



Incidencia de la distancia al Centro Comercial Viva Envigado en los valores comerciales de inmuebles residenciales en multifamiliares ubicados entre las calles 26sur y calle 40sur y entre las carreras 43A Avenida El Poblado y carrera 48 Avenida Las Vegas

Autores:

Lorena Strauch Salazar

Rubén Darío Rojas Restrepo

Institución Universitaria Esumer
Facultad de Estudios Empresariales y de Mercadeo
Medellín, Colombia
2019

Incidencia de la distancia al Centro Comercial Viva Envigado en los valores comerciales de inmuebles residenciales en multifamiliares ubicados entre las calles 26sur y calle 40sur y entre las carreras 43A Avenida el Poblado y carrera 48 Avenida Las Vegas

Autores:

Lorena Strauch Salazar

Rubén Darío Rojas Restrepo

Trabajo de grado presentado para optar al título de:

Especialista en Valoración Inmobiliaria

Tutor:

Alexander Zuluaga Correa.

Ingeniero Administrativo Especialista en Métodos y Técnicas de Valoración.

Línea de Investigación:

Marketing

Institución Universitaria Esumer

Facultad de Estudios Empresariales y de Mercadeo

Medellín, Colombia

2019

Incidence of the distance to Viva Envigado Shopping Center in the commercial values of residential properties located between calle 26sur and calle 40sur and between carreras 43A (Avenida El Poblado), and carrera 48 (Avenida Las Vegas)

Authors:

Lorena Strauch Salazar

Rubén Darío Rojas Restrepo

In partial fulfillment of the requirement for the award of

Especialista en Valoración Inmobiliaria

Consultant:

Alexander Zuluaga Correa

Ingeniero Administrativo Especialista en Métodos y Técnicas de Valoración.

Line of Research:

Marketing

Institución Universitaria Esumer

Facultad de Estudios Empresariales y de Mercadeo

Medellín, Colombia

2019

RESUMEN

El Centro Comercial Viva Envigado, debido a su magnitud, se ha convertido en una nueva centralidad urbana para el área Metropolitana del Valle de Aburrá y para otros Municipios cercanos. Mediante la aplicación de métodos hedónicos, y con el uso de modelos de regresión lineal múltiple, se analizó la incidencia de la distancia al Centro Comercial Viva Envigado en el valor comercial de los inmuebles multifamiliares localizados entre las calles 26sur y calle 40sur y entre las carreras 43A Avenida el Poblado y carrera 48 Avenida Las Vegas en el Municipio de Envigado, considerando en el análisis las distancia a pie, la distancia lineal (euclidiana) y la distancia en carro, junto con otras características físicas del inmueble. De acuerdo con los análisis realizados, el modelo que mejor explica el valor de los inmuebles incluye principalmente las variables área del inmueble, si tiene ascensor o no, el estrato socioeconómico, cantidad de parqueaderos y el estado de conservación. Se observa en los resultados que por cada 100m de distancia (a pie, lineal o en carro) al Centro comercial, la variación en el valor del inmueble está entre el 0.2% y el 0.4% y como esta variable tiene una mayor importancia comparadas con otras en el estudio.

Palabras clave: Distancia a pie, distancia lineal, incidencia de la distancia, valor comercial, centro comercial, precios hedónicos.

ABSTRACT

Viva Envigado Shopping Center has become a new metropolitan urban centrality and for nearby municipalities. Through the application of hedonic methods, by using multiple linear regression models, incidence of the distance to the Viva Envigado Shopping Center in the commercial value of multifamily real estate located between 26sur and 40sur streets and between 43A Avenida el Poblado and Carrera 48 Avenida Las Vegas in the Municipality of Envigado was analyzed, considering distance on foot, linear distance and distance by car, along with other physical characteristics of the property. According to the analyzes carried out, the model that best explains the value of the properties mainly includes the variables area of the property, whether it has an elevator or not, socioeconomic stratum, number of parking spaces and state of conservation of the property. It is observed in the results that for every 100m of distance (linear, on foot, or by road) to the Shopping Center, the variation in the value of the property is between 0.2% and 0.4%.

TABLA DE CONTENIDO

1	TÍTULO	14
2	TEMA.....	15
3	IDEA.....	16
4	OBJETO DE ESTUDIO	17
5	ANTECEDENTES.....	18
6	PLANTEAMIENTO DEL PROBLEMA.....	26
7	JUSTIFICACIÓN	30
8	OBJETIVO GENERAL	33
9	OBJETIVOS ESPECÍFICOS.....	34
10	ALCANCE	35
11	MARCO TEÓRICO.....	36
11.1	Definiciones	36
11.2	De la Teoría económica Urbana	38
11.3	El estudio de las variables	39
11.4	Teoría de la relación de la variable distancia y su incidencia en el valor de los inmuebles	41
12	MARCO METODOLÓGICO.....	44

13	DESARROLLO DE LA INVESTIGACIÓN	45
13.1	<i>Recolección de la información del mercado</i>	45
13.1.1	<i>Contacto con la población</i>	45
13.1.2	<i>Diseño de instrumentos para la recolección de información</i>	45
13.1.3	<i>Recolección de información</i>	51
13.2	<i>Modelos econométricos</i>	52
13.2.1	<i>Selección del modelo lineal</i>	52
13.2.2	<i>Comprobación de supuestos del modelo</i>	54
13.2.3	<i>Procesamiento de la información y preselección de las variables explicativas</i>	56
13.3	<i>Análisis de los datos de las ofertas recolectados en campo</i>	61
13.3.1	<i>Caracterización de las muestras</i>	61
13.3.2	<i>Estadística descriptiva variables preseleccionadas</i>	70
13.3.3	<i>Verificación de la distribución normal de las muestras</i>	71
13.4	<i>Modelaciones excluyendo datos atípicos</i>	74
13.4.1	<i>Modelaciones con distancia a pie</i>	75
13.4.1.1	<i>Modelaciones metodología paso a paso (Step Wise)</i>	80
13.4.1.2	<i>Modelo 1 - Modelación incluyendo variable distancia a pie</i>	83
13.4.2	<i>Modelaciones con distancia en carro</i>	89
13.4.2.1	<i>Modelaciones metodología paso a paso</i>	93

13.4.2.2	<i>Modelaciones incluyendo variable distancia en carro.</i>	98
13.4.3	<i>Modelaciones con distancia euclidiana (lineal).</i>	103
13.4.3.1	<i>Modelaciones metodología paso a paso.</i>	108
13.4.3.2	<i>Modelaciones incluyendo la variable distancia lineal Euclidiana.</i>	111
13.5	<i>Modelaciones con totalidad de los datos.</i>	116
13.5.1	<i>Modelaciones con distancia a pie.</i>	116
13.5.1.1	<i>Modelación metodología Paso a Paso.</i>	126
13.5.1.2	<i>Modelación incluyendo variable distancia a pie.</i>	131
13.5.2	<i>Modelaciones con distancia lineal.</i>	136
13.5.2.1	<i>Modelación metodología Paso a Paso.</i>	140
13.5.2.2	<i>Modelación incluyendo variable distancia lineal.</i>	145
13.5.3	<i>Modelaciones con distancia en carro.</i>	150
13.5.3.1	<i>Modelación metodología Paso a Paso.</i>	155
13.5.3.2	<i>Modelación Incluyendo Variable Distancia en Carro.</i>	160
13.5.4	<i>Resumen de las modelaciones.</i>	165
13.6	<i>Verificación de la incidencia de la distancia.</i>	169
14	CONCLUSIONES Y RECOMENDACIONES	173
15	REFERENCIAS	177

LISTADO DE TABLAS

Tabla 1. Centros Comerciales en Envigado.....	21
Tabla 2. Centros comerciales por clase en el Área Metropolitana del Valle de Aburrá	21
Tabla 3. Viviendas por estrato en Barrios El Portal, Jardines, San Marcos y Villagrande de Envigado.....	26
Tabla 4. Viviendas por estrato en Barrios Alcalá de Envigado.....	27
Tabla 5. Datos recolectados para cada muestra.....	46
Tabla 6. Caracterización de los datos recolectados. Fuentes de información	51
Tabla 7. Caracterización de los datos recolectados. Tipologías de vivienda	56
Tabla 8. Posibles variables explicativas	56
Tabla 9. Caracterización de los datos recolectados. Fuentes.....	61
Tabla 10. Caracterización de los datos recolectados. Estratos	61
Tabla 11. Caracterización de los datos recolectados. Vetustez	62
Tabla 12. Caracterización de los datos recolectados. Tipo de inmueble.....	62
Tabla 13. Caracterización de los datos recolectados. Nivel del piso.....	63
Tabla 14. Caracterización de los datos recolectados. Ascensor.	64
Tabla 15. Caracterización de los datos recolectados. Tipo de unidad (abierta/cerrada)	64
Tabla 16. Caracterización de los datos recolectados. Cantidad de habitaciones	64
Tabla 17. Caracterización de los datos recolectados. Número de baños completos	65
Tabla 18. Caracterización de los datos recolectados. Número de baños medios	65
Tabla 19. Caracterización de los datos recolectados. Balcón.....	65
Tabla 20. Caracterización de los datos recolectados. Estudio.....	65
Tabla 21. Caracterización de los datos recolectados. Poniente.	66
Tabla 22. Caracterización de los datos recolectados. Vista.....	66
Tabla 23. Caracterización de los datos recolectados. Cuartos útiles	67
Tabla 24. Caracterización de los datos recolectados. Tipo de parqueadero	67
Tabla 25. Caracterización de los datos recolectados. Cantidad de celdas de parqueo privadas para vehículos	68
Tabla 26. Caracterización de los datos recolectados. Acabado de pisos.....	68
Tabla 27. Caracterización de los datos recolectados. Tipo de cocina	69
Tabla 28. Caracterización de los datos recolectados. Estado de conservación de Fitto y Corvini	69
Tabla 29. Caracterización de los datos recolectados. Portería	69
Tabla 30. Caracterización de los datos recolectados. Zonas comunes	70
Tabla 31. Estadística descriptiva variables preseleccionadas.....	70
Tabla 32. Modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), incluyendo todas las variables	75
Tabla 33. Modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), metodología paso a paso.....	80
Tabla 34. Modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia a pie	83
Tabla 35. Tabla de correlación entre variables. Modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia a pie	87
Tabla 36. Modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras) incluyendo todas las variables.....	89

Tabla 37. Modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), metodología paso a paso	94
Tabla 38. Modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia en carro.....	99
Tabla 39. Modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras) incluyendo todas las variables.....	104
Tabla 40. Modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras), metodología paso a paso	109
Tabla 41. Modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia euclidiana	112
Tabla 42. Modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), incluyendo todas las variables	117
Tabla 43. Modelo de regresión lineal con distancia a pie (49 muestras luego de aplicar Bonferroni), incluyendo todas las variables	122
Tabla 44. Modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), metodología paso a paso.....	127
Tabla 45. Modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia a pie	132
Tabla 46. Modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), incluyendo todas las variables	137
Tabla 47. Modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), metodología paso a paso.....	141
Tabla 48. Modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia lineal	146
Tabla 49. Modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), incluyendo todas las variables.....	151
Tabla 50. Modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), metodología paso a paso	156
Tabla 51. Modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia lineal.....	161
Tabla 52. Resumen modelos paso a paso	166
Tabla 53. Resumen modelos distancia con todos los datos.....	167
Tabla 54. Resumen modelos distancia excluyendo datos atípicos	168
Tabla 55. Datos para verificación de la incidencia distancia	169
Tabla 56. Verificación incidencia distancia modelo 2	170
Tabla 57. Verificación incidencia distancia modelo 4	171
Tabla 59. Verificación incidencia distancia modelo 6	172

LISTADO DE GRÁFICOS

Gráfico 1. Localización Centro Comercial Viva Envigado.....	16
Gráfico 2. Localización Centros Comerciales del Valle de Aburrá	19
Gráfico 3. Oferta de Centros Comerciales por clase en el Valle de Aburrá.....	20
Gráfico 4. Plan de desarrollo 2008-2011 Zona 2 Envigado	27
Gráfico 5. Plan de desarrollo 2008-2011 Zona 8 Envigado	28
Gráfico 6. Aplicación para la recolección de información	45
Gráfico 7. Localización muestras	59
Gráfico 8. Rutas distancia a pie a Viva Envigado	60
Gráfico 9. Rutas distancia a pie a Viva Envigado	60
Gráfico 10. Histograma de frecuencias Valor del negocio.....	71
Gráfico 11. Diagrama de caja Valor del negocio	72
Gráfico 12. Histograma de frecuencias Valor del negocio excluyendo outliers	73
Gráfico 13. Diagrama de caja Valor del negocio excluyendo outliers	73
Gráfico 14. Matriz de correlación excluyendo outliers (44 muestras)	74
Gráfico 15. Comportamiento de los residuos del modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), incluyendo todas las variables	76
Gráfico 16. Gráfica de índices de influencia del modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), incluyendo todas las variables	77
Gráfico 17. Comportamiento de los residuos del modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), metodología paso a paso.....	81
Gráfico 18. Comportamiento de los residuos del modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia a pie	84
Gráfico 19. Gráfico de cuantiles del modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia a pie	85
Gráfico 20. Gráfico de índices de influencia con distancia a pie excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia a pie	86
Gráfico 21. Correlación del modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia a pie.....	88
Gráfico 22. Comportamiento de los residuos del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras) incluyendo todas las variables.....	90
Gráfico 23. Cuantiles del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), incluyendo todas las variables	91
Gráfico 24. Índices de influencia del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), incluyendo todas las variables.....	92
Gráfico 25. Comportamiento de los residuos del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), metodología paso a paso	95
Gráfico 26. Cuantiles del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), metodología paso a paso.....	96
Gráfico 27. Índices de influencia del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), metodología paso a paso	97
Gráfico 28. Comportamiento de los residuos del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia en carro	100

