table

```
Model 1: Valor ~ Paisaje + Portería + Ascensor + H..Servicio + CU + Adm. +
  Área..m2. + Parq. + D..Pie.media.m
Model 2: Valor ~ Paisaje + Portería + Ascensor + H..Servicio + Adm. +
  Área..m2. + Parq. + D..Pie.media.m
  Res.Df      RSS Df    Sum of Sq      F Pr(>F)
1      27 6.6796e+16
2      28 6.9154e+16 -1 -2.3579e+15 0.9531 0.3376
```

Anexo 4. Análisis R Modelo 1

```
Call:
lm(formula = Valor ~ Altura + Área..m2. + Edad + Portería + CU +
    Parq. + D..Pie.media.m + Distancia.Media.m + Distancia.Puntal.m,
    data = datosf)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-62643209 -18283914 -5082427  15596929  89922309
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  171936081  38609489  4.453 0.000132 ***
Altura       10363682   2330053  4.448 0.000134 ***
Área..m2.    2041764     234312   8.714 2.50e-09 ***
Edad        -1257533     594163  -2.116 0.043666 *
PorteríaDiurna -10726882  24863082 -0.431 0.669575
Portería24horas 29878465  22915801  1.304 0.203297
CUTiene     36900491  18791000  1.964 0.059935 .
Parq.       22340691  10415581  2.145 0.041113 *
D..Pie.media.m  160924     129991  1.238 0.226388
Distancia.Media.m  545818    191956  2.843 0.008403 **
Distancia.Puntal.m -876861    186587 -4.699 6.83e-05 ***
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 37810000 on 27 degrees of freedom
Multiple R-squared:  0.8537,    Adjusted R-squared:  0.7995
F-statistic: 15.76 on 10 and 27 DF,  p-value: 8.102e-09
```

Anexo 5. Análisis R Modelo 2

```
Call:
lm(formula = Valor ~ Paisaje + Portería + H..Servicio + Área..m2. +
    D..Pie.media.m, data = datosf)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-91186287 -31275923 -2302188  30821079 102432793
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  108620933  39887097  2.723 0.01052 *
PaisajeCon_paisaje  51212189  16685576  3.069 0.00443 **
PorteríaDiurna    15495422  31760941  0.488 0.62907
Portería24horas   69190907  28934183  2.391 0.02304 *
H..ServicioTiene -48908016  19279542 -2.537 0.01644 *
Área..m2.        1861400    272482   6.831 1.18e-07 ***
D..Pie.media.m    220233     62782   3.508 0.00140 **
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 48530000 on 31 degrees of freedom
Multiple R-squared:  0.7233,    Adjusted R-squared:  0.6697
F-statistic: 13.5 on 6 and 31 DF,  p-value: 1.78e-07
```

Anexo 6. Código R

```
# PAQUETES -----  
  
library(mctest)  
library(ggplot2)  
library(corrplot)  
library(MASS)  
library(leaps)  
library(pwr)  
  
# Datos -----  
  
datos <- read.table(file = "Apartamentos.txt", header = F) #Lectura de los datos  
como numeros  
head(datos)  
dt <- read.csv2(file = "Aptos.csv", header = T) #Lectura de los títulos  
names(dt)  
names(datos) = names(dt) #Asignación de los títulos  
head(datos)  
summary(datos)  
  
# Depuración y codificación -----  
  
datos1 <- datos[,-1] #Retirando columna del orden  
head(datos1)  
  
#Codificando variables factor  
  
datos1$Paisaje <- factor(datos1$Paisaje, levels = c(0,1),  
                        labels = c("Sin_paisaje", "Con_paisaje"))  
datos1$Portería <- factor(datos1$Portería, levels = c(0, 1, 2),  
                        labels = c("Sin_vigilancia", "Diurna", "24horas"))  
datos1$Ascensor <- factor(datos1$Ascensor, levels = c(0,1),  
                        labels = c("No_tiene", "Tiene"))  
datos1$H..Servicio <- factor(datos1$H..Servicio, levels = c(0,1),  
                        labels = c("No_tiene", "Tiene"))  
datos1$CU <- factor(datos1$CU, levels = c(0,1),  
                   labels = c("No_tiene", "Tiene"))  
head(datos1)  
  
#Retirando variables que implican el área en su cálculo
```

```

datosf <- datos1[,-c(2, 4)]
head(datosf)

# Análisis descriptivo y exploratorio -----

# Factores

factores <- datosf[,c(1, 4, 7, 8, 10, 11)]
head(factores)
gg1 <- ggplot(data = factores, aes(Paisaje, Valor)) + geom_boxplot()
gg1
gg2 <- ggplot(data = factores, aes(Portería, Valor)) + geom_boxplot()
gg2
gg3 <- ggplot(data = factores, aes(Ascensor, Valor)) + geom_boxplot()
gg3
gg4 <- ggplot(data = factores, aes(H..Servicio, Valor)) + geom_boxplot()
gg4
gg5 <- ggplot(data = factores, aes(CU, Valor)) + geom_boxplot()
gg5

# Numéricas

numericas <- datosf[,c(1,2,3,5,6,9,12,13,14,15,16,17)]
head(numericas)

plot(numericas)
pairs(numericas)
M <- cor(numericas)
corrplot.mixed(M, lower.col = "black")
mctest(numericas[,-1], numericas[,1])

#Incluyendo solo las más explicativas (y eliminando correlaciones muy altas)

explicativas <- numericas[,c(1, 2, 4, 7, 9)]
plot(explicativas)
mctest(explicativas[,-1], explicativas[,1])
sinadmon <- explicativas[,-2]
mctest(sinadmon[,-1], sinadmon[,1])
plot(sinadmon)

# Postulación de modelos iniciales -----