Gráfico 29. Cuantiles del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia en carro	101
Gráfico 30. Índices de influencia del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia en carro	102
Gráfico 31. Comportamiento de los residuos del modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras) incluyendo todas las variables	105
Gráfico 32. Cuantiles del modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras) incluyendo todas las variables	106
Gráfico 33. Índices de influencia del modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras) incluyendo todas las variables	107
Gráfico 34. Comportamiento de los residuos del modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras), metodología paso a paso	110
Gráfico 35. Comportamiento de los residuos del modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia euclidiana.....	113
Gráfico 36. Cuantiles del modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia euclidiana.....	114
Gráfico 37. Índices de influencia del modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia euclidiana	115
Gráfico 38. Comportamiento de los residuos del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), incluyendo todas las variables.....	118
Gráfico 39. Cuantiles del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), incluyendo todas las variables	119
Gráfico 40. Índices de influencia del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), incluyendo todas las variables	120
Gráfico 41. Comportamiento de los residuos del modelo de regresión lineal con distancia a pie (49 muestras luego de aplicar Bonferroni), incluyendo todas las variables	123
Gráfico 42. Cuantiles del modelo de regresión lineal con distancia a pie (49 muestras luego de aplicar Bonferroni), incluyendo todas las variables.....	124
Gráfico 43. Índices de influencia de regresión lineal con distancia a pie (49 muestras luego de aplicar Bonferroni), incluyendo todas las variables.....	125
Gráfico 44. Comportamiento de los residuos del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), metodología paso a paso	128
Gráfico 45. Cuantiles del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), metodología paso a paso.....	129
Gráfico 46. Índices de influencia del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), metodología paso a paso	130
Gráfico 47. Comportamiento de los residuos del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia a pie	133
Gráfico 48. Cuantiles del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia a pie.....	134
Gráfico 49. Índices de influencia del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia a pie.....	135
Gráfico 50. Comportamiento de los residuos del modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), incluyendo todas las variables.....	138
Gráfico 51. Cuantiles del modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), incluyendo todas las variables	139

Gráfico 52. Comportamiento de los residuos del modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), metodología paso a paso	142
Gráfico 53. Cuantiles del modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), metodología paso a paso	143
Gráfico 54. Índices de influencia del modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), metodología paso a paso	144
Gráfico 55. Comportamiento de los residuos del modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia lineal.....	147
Gráfico 56. Cuantiles de regresión lineal con distancia lineal incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia lineal.....	148
Gráfico 57. Índices de influencia del modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia lineal.....	149
Gráfico 58. Comportamiento de los residuos del modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), incluyendo todas las variables	152
Gráfico 59. Cuantiles del modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), incluyendo todas las variables	153
Gráfico 60. Índices de influencia modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), incluyendo todas las variables	154
Gráfico 61. Comportamiento de los residuos del modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), metodología paso a paso	157
Gráfico 62. Cuantiles del modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), metodología paso a paso.....	158
Gráfico 63. Índices de influencia del modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), metodología paso a paso	159
Gráfico 64. Comportamiento de los residuos del modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia lineal	162
Gráfico 65. Cuantiles del modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia lineal	163
Gráfico 66. Índices de influencia de los residuos del modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia lineal	164

1 TÍTULO

Incidencia de la distancia al Centro Comercial Viva Envigado en los valores comerciales de inmuebles residenciales en multifamiliares ubicados entre las calles 26sur y calle 40sur y entre las carreras 43A Avenida el Poblado y carrera 48 Avenida Las Vegas

Palabras clave: Distancia a pie, distancia media, incidencia de la distancia, valor comercial, efecto espacial, centro comercial.

2 TEMA

Las áreas del conocimiento que se abordarán serán:

Economía, Administración, Contaduría y Afines: Se trata el tema valoración inmobiliaria que implica teoría económica asociada a la determinación de precios de venta.

Ingeniería, Arquitectura, Urbanismo y Afines: Se revisa la incidencia de la distancia al Centro Comercial VIVA Envigado, desde el punto de vista de las tipologías constructivas.

Matemáticas y Ciencias Naturales: Se utilizan técnicas estadísticas para establecer el modelo de regresión espacial de los valores de vivienda.

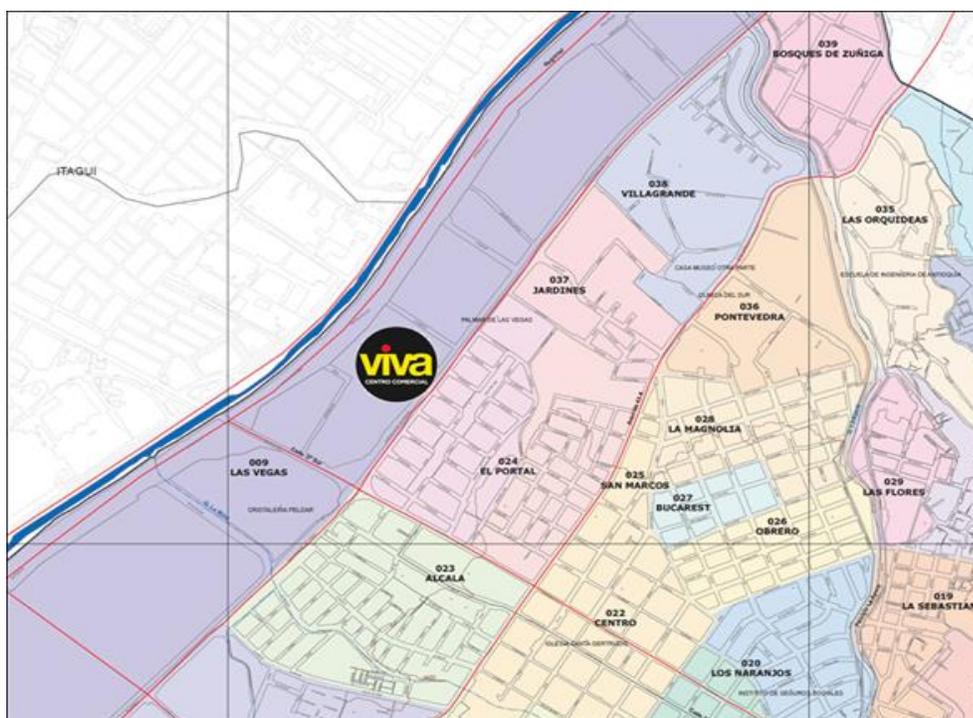
3 IDEA

Existen diversos estudios que establecen la incidencia del valor con relación a una variable explicativa. Viva Envigado es el centro comercial más grande del país, que a su vez tiene una conexión directa con el sistema de transporte masivo Metro y específicamente a la estación Envigado, condiciones que generan una centralidad para los municipios cercanos al municipio de Envigado y otros fuera del área Metropolitana del Valle de Aburrá.

Algunos estudios han encontrado que existe una relación directa o inversa entre la distancia y un punto de referencia.

Con el presente proyecto de investigación se estudia la relación que existe entre el valor de un inmueble residencial en multifamiliares y la distancia al Centro Comercial Viva Envigado, con el fin de determinar la incidencia que tiene esta sobre el valor de las viviendas multifamiliares ubicadas entre las calles 26sur y calle 40sur y entre las carreras 43ª, Avenida el Poblado, y carrera 48, Avenida Las Vegas.

Gráfico 1. Localización Centro Comercial Viva Envigado



Fuente: POT Acuerdo 10 2011 (Municipio de Envigado, 2011)

4 OBJETO DE ESTUDIO

Establecer, con la información de mercado recolectada entre el 28 de septiembre y el 15 de octubre de 2019, si se puede concluir si la variable distancia al centro comercial Viva Envigado tiene un comportamiento directo o inverso con relación a los valores de los inmuebles residenciales.

5 ANTECEDENTES

Historia de los centros comerciales en Colombia

La historia de los centros comerciales en el país comienza en el año 1972 con la construcción del centro comercial San Diego en la ciudad de Medellín y cuatro años después con la construcción del Centro Comercial Unicentro en la ciudad de Bogotá (Muñoz Echavarría, 2016). Para la época los centros comerciales enunciaban unas ventajas como “efectuar todas las compras en un solo sitio, y un gran almacén de departamentos, supermercado, cafetería, discoteca, auditorio, restaurante, servicios bancarios” (Muñoz Echavarría, 2016) y adicionalmente “sin peligro de tráfico automotor y sin congestión de peatones” (Muñoz Echavarría, 2016) en el cual “el público podría hacer sus compras confortablemente, transitando por amplias aceras cubiertas, que lo protegerán del sol y la lluvia.” (Muñoz Echavarría, 2016).

Análisis del mercado de los centros comerciales

De acuerdo con el artículo del portal America Retail, en Colombia actualmente hay 3,2 millones de metros cuadrados construidos en superficies de retail, de los cuales 891.870 metros cuadrados están en el Valle de Aburrá y el valle de San Nicolás en Rionegro (Suárez L., 2019).

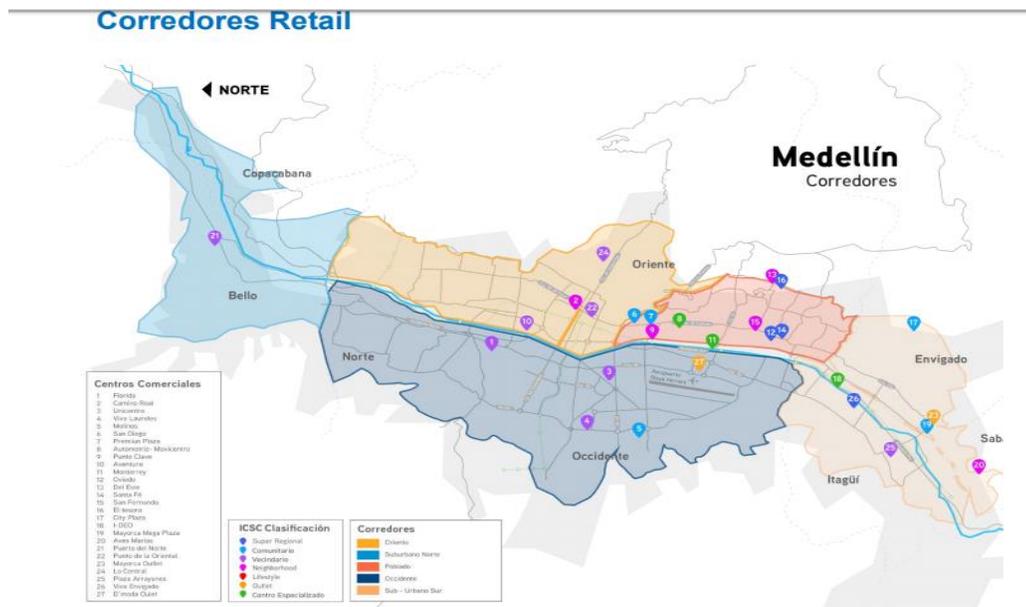
En el mismo artículo señalan que en el año 2018 se abrieron catorce centros comerciales en el territorio nacional, de los cuales tres fueron en Antioquia: Viva Envigado (137.000 m²), La Central en el barrio Buenos Aires (86.000m²) y D’ Moda Outlet (50.000m²). Así mismo establece que en lo que va de 2019 se hizo la apertura en Rionegro del mall Jardines de Llano Grande, con un área de 33.000 metros cuadrados. (Suárez L., 2019).

En el informe de Colliers de marzo de 2019, respecto del mercado del retail para la ciudad de Medellín para el primer trimestre de 2019, se observa que en la ciudad hay un inventario de 891.800 metros cuadrados de área rentable, presentándose un incremento del 27% al inicio del año comparado con el mismo trimestre del año inmediatamente anterior (2018), situación que se da especialmente por el ingreso a finales de 2018 del centro comercial D’ Moda Outlet, que brinda un área adicional para la ciudad de 12.000 metros cuadrados para este mercado. (Colliers International, 2019).

Adicionalmente realizan un análisis por clases en el cual encuentran que los centros comerciales Súper Regionales, caso Viva Envigado, fueron los que registraron un mayor crecimiento, (86%) y concluyendo que la mayor concentración de área libre alquilable (GLA) de la ciudad se encuentra en esta clase de centros comerciales (Colliers International, 2019). Se destaca además que en el año 2018 nuevos centros comerciales iniciaron operaciones en el área metropolitana, como fueron La Central en el barrio Buenos Aires y el centro comercial Plaza Arrayanes en el municipio de Itagüí (Colliers International, 2019).

De acuerdo con la multinacional Colliers, los centros comerciales denominados de “Clase Súper Regional” tienen características similares a las de los centros regionales, pero cuenta con: una mayor área para renta, representada normalmente en más de tres locales tipo ancla, que brindan mayores opciones de mercancía, y que tienen a su vez un mayor poder de atracción de clientes ubicados tanto a nivel zonal, como en otros municipios del área metropolitana (Colliers International, 2019).

Gráfico 2. Localización Centros Comerciales del Valle de Aburrá



Fuente: (Colliers International, 2019)

Gráfico 3. Oferta de Centros Comerciales por clase en el Valle de Aburrá



Fuente: (Colliers International, 2019)

En cuanto al mercado venidero, los expertos de Colliers International anuncian que en los próximos tres años desarrollos de centros comerciales en el área metropolitana de la ciudad de Medellín estarán representados en la construcción de alrededor de 173.000 metros cuadrados y en ampliaciones de algunos centros comerciales tradicionales, los cuales están ubicados especialmente en el occidente de la ciudad (39%), en la zona suburbana del norte (38%) y en la comuna del Poblado (24%), lo que representa un incremento del 2% del inventario para la región. (Colliers International, 2019).

El mercado de los centros comerciales comienza a tener cambios enfocados en los modelos de único propietario en los que un tercio del área neta arrendable corresponde a anclas de entretenimiento, priorizando entre otras, la gastronomía y los servicios experienciales. (Colliers International, 2019).

También muestran en su informe que uno de los fondos de inversiones más conocidos en el país compró en el año 2019 la mayoría de la participación de un centro comercial tipo outlet, buscando con ello realizar una expansión en las principales ciudades de un nuevo modelo en el cual combinan la comercialización en las tiendas físicas con el comercio electrónico (Colliers International, 2019), lo cual muestra como el sector comercial comienza a adaptarse a los cambios comportamentales de las sociedades dados especialmente por las nuevas formas de consumo y la tecnología.

Hasta octubre de 2018, Envigado contaba únicamente con el Centro Comercial City Plaza con 24.800 m2 de comercio. A partir de la apertura del Centro Comercial Viva Envigado, se cuenta con un total de 161.800 m2. Así el grado de saturación comercial pasó de 12 a 76 m2 por cada 100 habitantes en Envigado, teniendo en cuenta una población de 212.365 habitantes, resultado del censo 2018 realizado por el Dane.

Tabla 1. Centros Comerciales en Envigado

Centro comercial	Área comercial (m2)
Cityplaza	24800 ¹
Viva Envigado	137000 ²
Total	161800

¹ (Ospinas & Cia S.A., 2017)

² (La República, 2018)

El Área Metropolitana del Valle de Aburrá cuenta con 26 centros comerciales que suman un total de 879.870 m2, con un grado de saturación de 24 m2 por cada 100 habitantes, teniendo en cuenta una población de 3.725.682 habitantes, resultado del censo realizado por el Dane en 2018.

Tabla 2. Centros comerciales por clase en el Área Metropolitana del Valle de Aburrá

Clase	No centros comerciales	m2 comercio
Regional	5	259937
Comunitario	8	185542
Super regional	4	301871
Outlet	2	39931
Centro especializado	3	53163
Vecindario	5	51426
Area total comercio	27	891870

Fuente (Colliers International , 2019)

En Colombia el grado de saturación es de 11 m² por cada 100 habitantes, que es bajo comparado con países como Chile (22 m² por cada 100 habitantes) o México (18 m² por cada 100 habitantes), o Estados Unidos con 2000 m² por cada 100 habitantes, de acuerdo con las cifras reportadas por la revista Semana. (Semana, 2018).

Por lo anterior, se observa que el mercado de los centros comerciales continúa presentando inversiones importantes en Colombia y en la ciudad de Medellín y municipios del área metropolitana, con las cuales busca transformarse y adaptarse a las condiciones actuales y mantener así, los niveles de ocupación históricos.

Los centros comerciales en el imaginario de ciudad

Respecto al impacto de los centros comerciales en el urbanismo e imaginario de la ciudad, el sociólogo Jhon Muñoz Echavarría, Magister en Estudios Urbanos de la Universidad Nacional de Colombia, en su estudio “Los centros comerciales en la reconfiguración territorial de la ciudad de Medellín y los nuevos imaginarios urbanos 1990 – 2011” (Muñoz Echavarría, 2016) muestra como la relación entre los habitantes de la ciudad y los centros comerciales crearon una fragmentación de la ciudad, generando un nuevo modo de vida enfocado en el consumo, transformando y desplazando con ello las grandes empresas de producción industrial, y consiguiendo atraer a los nuevos consumidores a estos espacios. (Muñoz Echavarría, 2016). Sin embargo, con los avances tecnológicos de las últimas décadas que han modificado el comportamiento de los ciudadanos y especialmente las formas de compra y el comercio en general, el reto de los centros comerciales es transformarse en centros de servicios, con tendencia a ser de único propietario, que brinden experiencias a sus clientes y visitantes, y continuar siendo puntos de referencia importantes en la ciudad.

El comportamiento de las personas va cambiando en el tiempo, y de la misma manera los centros comerciales se van adaptando a esos nuevos comportamientos empujados especialmente por la tecnología y las condiciones del mercado. Las condiciones de tráfico de la ciudad también inciden en el comportamiento de los ciudadanos, haciendo que las personas valoren los inmuebles por la ubicación en la que se encuentren. Es por estos cambios comportamentales que, con el fin de combatir las vacancias en los centros comerciales, el negocio del retail se está adaptando a las nuevas condiciones del mercado,

pasando de la venta de productos a la prestación de servicios, donde priman las experiencias para los clientes.

De acuerdo con el ingeniero Jorge García Duque, experto con más de 30 años en el sector de los centros comerciales, los centros comerciales pasaron a ser “«centros de ciudad» con capacidad de ofrecer múltiples servicios y de cubrir los requerimientos del público” (La Nota Económica, 2019).

García Duque expresa adicionalmente que si bien hace unos 200 años las ciudades crecían teniendo como punto de referencia las iglesias y parques, en la actualidad lo hacen alrededor de los centros comerciales (La Nota Económica, 2019). Una evidencia de este hecho es observable en proyectos como Plaza Fabricato en el municipio de Bello, el cual hace parte de un proyecto urbanístico llamado Ciudad Fabricato, que consta además del centro comercial, de más de 2.000 apartamentos distribuidos en cuatro complejos residenciales y en los que se espera habiten más de 8.000 personas, tal como lo describen en el artículo del periódico La República del 29 de marzo de 2019 (La República, 2019) y en el que termina el patrocinador de la nota afirmando que “Plaza Fabricato no solo será una opción de entretenimiento, comercio y diversión para todos los habitantes de Medellín, Bello y sus alrededores, sino que brindará un desarrollo en infraestructura y cuidado de medio ambiente que pocas ciudades del país tienen y que abre las puertas a los turistas a una amplia oferta de servicios” (La República, 2019). De esto se colige que tendrá un impacto en el entorno del sector y por lo tanto, en alguna medida, una posible influencia en el valor de las viviendas circundantes, que debería ser medida.

Por lo tanto los centros comerciales tienen un impacto en el urbanismo y en los procesos de reconfiguración de la ciudad, como se ha observado en la historia de la misma, y por lo tanto hay una necesidad de medir los impactos que estos tienen sobre el entorno, tal como concluye y recomienda el Sociólogo Jhon Muñoz Echavarría en su estudio de tesis de maestría en estudios Urbanos, en la cual establece la necesidad de estudiar “la participación de los centros comerciales en la reorganización funcional-espacial de la ciudad, su relación con la provisión de infraestructura vial y de transporte, **incidencia en el mercado inmobiliario y efectos generales sobre las áreas circundantes** (negrilla nuestra)” (Muñoz Echavarría, 2016).

Estudios realizados en Colombia y en el mundo.

En Colombia se han realizado estudios como el de Vargas y Caicedo que analiza la influencia del Centro Comercial AltaVista de Usme en los precios de inmuebles cercanos bajo la metodología de diferencias en diferencias, comparando esta zona con la zona de San Cristóbal, de características similares pero que no fue afectada por una construcción de un centro comercial. Para este estudio se tuvo en cuenta la información catastral disponible en Bogotá para los años 2007 y 2009. (Vargas & Caicedo, 2013).

Otro caso estudiado en Colombia corresponde al realizado por Franco para obtener el título de Especialista en la Universidad Distrital Francisco José de Caldas, en el cual analiza el comportamiento del valor comercial de inmuebles residenciales sometidos a propiedad horizontal en el Municipio de Soacha, en un sector con centros comerciales y almacenes de gran superficie, encontrando que hay un mayor porcentaje de incremento en los valores comerciales de los conjuntos residenciales cuando éstos están más cercanos ya sea a centros comerciales o a almacenes de gran superficie. (Franco Polanco, 2018).

Otras metodologías utilizadas para evaluar la influencia de la cercanía de una centralidad o equipamiento en los precios de la vivienda cercana, han sido la de precios hedónicos espaciales y la metodología de Propensity Score Matching. En Colombia, así lo ha hecho Perdomo, quien con base en estas metodologías realizó estudios sobre el valor de los inmuebles cercanos a una Estación de Transmilenio en Bogotá (Perdomo, 2010).

A través del método analítico de Propensity Score Matching, Perdomo encontró que las propiedades son 13% más económicas comparadas con las que se encuentran dentro del área de influencia (500 metros) del sistema de transporte masivo TransMilenio, y por lo tanto, mientras más alejada esté la propiedad del sistema de transporte masivo, su valor por metro cuadrado disminuye en 0.05%. (Perdomo, 2010).

Otros autores como Pope y Pope encontraron que bajo la metodología de diferencias en diferencias, una nueva tienda de Walmart podría incrementar los precios de la vivienda entre un 2% y 3% para las viviendas localizadas a 0.5 millas y entre 1% y 2% para las viviendas localizadas a una distancia entre 0,5 y 1 milla de las tiendas. Este análisis fue realizado con base en alrededor de 1 millón de transacciones realizadas alrededor de 159 Walmarts en Estados Unidos. (Pope & Pope, 2014).

En el caso de estudio del centro comercial The West Intime en Hangzhou China, los autores encontraron a través de la aplicación de la teoría de precios hedónicos y de la aplicación del método de gradiente de precio, que hubo un efecto positivo en el valor de las viviendas en las cuadras cercanas y que al incrementar la distancia desde el mall, dicho efecto positivo disminuye y que por lo tanto éstos pueden generar nuevos puntos de desarrollo de las ciudades (Zhou, Zhang, Hui, & Wen, 2019).

En Suráfrica, el Dr. Sale desarrolló un estudio de caso de correlación entre la cercanía del centro comercial Walmer Park, localizado en la Bahía Nelson Mandela, y el precio de las propiedades residenciales a su alrededor, mediante la aplicación del modelo de precios hedónicos, encontrando en su estudio una correlación significativa entre la distancia al mall comercial y el precio de las propiedades. (Sale, 2015). También concluyó que una ubicación cercana al centro comercial pesa más que los efectos negativos generados por el tráfico, la contaminación ambiental y el ruido. (Sale, 2015).

En el Estado de Texas, en Estados Unidos, desarrollaron un estudio de correlación, en el cual evaluaron la cercanía de los desarrollos comerciales y su efecto en el valor de las viviendas. Con esto pretendían, especialmente los constructores, conocer el comportamiento de ésta variable y la afectación en las áreas circundantes, al tenerse como premisa que la cercanía tenía unos efectos negativos en las viviendas residenciales tales como el ruido, la luz, el tráfico, la disminución estética, y por lo tanto la legislación del estado preveía dejar parte de los ingresos para minimizar dicho efectos negativos. Sin embargo, encontraron en el estudio que los efectos positivos eran aún mayores a los negativos, inclusive en los sitios más cercanos al desarrollo comercial (Aydin, Crawford, & Smith, 2011).

En la investigación realizada por el Dr. Kwame Addae-Dapaah y Yu Siew Lan, respecto a “la relación entre los centros comerciales y el precio de las propiedades residenciales vecinas”, haciendo uso de un análisis hedónico, encontraron que “a pesar de las externalidades negativas de los centros comerciales, las propiedades residenciales dentro de un radio de 100 metros de los centros comerciales cuentan con una prima más alta que las que están más alejadas”, resultados que según los autores expresan, pueden ser de gran interés para inversionistas y para temas de políticas públicas. (Addae-Dapaah & Lan, 2010).

6 PLANTEAMIENTO DEL PROBLEMA

El centro comercial VIVA Envigado con cerca de 260.000 metros cuadrados de área construida, de los cuales 137.000 corresponden a espacios comerciales para 394 locales y concesionarios, abrió sus puertas el 5 de octubre de 2018 en el municipio de Envigado (La República, 2018), en el sitio que donde se encontraba localizado el ÉXITO de Envigado, y el cual había sido inaugurado en 1981.

Viva Envigado incluye zonas diferenciales en el segmento de centros comerciales como un parque de diversiones, un bulevar de 21 restaurantes con concepto urbano a cielo abierto que complementará la plazoleta de comidas, un market gourmet, un complejo deportivo y un lugar que se enfoca en el desarrollo artístico y tiene su propio teatro.

El centro comercial VIVA Envigado, acorde con el artículo del diario La República, tiene un potencial de visitantes de 3.1 millones de personas, considerando los habitantes de los municipios del área metropolitana tales como Envigado, Medellín, Sabaneta, Itagüí, Caldas y La Estrella. (La República, 2018).