```

```

# Para evaluación de variables factor

# Individualmente
modf0 <- lm(Valor ~ Paisaje + Portería + Ascensor + H..Servicio + CU, data =
datosf)
summary(modf0)
modf1 <- lm(Valor ~ Paisaje, data = datosf)
summary(modf1) # No significativa
modf2 <- lm(Valor ~ Portería, data = datosf)
summary(modf2) # Significativa
modf3 <- lm(Valor ~ Ascensor, data = datosf)
summary(modf3) # Significativa
modf4<- lm(Valor ~ H..Servicio, data = datosf)
summary(modf4) # No significativa
modf5 <- lm(Valor ~ CU, data = datosf)
summary(modf5) # No significativa
# Con Anova
modaf1 <- lm(Valor ~ Portería + Ascensor + H..Servicio + CU, data = datosf)
modaf2 <- lm(Valor ~ Paisaje + Ascensor + H..Servicio + CU, data = datosf)
modaf3 <- lm(Valor ~ Paisaje + Portería + H..Servicio + CU, data = datosf)
modaf4 <- lm(Valor ~ Paisaje + Portería + Ascensor + CU, data = datosf)
modaf5 <- lm(Valor ~ Paisaje + Portería + Ascensor + H..Servicio, data = datosf)
anova(modf0, modf1) # No significativa
anova(modf0, modf2) # Significativa
anova(modf0, modf3) # Significativa
anova(modf0, modf4) # No significativa
anova(modf0, modf5) # Significativa

#En presencia de las demás

modd1 <- lm(Valor ~ ., data = datosf)
modd2 <- lm(Valor ~ Paisaje + Portería + Ascensor + H..Servicio + CU + Adm. +
Área..m2. + Parq.
+ D..Pie.media.m, data = datosf)
summary(modd2)
modd3 <- lm(Valor ~ Portería + Ascensor + H..Servicio + CU + Adm. + Área..m2. +
Parq.
+ D..Pie.media.m, data = datosf)
modd4 <- lm(Valor ~ Paisaje + Ascensor + H..Servicio + CU + Adm. + Área..m2. +
Parq.
+ D..Pie.media.m, data = datosf)
modd5 <- lm(Valor ~ Paisaje + Portería+ H..Servicio + CU + Adm. + Área..m2. +
Parq.

```

```

      + D..Pie.media.m, data = datosf)
modd6 <- lm(Valor ~ Paisaje + Portería + Ascensor + CU + Adm. + Área..m2. +
Parq.
      + D..Pie.media.m, data = datosf)
modd7 <- lm(Valor ~ Paisaje + Portería + Ascensor + H..Servicio+ Adm. + Área..m2.
+ Parq.
      + D..Pie.media.m, data = datosf)
anova(modd2, modd3) # No significativa
anova(modd2, modd4) # Significativa
anova(modd2, modd5) # Signifivativa
anova(modd2, modd6) # Significativa
anova(modd2, modd7) # Significativa

# Se seleccionan: Portería, Ascensor, CU

# Modelos stepwise -----

# Modelo completo

modc <- lm(Valor ~ ., data = datosf)

# Modelo stepwise

stepmc <- stepAIC(modc, direction = "both", trace = FALSE)
summary(stepmc)

# subsets

modrs1 <- regsubsets(Valor ~ ., data = datosf, nvmax = 5, method = "seqrep")
summary(modrs1)
# con 5 Elige Area, Porteria, P..parque, D..pie.media.m, altura

modrs2 <- regsubsets(Valor ~ ., data = datosf, nvmax = 10, method = "seqrep")
summary(modrs2)

# Modelo prefiltrado
modp <- lm(Valor ~ Paisaje + Portería + Ascensor + H..Servicio + CU + Adm. +
Área..m2. + Parq.
      + D..Pie.media.m, data = datosf)

stepmp <- stepAIC(modp, direction = "both", trace = FALSE)
summary(stepmp)