Con esta amplia oferta de comercio y servicios es de esperar que se presente una incidencia en los valores de la vivienda en la zona de influencia directa, de acuerdo con la distancia al centro comercial.

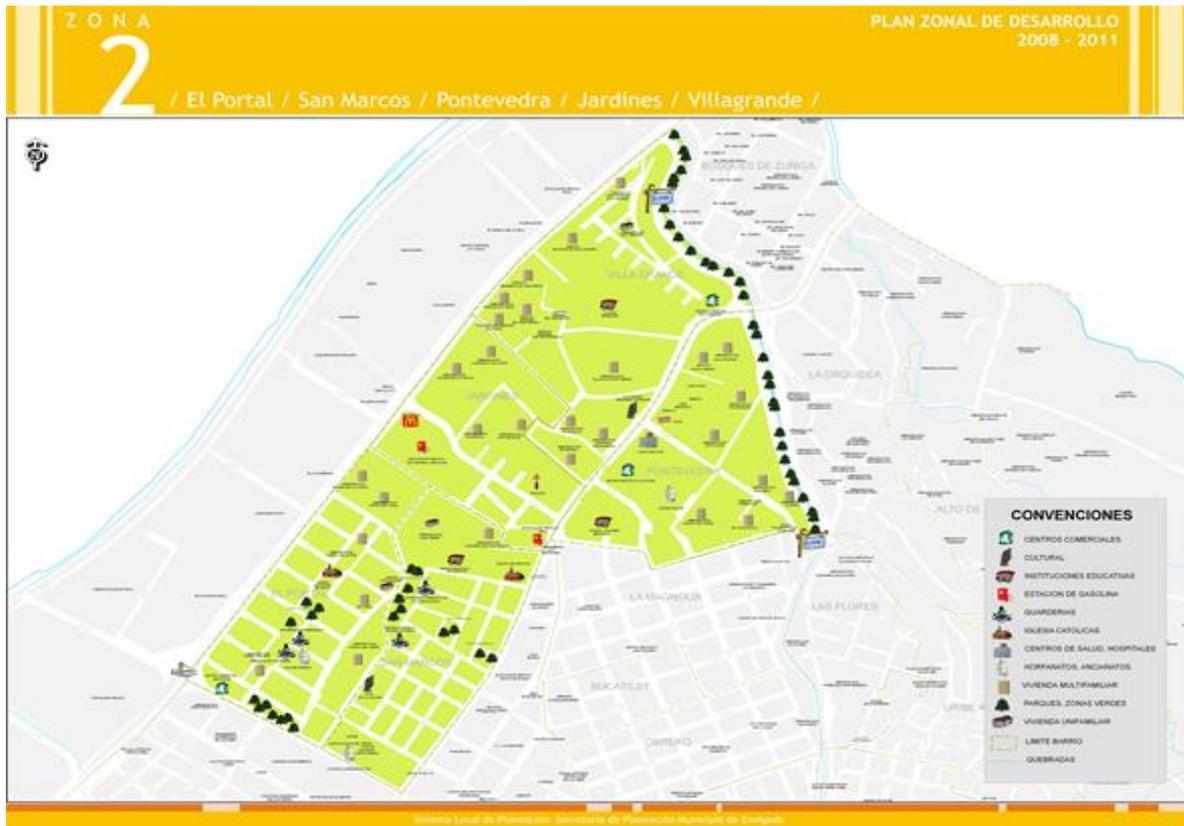
La zona de influencia del Centro Comercial VIVA Envigado está conformada por los barrios El Portal, Jardines, San Marcos y Villagrande, que hacen parte de la Zona 2 de Envigado. A continuación, se presenta el número de viviendas por estrato de estos barrios y el plano del Plan Zonal de Desarrollo 2008-2011 de dicha zona. (Municipio de Envigado, 2017).

Tabla 3. Viviendas por estrato en Barrios El Portal, Jardines, San Marcos y Villagrande de Envigado

Zona	Barrio	Estrato 2	Estrato 3	Estrato 4	Estrato 5	Estrato 6	Total
2	El Portal			487	885	97	1469
	San Marcos	1	174	768	64	0	1007
	Jardines		159	143	966	0	1268
	Villagrande	1	3	174	1614	0	1792
	Total	1	333	1398	1915	97	3744

Fuente: (Departamento Administrativo de Planeación 2016, citado por Municipio de Envigado, 2017)

Gráfico 4. Plan de desarrollo 2008-2011 Zona 2 Envigado



Fuente: (Municipio de Envigado, 2017)

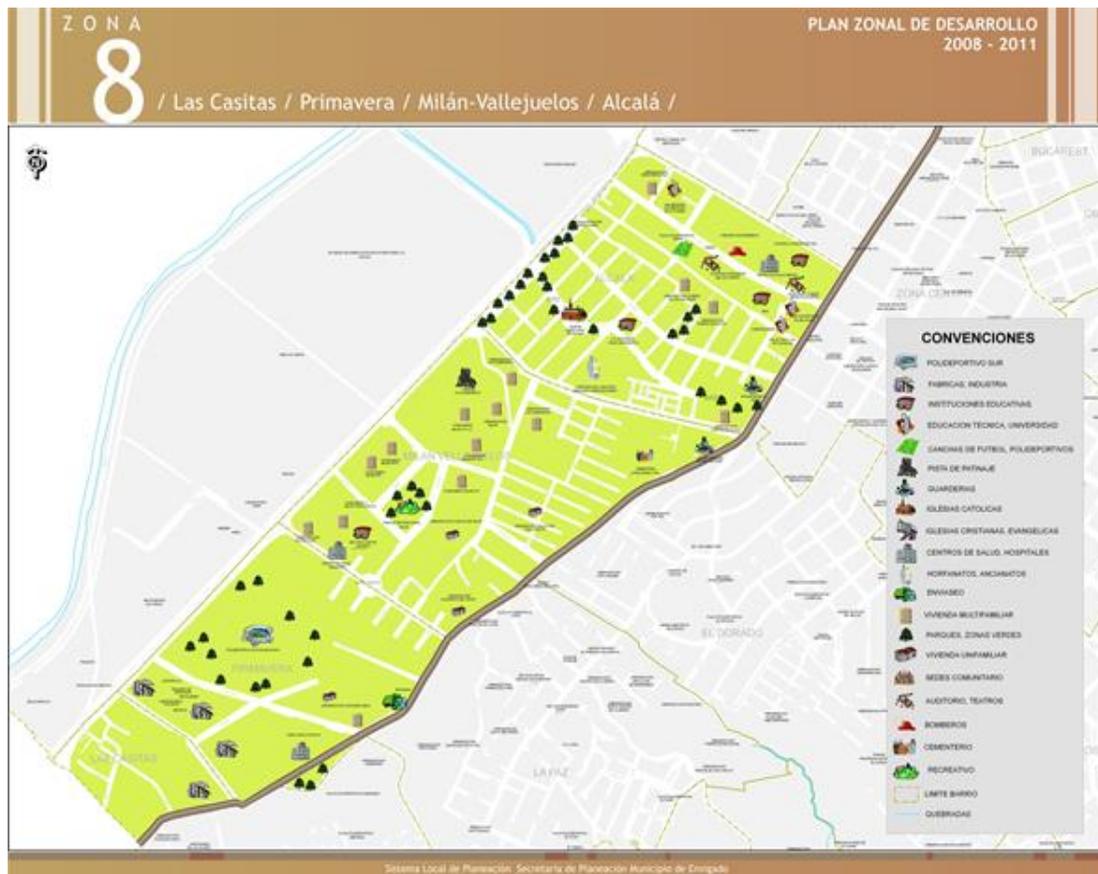
Adicionalmente, el Barrio Alcalá también se incluyó como parte de la zona de estudio. En la tabla a continuación se muestra el número de viviendas por estrato de Alcalá y el plano del Plan Zonal de Desarrollo 2008-2011 (Municipio de Envigado, 2017) de dicha zona:

Tabla 4. Viviendas por estrato en Barrios Alcalá de Envigado

Zona	Barrio	Estrato 2	Estrato 3	Estrato 4	Estrato 5	Estrato 6	Total
8	Alcalá	3	1597	273			1873

Fuente: (Departamento Administrativo de Planeación 2016, citado por Municipio de Envigado, 2017)

Gráfico 5. Plan de desarrollo 2008-2011 Zona 8 Envigado



Fuente: (Municipio de Envigado, 2017)

La valoración inmobiliaria tiene en cuenta atributos de los bienes y, en metodologías como el método de mercado, hace uso del principio de sustitución, el cual se basa en que un comprador no pagará más por un bien que puede ser sustituido por otro, lo cual requiere que el inmueble comparable tenga características físicas, legales y de entorno similares al inmueble objeto de estudio. Una de las características más importantes que define el valor de un bien inmueble es su localización. Al realizar avalúos, es usual que esta variable sea analizada cualitativamente y/o subjetivamente.

De acuerdo con lo anteriormente expuesto, surgieron las siguientes preguntas a responder con la investigación:

¿Existe incidencia de la distancia al Centro Comercial Viva Envigado en los valores comerciales de los inmuebles residenciales ubicados entre las calles 26sur y calle 40sur y entre las carreras 43^a, Avenida el Poblado, y la carrera 48 Avenida Las Vegas?

¿Es posible explicar la incidencia de la distancia al centro comercial en los valores de los inmuebles residenciales a través de un modelo o una ecuación?

¿El uso de metodologías espaciales podría facilitar la determinación del modelo o ecuación?

Para abordar el problema, se revisó inicialmente con qué información se podía contar:

- Datos de mercado de ofertas
- Información catastral. La última actualización catastral de Envigado fue realizada en el año 2014.
- Plan de Ordenamiento Territorial
- Encuesta de calidad de vida.

A partir de la información disponible y con la ayuda de sistemas de información geográfica (QGis) se determinó la ubicación y distancia (distancia a pie, distancia en carro y distancia media) con respecto al Centro Comercial.

Finalmente se realizaron regresiones multivariadas para encontrar un modelo, teniendo en cuenta la información disponible y el fenómeno que se pretendía estudiar.

7 JUSTIFICACIÓN

Los Centros Comerciales son y seguirán siendo un punto de referencia en el desarrollo de las ciudades, impactando en las costumbres de los habitantes, en el desarrollo urbanístico de las mismas y de los sectores donde están ubicados, en los sistemas e infraestructura de transporte y en el tráfico de las zonas aledañas. En ese interactuar con el entorno tienen un impacto en los inmuebles residenciales a su alrededor, lo que podría llegar a incidir en su valor comercial en función de su distancia.

La actividad de la valoración inmobiliaria requiere del uso de metodologías que permitan establecer el valor más probable por el cual se puede transar un inmueble en el mercado, siendo indispensable conocer el comportamiento del mismo, al ser éste quien establece las variables explicativas del valor dentro del proceso valuatorio. Es por ello que en los inmuebles residenciales, una de las variables explicativas más comúnmente encontradas es la localización espacial del inmueble. Por lo tanto, es importante explicar el valor que le da el mercado a un inmueble por su localización respecto a un punto de importancia en el desarrollo de las ciudades, como es el caso de un Centro Comercial.

En el sector inmobiliario es normal encontrar en las ofertas la alusión que el bien se encuentra cerca de un centro comercial, o de un mall comercial o de un almacén de grandes superficies, por lo tanto es una característica que es percibida como algo positivo en el mercado. A pesar de lo anterior, no se encuentran estudios que indiquen o cuantifiquen el peso que se le da a esta variable, por ende es pertinente para el sector inmobiliario determinar en qué magnitud esa cercanía, medida a través de la variable distancia, incide en el valor de un inmueble residencial.

El análisis del mercado es fundamental para la aplicación de las metodologías de valoración inmobiliaria y para ello es indispensable contar con inmuebles sustitutos que permitan evaluar dicho mercado de una manera homogénea, y por lo tanto los resultados de la investigación podrían ser un soporte para la determinación de los bienes comparables, basados en la ubicación del inmueble residencial con referencia al centro comercial.

Socialmente el resultado de la investigación permite que las personas tengan una herramienta para el análisis de inversión y, como comunidad, a determinar el posible impacto que un proyecto inmobiliario de estas magnitudes puede tener en el valor de sus

viviendas, al ser éstas parte del patrimonio familiar. Adicionalmente, es un resultado útil para constructores e inversionistas de este tipo de proyectos como herramienta o insumo para el análisis de sus impactos y el área de influencia de los mismos.

Hoy en día prima el mundo digital, el acceso al conocimiento y a la información. Esta generación está basada en el conocimiento, en la cual la tecnología es un apoyo para las personas y las empresas, creando nuevas experiencias y ventajas competitivas con el fin de facilitar la vida, mejorar formas de comunicación y la practicidad en los procesos de la vida diaria. Los cambios tecnológicos y en la sociedad han sido apreciables en los últimos 20 años con la masificación de las tecnologías de la información y comunicación (TIC) generado cambios en la sociedad y su comportamiento, e influyendo en la creciente y vertiginosa generación de datos e información que comienzan a ser y serán la base para la toma de decisiones. Es importante, entonces aplicar estas tecnologías al campo de la valoración inmobiliaria y contribuir a su masificación.

En la actualidad las metodologías de valoración pueden ser agrupadas en dos clases: en los métodos clásicos y tradicionales y en los métodos avanzados de valoración, tal como lo expresa el artículo Real estate appraisal: a review of valuation methods (Pagourtzi, Assimakopoulos, Hatzichristos, & French, 2018) donde las metodologías avanzadas usan en general modelos de valoración haciendo uso de la inteligencia artificial, como es el caso de las redes neurales, o de estimaciones hedónicas y de tecnología GIS como es el caso de los métodos de análisis espacial. (Pagourtzi, Assimakopoulos, Hatzichristos, & French, 2018).

Teniendo en cuenta todo este marco tecnológico que se está presentando en la actualidad a nivel mundial y local, se avecina una transformación de la actividad valuatoria. Es por ello que se espera que con el tiempo se ingresen otros componentes al mercado y se consoliden nuevas tecnologías y metodologías de valoración avanzada de inmuebles basados en estos modelos en computadora (Automated Valuation Models), que requerirán entre otras de información espacial.

De igual manera esta investigación permite explorar el mercado inmobiliario y conocer a mayor profundidad nuevas metodologías de valoración inmobiliaria que permitan mejorar los procesos valuatorios en el día a día de la profesión, así como buscar herramientas para

desarrollar nuevas tecnologías, que son la tendencia en el presente y el futuro de nuestra sociedad.

8 OBJETIVO GENERAL

Conocer cómo incide la distancia al Centro Comercial Viva Envigado en los valores comerciales de los inmuebles residenciales ubicados entre las calles 26sur y calle 40sur y entre las carreras 43A Avenida el Poblado y carrera 48 Avenida Las Vegas, a través del análisis de mercado actual.

9 OBJETIVOS ESPECÍFICOS

- Recolectar la información del mercado entre las calles 26sur y calle 40sur y entre las carreras 43ª, Avenida el Poblado, y carrera 48 Avenida Las Vegas.
- Analizar los datos del mercado recolectados en campo, clasificando las variables explicativas de valor.
- Realizar las modelaciones de la información recolectada y analizada, que permita establecer un modelo econométrico con el cual se encuentre el valor comercial de las viviendas.
- Concluir si existe incidencia de la distancia al Centro Comercial Viva Envigado en los valores de los inmuebles residenciales, realizando simulaciones en el modelo econométrico

10 ALCANCE

Desde un enfoque correlacional, la presente investigación pretende conocer la relación o grado de asociación que existe entre la distancia de un inmueble residencial ubicado entre las calles 26sur y calle 40sur y entre las carreras 43A Avenida el Poblado y carrera 48 Avenida Las Vegas al Centro Comercial Viva Envigado, y el valor comercial del inmueble.

El grado de asociación permitirá estimar el valor de un inmueble en la zona de estudio a partir de su localización frente al centro comercial.

El alcance de este estudio se encuentra limitado por la obtención de datos de características y precios de inmuebles.

11 MARCO TEÓRICO

En el presente capítulo desarrollaremos las bases teóricas sobre las cuales se sustenta la investigación de la Incidencia de la distancia al Centro Comercial Viva Envigado en los valores comerciales de inmuebles residenciales ubicados entre las calles 26sur y calle 40sur y entre las carreras 43A Avenida el Poblado y carrera 48 Avenida Las Vegas, partiendo inicialmente de las bases de la teoría económica urbana, para luego revisar las variables que intervienen en el valor de los inmuebles, y finalizar abordando las teorías utilizadas en diferentes investigaciones, en las que toman la variable distancia a un punto específico (parque, sistema de transporte, áreas comerciales, etc.) como crítica para determinar su posible efecto en el valor de los inmuebles.

11.1 Definiciones

Valor de mercado: Según María del Pilar García, “se refiere al precio más probable por el cual un inmueble se venderá en un mercado caracterizado por un funcionamiento económicamente racional entre oferta y demanda” (García, 2007, pág. 14).

Precio: María del Pilar García, indica que “se refiere a la cantidad de dinero que retribuye una transacción concreta, por lo que depende del pacto entre comprador y vendedor”. (García, 2007, pág. 14).

Modelos econométricos: De acuerdo con Luis Quintana Romero y Miguel Ángel Mendoza, “los modelos econométricos son una simplificación de la realidad que se compone de relaciones entre variables. Dichas relaciones son no exactas y, por ello, se les llama relaciones estadísticas y pueden describirse en términos probabilísticos. Este tipo de relaciones funcionales pueden expresarse como un modelo estadístico para una variable dependiente y_i y un conjunto de $k-1$ variables explicativas o regresores X_{ki} : En donde el término u_i es un error o perturbación aleatoria y $\beta_1 \dots \beta_k$ son los parámetros desconocidos a estimar por el modelo.” (Quintana & Mendoza, 2016, pág. 20).

Estado de conservación Fitto y Corvini: Es una clasificación de inmuebles de acuerdo con su estado de conservación que hace parte de un sistema de depreciación elaborado por Fitto y Corvini. De acuerdo con la Resolución 620 de 2008 del IGAC, los inmuebles pueden corresponder a una las siguientes clases:

- Clase 1: El inmueble está bien conservado y no necesita reparaciones ni en su estructura ni en sus acabados.
- Clase 2: El inmueble está bien conservado pero necesita reparaciones de poca importancia en sus acabados especialmente en lo que se refiere al enlucimiento.
- Clase 3: El inmueble necesita reparaciones sencillas por ejemplo en los pisos o pañetes.
- Clase 4: El inmueble necesita reparaciones importantes especialmente en su estructura.
- Clase 5: El inmueble amenaza ruina por tanto su depreciación es del 100% conservación basadas en las tablas de Fitto y Corvini.

Barrio: División de una población con características comunes

Estratificación socioeconómica (Estrato): Según el Dane, “es una clasificación en estratos de los inmuebles residenciales que deben recibir servicios públicos. Se realiza principalmente para cobrar de manera diferencial por estratos los servicios públicos domiciliarios permitiendo asignar subsidios y cobrar contribuciones en esta área.” (Dane, 2019).

Distancia a pie: Es la distancia recorrida haciendo uso del entramado urbano de vías públicas (andenes y vías vehiculares)

Distancia recta: Es la distancia en línea recta existente entre dos centroides, en este caso, entre el centroide de la vivienda y el punto de acceso principal al Centro Comercial Viva Envigado localizado sobre la Avenida las Vegas.

Zona de influencia: Es el área en la cual se espera haya un efecto o impacto generado por el desarrollo de una actividad.

Criterio de información de Akaike: es “un estimador insesgado asintótico de la información de Kullback-Leibler esperada, entre un modelo candidato ajustado y el verdadero modelo. El AIC fue propuesto por Akaike en 1974” (Amaya, 2018).

Método de comparación o de mercado. De acuerdo con la Resolución 620 de 2008 del IGAC, el método de comparación o de mercado “es la técnica valuatoria que busca establecer el valor comercial del bien, a partir del estudio de las ofertas o transacciones recientes, de bienes semejantes y comparables al del objeto de avalúo. Tales ofertas o

transacciones deberán ser clasificadas, analizadas e interpretadas para llegar a la estimación del valor comercial.”

11.2 De la Teoría económica Urbana

De acuerdo con el economista Oscar Borrero en su libro “Economía Urbana y Plusvalía del suelo”, la valoración es “una disciplina que pertenece a las ciencias económicas y tiene por objeto determinar los precios del mercado y valores físicos de los activos.” (Borrero Ochoa, 2018) y, por lo tanto, toda actividad desarrollada por un evaluador estará basada en las leyes y teorías de la economía urbana.

Tal como Borrero enuncia en su libro, la teoría económica urbana aparece a mediados del siglo XX con el fin de estudiar el fenómeno de los asentamientos urbanos y como consecuencia del traslado del valor del suelo y la plusvalía del campo a las ciudades, (Borrero Ochoa, 2018) y es por este motivo que en las ciudades aparece una competencia por el uso del suelo y por la localización de los diferentes sectores (industrial, comercio y servicios, y residencial).

Específicamente en el sector de comercio y servicios, Borrero enuncia que en las ciudades pluricéntricas, como es el caso de Bogotá, Medellín, México, Sao Paulo, Santiago, Caracas, se van conformando varios centros de diferentes grados, de acuerdo con su localización, con los cuales se pretende evitar lo que llama deseconomías externas (altos costos del suelo, aspectos ambientales, ruido, contaminación, tráfico, entre otras) (Borrero Ochoa, 2018), y en este sentido “la construcción de grandes o medianos Centros Comerciales ayuda a concentrar estas clases de subcentros en medio de las zonas residenciales o de servicios.” (Borrero Ochoa, 2018).

Borrero en su libro explica cómo la economía urbana se basa en cinco principios como son: Principio de aglomeración, de accesibilidad, de interacción espacial, de jerarquía y competitividad, los cuales se fueron presentados inicialmente por el profesor Roberto Carmingni en su libro de “Economía Urbana” (Borrero Ochoa, 2018).

Dentro de estos principios destacamos dos por su alta incidencia en el tema de la investigación:

1. El principio de aglomeración del cual Borrero expone que “se basa en las ventajas que genera la aglomeración de una población en una ciudad, porque con ello se concentra el mercado: la oferta y la demanda”, de gran importancia para los diferentes centros

comerciales o almacenes de gran superficie (Borrero Ochoa, 2018). Pero también destaca que, dependiendo del crecimiento, las aglomeraciones pueden generar lo que él llama “deseconomías”, dentro de las cuales contempla los problemas generados por el tráfico como la congestión, la contaminación y el ruido, aunado a la falta de espacio público y finalmente a una mala calidad de vida. (Borrero Ochoa, 2018).

2. Respecto al principio de accesibilidad Borrero explica que “...nace de la competencia entre las diversas actividades económicas para asegurarse las localizaciones más ventajosas” y por lo tanto es correspondiente con las ventajas que da la localización (Borrero Ochoa, 2018). En este sentido Borrero manifiesta que “las familias buscarán los mejores lugares para sus viviendas de acuerdo a las distancias al trabajo o a zonas comerciales y de servicios.”, citando además la frase de Hurd en 1903 cuando habla de la importancia de la accesibilidad que dice que “dado que el valor del suelo urbano depende de la renta, y la renta de la localización, y la localización de la conveniencia y la conveniencia de la cercanía, podemos eliminar los pasos intermedios y decir que el VALOR DEPENDE DE LA CERCANÍA”, y finaliza Borrero afirmando “que el valor del suelo depende de la cercanía al centro o al Centro de Negocios, o de la cercanía a las actividades que generan demanda (un centro comercial), un distrito turístico, un aeropuerto, un nodo de transporte, una estación de transporte masivo, etc.” (Borrero Ochoa, 2018).

11.3 El estudio de las variables

El Dr. Francisco Rey Carmona en su informe de Tesis doctoral al hablar de las metodologías avanzadas de valoración enuncia que “los bienes inmuebles pueden ser considerados como el prototipo de bienes que son transmitidos como un conjunto de atributos ligados a la unidad física que se transmite” (Rey Carmona, 2014) y explica que según la teoría de Lancaster de 1966, “cada bien tiene una cesta de características que le dan utilidad, asimismo cada bien posee más de una característica y muchas de ellas son compartidas por más de un bien” (Rey Carmona, 2014).

Del mismo modo Rey Carmona explica que acorde con la teoría de Sherwin Rosen del año 1974, “los productos heterogéneos están constituidos por un conjunto de atributos y que el precio marginal implícito de los mismos puede ser determinado mediante la estimación de un modelo que explica el precio de un producto a partir de los atributos del mismo” (Rey Carmona, 2014). Rey Carmona muestra como Rosen da un punto de vista

diferente respecto al funcionamiento del mercado para bienes heterogéneos expresando que “las decisiones de los agentes (consumidores y oferentes) se evalúan no en el espacio de las distintas variedades de bienes similares, sino en el espacio de las características que los componen” (Rey Carmona, 2014).

Siguiendo la misma línea, encontramos autores como Aznar & Guijarro quienes exponen en su libro “Nuevos Métodos de Valoración. Valoración Multicriterio” quienes expresan que el valor de un activo ya no solo depende de sus características cuantitativas, sino que hay una tendencia actual a incorporar nuevas variables cualitativas dentro de las bases de valor como lo son el paisaje, el vecindario y la calidad del aire, surgiendo así la necesidad de incorporarlas en los procesos valuatorios. (Aznar Bellver & Guijarro Martínez, 2012).

Autores como George Lever describen que: “el valor de un bien raíz no sólo está determinado por sus características estrictamente residenciales, comerciales o industriales, sino también por los complejos procesos de inversión, especulación y arbitraje que tienen lugar con el crecimiento y desarrollo urbano de las ciudades, con la congestión de las áreas urbanas inducida por el mayor poblamiento y el crecimiento de la edificación, y con las políticas de regulación urbana que implementa la autoridad, entre otros” (Lever D.).

De acuerdo con Castaño, Laverde, Morales y Yaruro, existen metodologías que parten del supuesto de que la vivienda es un bien diferenciado y que por lo tanto “su precio debe evaluarse en función de sus características” calculando a partir de “una estimación econométrica su contribución marginal de cada atributo en el precio, lo que se conoce como precio implícito o hedónico” (Castaño, Laverde, Morales, & Yaruro, 2013).