```



```

# subsets

modrs3 <- regsubsets(Valor ~ Paisaje + Portería + Ascensor + H..Servicio + CU +
Adm. + Área..m2. + Parq.
                + D..Pie.media.m, data = datosf, nvmax = 5, method = "seqrep")
summary(modrs3)
# Entran Paisaje, porteria, ascensor, area, d..pie.media.m, H..Servicio

modrs4 <- regsubsets(Valor ~ Paisaje + Portería + Ascensor + H..Servicio + CU +
Adm. + Área..m2. + Parq.
                + D..Pie.media.m, data = datosf, nvmax = 10, method = "seqrep")
summary(modrs4)
# Entran paisaje, Porteria, Ascensor, HS, area, D..pie.media.m

# Mejores modelos -----

mod1 <- stepmc
mod2 <- stepmp
mod3 <- lm(Valor ~ Área..m2. + Portería + Ascensor + Parq. + D..Pie.media.m, data
= datosf)
mod4 <- lm(Valor ~ Paisaje + Área..m2. + Portería + Ascensor+ H..Servicio + Parq.
+ D..Pie.media.m, data = datosf)
mod5 <- lm(Valor ~ Paisaje + Área..m2. + Portería + H..Servicio + Parq. +
D..Pie.media.m, data = datosf)
mod6 <- lm(Valor ~ Paisaje + Área..m2. + Portería + H..Servicio + D..Pie.media.m,
data = datosf)
summary(mod3)
summary(mod4)
summary(mod5)
summary(mod6)
summary(mod1)
summary(mod2)
# Los modelos 3, 4 y 5 se deputan hasta el 6 que es igual al modelo 2
# Solo queden modelo 1 y 2
mod1s <- lm(Valor ~ Altura + Área..m2. + Edad + Portería + CU + Parq. +
D..Pie.media.m, data = datosf)
summary(mod1s)

# Análisis de supuestos -----

# Variables en modelo 1
datosmod1 <- subset(datosf, select = c(Valor, Altura, Área..m2., Edad, Portería, CU,
Parq.,

```

```

D..Pie.media.m, Distancia.Media.m, Distancia.Puntal.m))
summary(datosmod1)
numericasmod1 <- datosmod1[-c(5,6)]
numericasmod1

# Variables en modelo 2

datosmod2 <- subset(datosf, select = c(Valor, Paisaje, Portería, H..Servicio,
Área..m2., D..Pie.media.m))
summary(datosmod2)
numericasmod2 <- datosmod2[,c(1, 5, 6)]
head(numericasmod2)

# Independencia covariables

# Modelo 1

M1 <- cor(numericasmod1)
corrplot.mixed(M1, lower.col = "black")
mctest(numericasmod1[,-1], numericasmod1[,1])

# Modelo 2

M2 <- cor(numericasmod2)
corrplot.mixed(M2, lower.col = "black")
mctest(numericasmod2[,-1], numericasmod2[,1])

# Supuestos sobre los errores

#Modelo 1
resmod1 <- mod1$residuals
plot(mod1)
shapiro.test(resmod1)
acf(resmod1)
pacf(resmod1)
rsmo1 <- rstandard(mod1)
qqnorm(rsmo1)
qqline(rsmo1, col = "red")
plot(mod1$fitted.values, rsmo1, ylim = c(-3,3))
abline(0,0)
abline(h = c(-2, 2), col = "red")

#Modelo 2
resmod2 <- mod2$residuals

```

```
plot(mod2)
shapiro.test(resmod2)
acf(resmod2)
pacf(resmod2)
rsm2 <- rstandard(mod2)
qqnorm(rsm2)
qqline(rsm2, col = "red")
plot(mod1$fitted.values, rsm2, ylim = c(-3,3))
abline(0,0)
abline(h = c(-2, 2), col = "red")

rmod1 <- summary(mod1)
rmod2 <- summary(mod2)
rmod1$sigma
rmod2$sigma
```