Dentro de estas metodologías que parten del bien como un conjunto de atributos, y a las que se refieren diferentes autores en sus investigaciones del valor de los activos, se encuentran los modelos hedónicos. Al respecto Castaño et al describen en su informe que “En términos generales, una ecuación hedónica permite explicar el precio de un bien en función de sus características, logrando medir la contribución marginal de cada una de estas en el precio. De esta manera, los precios hedónicos son definidos como los precios implícitos de los atributos de los bienes.” (Castaño, Laverde, Morales, & Yaruro, 2013).

De esta misma forma, Lever establece que: “la Teoría de Precios Hedónicos pretende explicar el valor de un bien raíz, entendido como un conjunto de atributos (superficie,

aptitud de uso del suelo, calidad de la construcción, diseño interior y exterior, áreas verdes, ubicación, características del vecindario, etc.), en función de cada uno de ellos, obteniendo sus respectivas valoraciones y, por ende, demandas implícitas” (Lever D.).

11.4 Teoría de la relación de la variable distancia y su incidencia en el valor de los inmuebles

De acuerdo con John W. Matthews dentro de la teoría económica urbana existen dos puntos de vista respecto a la relación de la distancia entre las zonas de comercio y el valor de las viviendas residenciales. Expresa en el informe de su investigación que desde el punto de vista de la “teoría microeconómica aplicada al terreno urbano” el valor de la tierra está “determinado por el costo del transporte”, lo que implica que las propiedades residenciales mejoran su valor a medida que está más cerca de los centros de comercio, debido a la reducción del costo de viaje (Matthews, 2007). Enuncia además que existe una segunda teoría en la que las propiedades residenciales pierden valor a medida que se encuentran más cerca del centro de comercio debido a las “externalidades”, como “ruido” y “congestión” que genera el comercio, que producen el efecto contrario (Matthews, 2007).

Matthews enuncia la necesidad de tomar como variables para el análisis la distancia de viaje y la distancia recta. En el primer caso la distancia de viaje permite medir la “conveniencia” por la localización del sitio de comercio (Matthews, 2007). En el segundo caso, la distancia recta permite medir los efectos negativos generados como “consecuencia de una actividad económica” y cuyo efecto se observa en línea recta como es por ejemplo el ruido, la contaminación o la congestión (Matthews, 2007).

En el documento presentado por Castaño, Laverde, Morales y Yaruro, para el estudio de los índices de precios de vivienda, enunciaron diferentes investigaciones realizadas para la ciudad de Bogotá en las que se tuvieron en cuenta la variable ubicación. En su informe manifiestan que “hay diferentes trabajos para Bogotá, que junto con herramientas de la economía espacial, evalúan el efecto que tienen sobre el precio de las viviendas sus atributos y variables asociadas a la ubicación. Por ejemplo, Revollo (2009), Carriazo et al. (2011) y Santana & Núñez (2011) encuentran que variables como el área, el estrato y la cercanía a establecimientos de recreación, zonas verdes y servicios públicos tienen un efecto positivo en el precio de los inmuebles, mientras que elementos como la contaminación e inseguridad contribuyen de manera negativa. Por su parte, Morales &

Arias (2011) comprueban que el número de cuartos, el número de baños, el estrato, la disponibilidad de garajes, zona verde y servicio telefónico tienen un efecto positivo en el valor de la vivienda.” (Castaño, Laverde, Morales, & Yaruro, 2013).

Tekel & Akbarishahabi estudiaron el efecto que tenía la cercanía a un campo verde (parque), en su caso el parque Ankara/Botanik, en el precio de las casas cercanas, con el fin de determinar el valor económico creado por ésta zona verde. En su investigación partieron de la teoría de que los parques aumentan el valor de las viviendas a su alrededor, y buscaron comprobarlo de manera racional a través del método de precios hedónicos (Tekel & Akbarishahabi, 2013).

Trojanek, Gluszak, & Tanas analizaron en su investigación el impacto de la proximidad a las áreas verdes urbanas sobre los precios de los apartamentos en la ciudad de Varsovia (Polonia), partiendo del concepto de que los valores del mercado de una propiedad dependen principalmente de sus características físicas y especialmente de la ubicación. Enuncian en su artículo que estas decisiones de ubicación de los individuos están influenciadas principalmente por factores como la distancia al trabajo, a las escuelas, a los hospitales, al sistema de transporte público y al entorno (ambiente) construido o natural, y por ello parten su investigación desde la teoría de que la cercanía a las áreas verdes urbanas incrementa el valor de las viviendas.

Encontraron en su investigación, por medio de modelos Hedónicos tales como Ordinary Least Square, Weighted Least square y Median quantil regression, que “en la vecindad cercana (menos de 100m de distancia)” al área verde, los precios de venta se incrementaron entre 8.0 – 8.6% dependiendo del modelo utilizado. (Trojanek, Gluszak, & Tanas, 2017)

Quintero García & Rivera Grajales analizaron en su investigación el efecto de la distancia al acceso del Megabus en el valor de la propiedad en la zona subcentro del barrio Cuba en la ciudad de Pereria (Colombia), con el fin de cuantificar el valor que le da el sistema de transporte a la vivienda. En sus estudios utilizaron la metodología de precio hedónicos. (Quintero García & Rivera Grajales, 2012).

Muñoz-Raskin realizaron un estudio con el fin de determinar si la accesibilidad a pie al sistema de transporte masivo, Transmilenio de la ciudad de Bogotá, afecta el valor de la propiedad raíz, encontrando que el mercado agrega valor en las proximidades de las líneas alimentadoras, sin embargo hallaron diferencias en los resultados cuando éstos los

analizaron por estrato socio-económico, lo cual sería un factor a tener en cuenta en la investigación (Munoz-Raskin, 2009).

En el mismo sentido, Rodríguez & Targa encontraron en su investigación a través de un modelo de precios hedónicos, que “por cada 5 minutos de tiempo de caminata adicional a una estación del sistema de transporte masivo, el precio de alquiler de una propiedad disminuía entre 6.8 y 9.3%” (Rodríguez & Targa, 2004).

Autores como Kwame Addae-Dapaah expresan que las teorías urbanas modernas, a diferencia de las tradicionales, tienen en cuenta el comportamiento del mercado en términos espaciales y la preferencia de ubicación del consumidor y enuncian que “el precio de la vivienda debe, entre otras cosas, reflejar la influencia combinada de las externalidades positivas y negativas de varios atributos de ubicación” considerando dentro de esas “externalidades” a los centros comerciales, los cuales “ejercen simultáneamente efectos atractivos y repulsivos que pueden afectar la elección de ubicación de una vivienda” (Addae-Dapaah & Lan, 2010)

Votsis en su investigación sobre el efecto espacial de parques, bosques y campos en el precio de los apartamentos en la ciudad de Helsinki, expresa la importancia de la ubicación relativa en la prosperidad socioeconómica de las ciudades. Enuncia en su artículo que “esta lógica espacial históricamente ha brindado beneficios fundamentales, como la prestación óptima de servicios y empleo, redes sociales y una distribución e intercambio eficientes de bienes” y que la lógica de optimización espacial ha buscado mantener “una cantidad suficiente de beneficios por la aglomeración, asignando al espacio funciones altamente productivas y competitivas” (Votsis, 2016).

12 MARCO METODOLÓGICO

Con este estudio se realizó un análisis cuantitativo y espacial que permitiera estudiar objetivamente la variable distancia, específicamente, con respecto al centro comercial VIVA Envigado.

Por lo tanto, la investigación desde un enfoque correlacional permite pasar del conocimiento empírico generado por el mercado a través de la experiencia y de la visión individual, a un conocimiento científico basado en hechos y datos, a través de una mirada sistémica, que permita teorizar al respecto.

El método de razonamiento utilizado es el deductivo que “consiste en tomar conclusiones generales para obtener explicaciones particulares. El método se inicia con el análisis de los postulados, teoremas, leyes, principios, etcétera, de aplicación universal y de comprobada validez, para aplicarlos a soluciones o hechos particulares.” (Bernal, 2010).

La población corresponde a las 5617 viviendas de los Barrios El Portal, Alcalá, San Marcos, Jardines y Villagrande, determinadas por el Departamento Nacional de Planeación en 2016.

Con el muestreo se recopiló la información relativa a las características constructivas, cualitativas y económicas de las ofertas de mercado. La muestra es representativa de la diversidad de oferta en mercado. El muestreo amplio se utiliza “cuando se trata de establecer la estructura de un ámbito urbano, por lo que deberá obtenerse información de los diversos tipos constructivos en oferta, así como de las diferentes localizaciones posibles atendiendo el nivel de centralidad y demás externalidades urbanísticas” (García, 2007).

A través de portales de internet especializados se realizó una búsqueda de las ofertas del mercado entre septiembre y octubre de 2019 y se realizó una depuración de la información. Posteriormente, se realizó una investigación del mercado con el fin de establecer las características particulares que se utilizarían en el análisis.

Cada uno de los datos de mercado son georeferenciados con el fin establecer las distancias a pie, distancia en carro y distancia media con respecto al Centro Comercial Viva Envigado.

13 DESARROLLO DE LA INVESTIGACIÓN

13.1 Recolección de la información del mercado

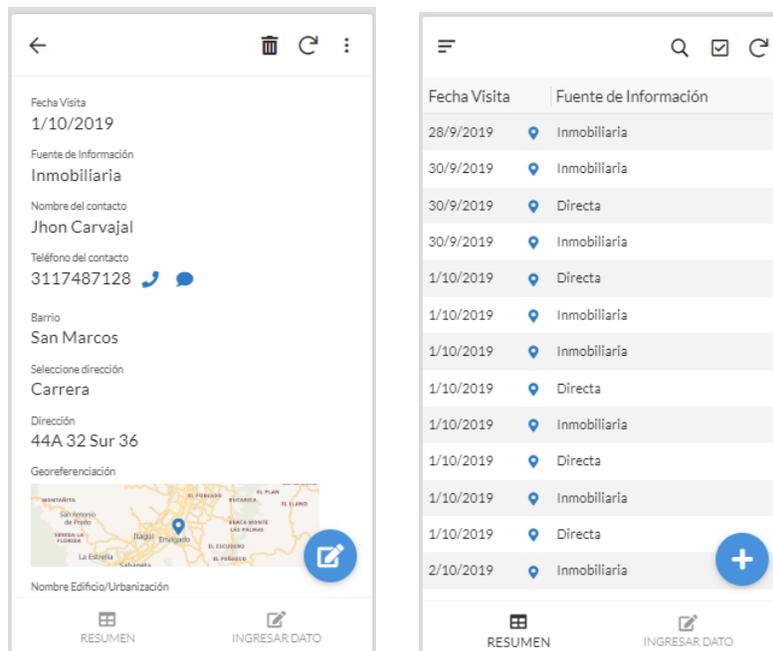
13.1.1 Contacto con la población.

El primer contacto con la población se realizó en el momento de realizar la delimitación de la zona de influencia. Luego se realizó un recorrido por el sector y se consultó sobre los estratos existentes en la zona como se detalló anteriormente.

13.1.2 Diseño de instrumentos para la recolección de información.

Para la toma de información se creó una aplicación para usar en dispositivos móviles haciendo uso de la plataforma para el desarrollo de aplicaciones de código abierto AppSheet, que permitiera tomar en el sitio tanto la georreferenciación de los inmuebles como las características de cada uno de ellos, incluyendo las variables que serían utilizadas para la investigación, facilitando la digitalización de la información, evitando con ello reprocesos y minimizando así la probabilidad de errores en la toma de datos.

Gráfico 6. Aplicación para la recolección de información



A continuación, se detalla la información que fue recolectada para cada inmueble en oferta:

Tabla 5. Datos recolectados para cada muestra

Dato	Descripción
Fecha Visita	Fecha en la que se visita la vivienda
Fuente de Información	Si el propietario entrega la información es "Directa"; si algún agente inmobiliario la entrega es " Inmobiliaria"
Nombre del contacto	Nombre de la fuente
Teléfono del contacto	Teléfono de la fuente
Barrio	Barrio en el que se localiza la vivienda. Puede ser "Alcalá", "El Portal", "Jardines" o "San Marcos"
Seleccione dirección	Vía sobre la que se ubica la vivienda según nomenclatura. Puede tomar los valores de "Calle", "Carrera", "Transversal" "Diagonal"
Dirección	Número de la vivienda según nomenclatura
Georeferenciación	Latitud, Longitud
Nombre Edificio/Urbanización	Nombre de la propiedad horizontal (si lo tiene)

Estrato	Estrato socioeconómico establecido por la Alcaldía de Envigado. Puede tomar valores enteros de 1 a 6.
Valor del negocio	Precio del inmueble si el tipo de negocio es "venta"
Área Construida	Área techada de la vivienda incluyendo elementos como columnas, muros, ductos y falseados, que no hacen parte del espacio útil de la vivienda
Área Lote	Área de lote de la vivienda, si no está sometida a Reglamento de propiedad horizontal
Valor Administración	Valor mensual de administración
Valor Impuesto Predial	Valor anual de administración
Año de construcción	Año en que fue construida la estructura del inmueble
Edad	Tiempo transcurrido entre el año de construcción del inmueble y la fecha de la visita. Puede tomar uno de los siguientes valores " Nuevo", "Inferior a un año", "entre 1 y 5 años", "entre 5.1 y 10 años", "entre 10 y 20 años", "mayor a 20 años".
Tipo Inmueble	Puede ser "apartamento", "casa", "duplex" . Hasta trifamiliar se considera casa.
Nivel del piso	Piso en el que se encuentra el apartamento

Ascensor	Puede ser "si" si lo tiene o "no" si no cuenta con ascensor
Cantidad total de pisos	Número total de pisos de la edificación
Tipo de Vivienda	Puede ser "unifamiliar", "bifamiliar", "trifamiliar", "multifamiliar"
Régimen	Se refiere a si la propiedad está sometido a régimen de propiedad horizontal o no. Las opciones con "Con RPH" o "Sin RPH"
Unidad Cerrada / Abierta	Indicar "cerrada" si se cuenta con un acceso o varios accesos comunes desde la calle para el edificio o unidad, de lo contrario indique "abierta"
Número de habitaciones	Número de alcobas, cuartos o habitaciones
Número de baños completos	Número de baños con batería sanitaria, lavamanos y ducha
Número de baños medios	Número de baños sin ducha
Número de Closets	Número de espacios destinados a almacenamiento, dentro de la vivienda, son contar el espacio de almacenamiento de cocina
Balcón	Indicar "si", si tiene balcón o "no" si no lo tiene "NA" si no aplica
Estudio	Indicar "si" si tiene estudio o "no" si no lo tiene o "NA" si no aplica

Sala Independiente	Indicar "si" si tiene sala independiente o "no" si no es independiente o "NA" si no aplica
Comedor independiente	Indicar "si" si tiene comedor independiente o "no" si no es independiente o "NA" si no aplica
Sala/comedor	Indicar "si" si la sala y comedor están integradas o "no" si no lo están o "NA" si no aplica
Zona de ropas	Indicar "si" si tiene zona de ropas o "no" si no la tiene.
Patio	Indicar "si" si tiene patio o "no" si no lo tiene.
Gas natural	Indicar "si" si cuenta con servicio de gas natural, "no" si no lo tiene
Poniente	Indicar "si" si la vivienda recibe directamente el sol de la tarde, o "no" en caso contrario
Vista	Indicar si la vista frontal de la vivienda es hacia una "zona verde", "Parque", "vista a la calle", "vista interna", "vista externa", "oriente", "occidente", "norte", "sur" u "otro edificio. Permite seleccionar varias opciones.
Cantidad de Cuartos útiles	Indicar el número de cuartos útiles o espacios destinados a almacenamiento externos a la vivienda

Parqueaderos	Seleccionar entre "comunes", "descubierto", "garaje cerrado", "doble lineal", "doble paralelos", "sencillo", "No tiene", "Moto", "Visitantes". Permite seleccionar varias opciones
Estructura	Seleccionar entre "aportificada concreto", mampostería simple", "mampostería estructural", "metálica", "mixta", "muros vaciados", "prefabricada"
Acabados Pisos	Seleccionar entre "cerámica", "madera", "madera laminada", "mortero", "porcelanato", "vinilo", "mármol"
Tipo de Cocina	Seleccionar entre "integral", "semi-integral" o "sencilla"
Acabados Paredes	Seleccionar entre " ladrillo", "revoque", "estuco", "pintura" u "otros acabados"
Estado de conservación Fitto y Corvini	Seleccionar de acuerdo con la tabla de Fitto y Corvini en una escala que va entre 1 para los inmuebles que "no ha sufrido ni necesita reparaciones" y 5 para los que "requieren demolición"
Carros	Número de celdas de parqueo privadas
Portería	Indicar si la unidad residencial en la que se encuentra el inmueble cuenta con " portería 24 horas", "portería sólo diurna", "portería sólo nocturna", "sin portería"

Zonas comunes

Indicar si la unidad residencial en la que se encuentra el inmueble cuenta con "zona común completa", "zona común sencilla" o "sin zona común"

13.1.3 Recolección de información.

El proceso de captura de información se inició con la búsqueda en portales como Fincaraíz, Metrocuadrado y Espacio Urbano, entre otros, de las ofertas de inmuebles residenciales en la zona de estudio.

Con esta información se procedió a realizar llamadas para confirmar los datos y programar visitas a los inmuebles para conocerlos internamente, poderlos describir correctamente, tomar las fotos y tomar las coordenadas a través de la aplicación para móviles ya descrita.

Teniendo en cuenta que muchos de los inmuebles encontrados en los portales ya no se encontraban disponibles y que consideramos importante obtener más datos, se procedió a recorrer el sector establecido para mirar avisos físicos de venta. En estos recorridos se obtuvieron las coordenadas, y se procedió a realizar las llamadas telefónicas para solicitar la información requerida y las fotos. En los casos en que fue necesario también se visitaron estos inmuebles para conocerlos internamente.

En unos casos la toma de la información fue de forma directa con el propietario del inmueble, caso en el que pudimos obtener la mayor cantidad de información, y en otros casos a través de inmobiliarias o de corredores inmobiliarios independientes a quienes los propietarios les habían consignado sus inmuebles.

Tabla 6. Caracterización de los datos recolectados. Fuentes de información

Fuente Información	Cantidad Muestras	Porcentaje de la muestra
Directa	20	33,90%
Inmobiliaria	39	66,10%
Total general	59	100,00%

El proceso de recolección de información se realizó entre el 28 de septiembre y el 15 de octubre de 2019 en el cual se obtuvieron 58 ofertas de inmuebles disponibles para la venta.

13.2 Modelos econométricos

De acuerdo con Quintana Romero y otros, los modelos econométricos están compuestos por relaciones estadísticas de las variables, descritas en términos probabilísticos, con las cuales se busca simplificar la realidad a través de dichas relaciones existentes entre las variables, en cual hay una variable dependiente y_i , y un conjunto de $k-1$ variables explicativas o también llamados regresores X_{ki} como se muestra en la ecuación a continuación:

$$y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i$$

Al término u_i se le denomina error o perturbación aleatoria, y donde los coeficientes β_1 a β_k son los parámetros a estimar en el modelo. (Quintana Romero, y otros, 2016).

Así mismo, dichos autores establecen que con el fin de evitar problemas de especificación del modelo, al modelo econométrico estimado es necesario realizarle una evaluación tanto económica como econométrica, en la que no solo se verifique la evidencia a favor o en contra que produce respecto a la teoría económica, sino también los supuestos que se utilizaron tanto en la parte determinista como en el término estocástico. (Quintana Romero, y otros, 2016).

En este caso se verifica en cada modelo las características de no normalidad, la autocorrelación, la heteroscedasticidad, la multicolinealidad, y por lo tanto se debe verificar el término u_i de la ecuación para comprobar que su media es cero, su homoscedasticidad, que no presente autocorrelación y que se distribuya de manera normal, para que el modelo tenga validez, (Quintana Romero, y otros, 2016).

13.2.1 Selección del modelo lineal.

El procedimiento general utilizado para definir el modelo lineal final con el cual se verificará la incidencia de la distancia al centro comercial Viva Envigado en los precios de las viviendas se describe a continuación, teniendo en cuenta que el mejor modelo es el que explique el fenómeno estudiado con la menor cantidad de variables.

Para la determinación del mejor modelo se utilizó el modelo step – step en el programa R, con la metodología de atrás/adelante y utilizando el criterio de Akaike. Una vez

obtenidos los coeficientes del modelo se aplicó un modelo lineal a los datos correspondientes a las variables preseleccionadas mediante el modelo paso-paso haciendo uso del programa estadístico de código abierto R, con el fin de identificar los resultados de p-crítico de cada variable, el coeficiente de determinación r^2 del modelo, y más específicamente el r^2 ajustado el cual penaliza el resultado del coeficiente de determinación r^2 por la cantidad de variables utilizadas. Así mismo se obtuvo el resultado del p-crítico del modelo, del estadístico F del modelo y el residuo generado por el mismo.

Se verificó en cada modelación realizada, el resultado del coeficiente de determinación r^2 , el estadístico F mayor a 10, y un p-crítico del modelo inferior al error máximo acorde con el nivel de confianza utilizado de 95%.

Posteriormente, una vez establecido el mejor modelo, es decir aquel cuyas variables explicativas tengan un p-crítico menor al error, un p-crítico del modelo inferior al nivel de significancia, un alto coeficiente de determinación r^2 ajustado y un estadístico F superior a 10, se le adicionó la variable distancia (distancia a pie, distancia lineal, distancia en carro) y se verificaron estos mismos criterios.

A cada modelo se le evaluó el cumplimiento de los supuestos del modelo como son: la colinealidad, la igualdad en las varianzas u homocedasticidad, a lo largo de los predictores, la normalidad de los residuos, pruebas de intercorrelación de las variables con el fin de identificar si estas están relacionadas entre sí, en que magnitud y si esto puede afectar la validez del modelo. Para ello se aplicaron pruebas como:

- Gráficos de diagnóstico en R.
- Prueba Durbin-Watson, la cual es una estadística de prueba utilizada para detectar la presencia de autocorrelación entre las variables. El estadístico de la prueba debe estar en el rango de 0 a 4. Mientras más cercano a 2 es el resultado del estadístico es un indicativo que no hay autocorrelación. Los valores sustancialmente inferiores a 2 indican que hay correlación serial positiva, y de la misma manera, valores sustancialmente superiores a 2 indican que hay correlación negativa. Valores del estadístico Durbin – Watson inferiores a 1 indican que hay alta probabilidad de autocorrelación y es una señal de alarma para el modelo evaluado.
- A través de la prueba de factor de inflación de varianzas se verifica la multicolinealidad del modelo, con el fin de identificar que las variables explicativas del

modelo son independientes y muestran diferentes aspectos en el modelo. El resultado debe ser inferior a 5 (resultados mayores a 5 indican que las variables pueden estar correlacionadas).

13.2.2 Comprobación de supuestos del modelo.

Con el fin de comprobar los supuestos del modelo de regresión lineal se utilizaron métodos gráficos y a su vez tests de hipótesis desarrollados por diferentes investigadores como se muestra a continuación, teniendo en cuenta que para cualquiera de los test de hipótesis realizados para verificar los supuestos del modelo de regresión línea se tiene que:

- Si $p\text{-value} < \text{Nivel de significancia (alfa)}$; entonces la hipótesis nula se rechaza y se debe especificar nuevamente el modelo.
- Si $p\text{-value} > \text{Nivel de significancia (alfa)}$; no se puede rechazar la hipótesis nula. Siendo el nivel de significancia alfa en nuestro caso de 0.05.

Gráficas de diagnóstico

En las gráficas de diagnóstico arrojadas por el software R, se puede observar visualmente el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, en las cuales, acorde con lo descrito por Romero Mendoza y otros (Quintana Romero, y otros, 2016) se comparan los residuos del modelo respecto a los estimados de la regresión en el primer caso, y en las otras, se analizan los residuales estandarizados de la regresión

Comprobación de autocorrelación serial

De acuerdo con la cita realizada por Quintana Romero y otros, de Gujarati, 2004 Griffiths y Judge, 1993, “La autocorrelación es un caso particular del modelo de regresión generalizado que se produce cuando las perturbaciones del modelo presentan correlaciones entre ellas. La autocorrelación supone que la matriz de varianzas y covarianzas de las perturbaciones presentan valores distintos de cero en los elementos que están fuera de la diagonal principal” (Quintana Romero, y otros, 2016).

En este caso se utilizó el estadístico de Durbin-Watson con el fin de detectar la presencia de autocorrelación, cuya interpretación se explicó anteriormente.

Comprobación de multicolinealidad

De acuerdo con Quintana Romero y otros, la multicolinealidad es un problema de grado en el establecimiento de los modelos económicos y describen que “el comportamiento de

buen parte de las variables económicas guarda algún tipo de relación unas con otras y esa relación puede ser de menor o mayor grado. Solamente cuando dicha relación es de mayor grado podría ser un problema dentro de la modelación econométrica” (Quintana Romero, y otros, 2016).

En este caso utilizamos la prueba de factor de inflación de varianza se verifica la multicolinealidad del modelo, con el fin de identificar la independencia de las variables explicativas, ya que acorde con Quintana Romero y otros, “El factor de inflación varianza (VIF) se utiliza como una medida del grado en que la varianza del estimador de mínimos cuadrados es incrementada por la colinealidad entre las variables.” (Quintana Romero, y otros, 2016). De acuerdo con estos autores, los resultados del test realizado deben ser inferiores a 5, ya que valores mayores a 5 e incluso a 5, indican que las variables pueden tener una fuerte colinealidad (Quintana Romero, y otros, 2016).

Comprobación de Heterocedasticidad

Según Quintana Romero y otros respecto a la heteroscedasticidad establecen que “El origen de la heterocedasticidad, está asociado a la varianza creciente de las perturbaciones aleatorias de los valores de algunas de las variables, incluidas en el modelo.” Lo cual puede traer como consecuencia la ineficiencia del método de mínimos cuadrados ordinarios en la determinación de los estimadores, y error en el cálculo de la varianza y por ende en el cálculo de la t de student y otros cuyo cálculo esté basado en la varianza. Por lo tanto Quintana Romero y otros, describen que “En términos econométricos los errores de la estimación, no deben crecer a medida que lo hace el tamaño de la muestra de (x) , la dispersión de los errores en la estimación, debe mantenerse estable y no debe dispersarse en el tiempo.”

El test de Breusch-Pagan es usado en este caso para determinar si hay heterocedasticidad en las muestras, a través del análisis de varianza de los residuos. Al aplicar el test, si p-value obtenido es inferior a 0.05, se rechaza la hipótesis nula y por lo tanto se asume la heterocedasticidad.

Comprobación de la No linealidad

Tal como lo describe en su libro Quintana Romero y Otros, la “prueba de especificación de Ramsey, denominada RESET, por sus siglas en inglés Regression Equation Specification Error Test, debida a Ramsey (1969), que sirve para detectar errores de especificación

ocasionados por la omisión de variables independientes, por la posible existencia de correlación entre las variables en X y ε o bien, porque la forma funcional de las variables independientes no es la apropiada” (Quintana Romero & Mendoza, 2016).

13.2.3 Procesamiento de la información y preselección de las variables explicativas.

Inicialmente, se tomaron datos de inmuebles unifamiliares, bifamiliares, trifamiliares y multifamiliares. Sin embargo, los inmuebles unifamiliares y bifamiliares tienen características que los hacen especiales, asociadas al valor del lote, y el número de muestras obtenidas de esta tipología no fue representativa dentro del muestreo obtenido.

Tabla 7. Caracterización de los datos recolectados. Tipologías de vivienda

Tipo Vivienda	Cantidad de datos	Porcentaje (%)
Bifamiliar	3	5,08%
Multifamiliar	53	89,83%
Trifamiliar	1	1,69%
Unifamiliar	2	3,39%
Total general	59	100,00%

Por lo anterior, se incluyeron en las modelaciones del presente estudio únicamente los inmuebles multifamiliares, para un total muestral de 53 ofertas.

De toda la información capturada con la Aplicación en campo se descartaron para el análisis las características que no estuvieran completas en todas las muestras, porque los propietarios actuales o la inmobiliaria o asesores inmobiliarios no tenían información certera o verificable. Ejemplo de esto fueron algunos datos de valores de administración de los inmuebles y edades de los inmuebles al no tener certeza con el valor otorgado de los mismos.

En la tabla a continuación se relacionan las posibles variables explicativas consideradas para ejecutar los diferentes análisis de modelación, clasificándolas en cuantitativas y cualitativas según el caso.

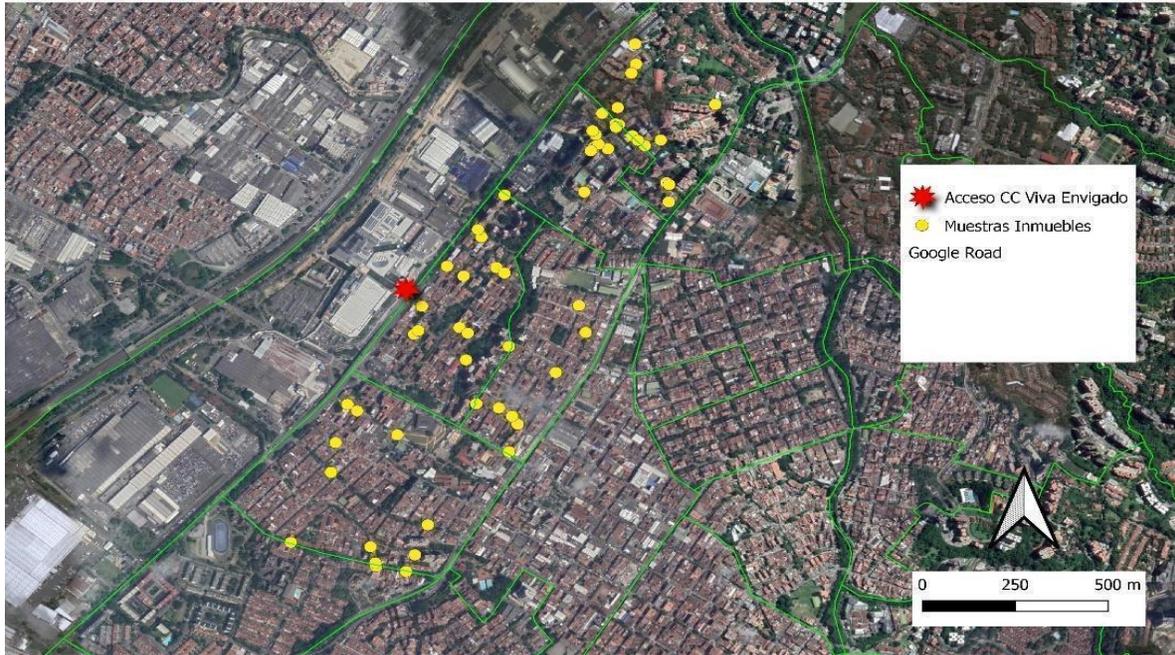
Tabla 8. Posibles variables explicativas

Variable	Descripción y/o tratamiento	Clasificación (cuantitativa-cualitativa)	Unidad
Valor del negocio	Valor de oferta del inmueble	Cuantitativa	NA
Estrato	Estrato socioeconómico establecido por la Alcaldía de Envigado. Puede tomar valores enteros de 1 a 6.	Cuantitativa	NA
Área Construida	Área techada de la vivienda incluyendo elementos como columnas, muros, ductos y falseados, que no hacen parte del espacio útil de la vivienda	Cuantitativa	m2
Nivel del piso	Piso en el que se encuentra el apartamento	Cuantitativa	NA
Ascensor	Puede ser "1" si tiene ascensor o "0" si no lo tiene	Cuantitativa	NA
Número de habitaciones	Cantidad de alcobas del inmueble	Cuantitativa	unidades
Cantidad de Cuartos útiles	Indicar el número de cuartos útiles	Cuantitativa	unidades
No Parquaderos (carros)	Indicar el número de celdas de parqueo	Cuantitativa	unidades
Estado de conservación	Seleccionar de acuerdo con la tabla de Fitto y Corvini	Cualitativa	NA
Zonas comunes	Puede ser "2" si tiene zonas comunes completas, "1" si tiene zonas comunes sencillas, "0" si no tiene zonas comunes	Cuantitativa	NA

Portería	Puede ser "2" si tiene portería 24 horas, "1" si tiene portería medio tiempo, "0" si no tiene portería	Cuantitativa	NA
Poniente	Indicar "1" si la vivienda recibe directamente el sol de la tarde, o "0" en caso contrario sobre la Avenida Las Vegas	Cuantitativa	NA
Distancia a pie	Distancia a pie desde la portería de la edificación hasta el acceso principal de Viva Envigado sobre la Avenida Las Vegas	Cuantitativa	m
Distancia lineal	Distancia lineal desde la portería de la edificación hasta el acceso principal de Viva Envigado	Cuantitativa	m
Distancia en carro	Distancia lineal desde la portería de la edificación hasta el acceso principal de Viva Envigado sobre la Avenida Las Vegas	Cuantitativa	m

Gráfico 7. Localización muestras

LOCALIZACIÓN MUESTRAS INMUEBLES



A partir de la herramienta Quantum GIS (QGIS) y usando los complementos: OSM downloader, ORS Tools y QuickMapServices, con los cuales se trazaron las rutas para determinar la distancia a pie y la distancia en carro, haciendo uso de las funciones de openrouteservice.org, las cuales, acorde con el complemento, están basadas en OpenStreetMap. El cálculo de la distancia lineal o euclidiana se obtuvo de igual forma con el uso de QGIS con las herramientas de análisis vectorial de QGIS.

Gráfico 8. Rutas distancia a pie a Viva Envigado

RUTAS DISTANCIA A PIE



Gráfico 9. Rutas distancia a pie a Viva Envigado

RUTAS PARA DISTANCIA EN CARRO



13.3 Análisis de los datos de las ofertas recolectados en campo

13.3.1 Caracterización de las muestras.

Para el total de la información recolectada en campo para los inmuebles multifamiliares con los que se realiza el estudio (53 Inmuebles de la muestra inicial) en las tablas a continuación presentamos, los resultados informativos por característica tomada en campo:

Tabla 9. Caracterización de los datos recolectados. Fuentes

Etiquetas de fila	Cantidad Datos	Cuenta de Fuente
Directa	18	33,96%
Inmobiliaria	35	66,04%
Total general	53	100,00%

Respecto al estrato, la distribución de las muestras se enseña en la tabla a continuación:

Tabla 10. Caracterización de los datos recolectados. Estratos

Estrato	Cantidad Datos	% respecto al total
---------	----------------	---------------------

Estrato 3	10	12,77%
Estrato 4	10	17,02%
Estrato 5	33	70,21%
Total general	53	100,00%

Respecto a la vetustez, dada la imposibilidad de obtener el 100% de los resultados confiables, se elaboraron rangos con edades aparentes para poder analizar las muestras, encontrando que:

Tabla 11. Caracterización de los datos recolectados. Vetustez

Vetustez	Cantidad Datos	% respecto al total
Entre 1 y 5 años	6	11,32%
Entre 10 y 20 años	28	52,83%
Entre 5.1 y 10 años	6	11,32%
Inferior a 1 año	2	3,77%
Mayor a 20 años	8	15,09%
Nuevo	3	5,66%
Total general	53	100,00%

Tabla 12. Caracterización de los datos recolectados. Tipo de inmueble

Tipo Inmueble	Cantidad Datos	% respecto al total
Apartamento	48	90,57%

Apartamento Dúplex	5	9,43%
Total general	53	100,00%

Tabla 13. Caracterización de los datos recolectados. Nivel del piso

Nivel del Piso	Cantidad Datos	% respecto al total
1	4	1,39%
2	6	4,17%
3	3	3,13%
4	14	19,44%
5	7	12,15%
6	4	8,33%
7	4	9,72%
8	2	5,56%
9	3	9,38%
11	2	7,64%
12	1	4,17%

13	2	9,03%
17	1	5,90%
Total general	53	100,00%

Tabla 14. Caracterización de los datos recolectados. Ascensor.

Ascensor	Cantidad Datos	% respecto al total
No	17	32,08%
Si	36	67,92%
Total general	53	100,00%

Tabla 15. Caracterización de los datos recolectados. Tipo de unidad (abierta/cerrada)

Tipo de unidad	Cantidad Datos	% respecto al total
Abierta	31	58,49%
Cerrada	22	41,51%
Total general	53	100,00%

Tabla 16. Caracterización de los datos recolectados. Cantidad de habitaciones

Cantidad Habitaciones	Cantidad Datos	% respecto al total
2	6	11,32%
3	31	58,49%
4	15	28,30%
6	1	1,89%

Total general	53	100,00%
----------------------	-----------	----------------

Tabla 17. Caracterización de los datos recolectados. Número de baños completos

Baños Completos	Cantidad Datos	% respecto al total
2	31	58,49%
3	17	32,08%
4	4	7,55%
5	1	1,89%
Total general	53	100,00%

Tabla 18. Caracterización de los datos recolectados. Número de baños medios

Baños Medios	Cantidad Datos	% respecto al total
0	46	86,79%
1	7	13,21%
Total general	53	100,00%

Tabla 19. Caracterización de los datos recolectados. Balcón

Balcón	Cantidad Datos	% respecto al total
NO	4	7,55%
SI	49	92,45%
Total general	53	100,00%

Tabla 20. Caracterización de los datos recolectados. Estudio

Estudio	Cantidad Datos	% respecto al total
NO	29	54,72%
SI	24	45,28%
Total general	53	100,00%

Tabla 21. Caracterización de los datos recolectados. Poniente.

Poniente	Cantidad Datos	% respecto al total
NO	41	77,36%
SI	12	22,64%
Total general	53	100,00%

Tabla 22. Caracterización de los datos recolectados. Vista

Vista	Cantidad Datos	% respecto al total
Norte	1	1,89%
Occidente	1	1,89%
Oriente	1	1,89%
Parque , Vista a la calle , Vista Externa , Oriente	1	1,89%
Vista a la calle , Occidente	4	7,55%
Vista a la calle , Occidente , Norte	1	1,89%
Vista a la calle , Occidente , Otro edificio	3	5,66%
Vista a la calle , Oriente	4	7,55%
Vista a la calle , Oriente , Otro edificio	1	1,89%
Vista a la calle , Otro edificio	1	1,89%
Vista a la calle , Sur	4	7,55%

Vista a la calle , Sur , Otro edificio	5	9,43%
Vista a la calle , Vista Externa , Norte	12	22,64%
Vista a la calle , Vista Externa , Occidente	2	3,77%
Vista a la calle , Vista Externa , Sur	4	7,55%
Vista Externa	1	1,89%
Vista Interna , Norte	1	1,89%
Vista Interna , Oriente	1	1,89%
Vista Interna , Sur	2	3,77%
Vista Interna , Vista Externa , Occidente , Norte , Sur	1	1,89%
Zona verde , Vista a la calle , Vista Externa , Oriente	1	1,89%
Zona verde , Vista Interna , Sur	1	1,89%
Total general	53	100,00%

Si bien se tomó la información en campo para la variable cualitativa vista, la dificultad para categorizar la información y las múltiples opciones que se derivan de la misma, no se utilizó esta variable para el análisis.

Tabla 23. Caracterización de los datos recolectados. Cuartos útiles

Cuartos útiles	Cantidad Datos	% respecto al total
0	13	24,53%
1	38	71,70%
2	2	3,77%
Total general	53	100,00%

Tabla 24. Caracterización de los datos recolectados. Tipo de parqueadero

Tipo Parqueadero	Cantidad Datos	% respecto al total
-------------------------	-----------------------	----------------------------

Doble Lineal	12	22,64%
Doble Paralelos	3	5,66%
No tiene	6	11,32%
Sencillo	32	60,38%
Total general	53	100,00%

Pero más importante que el tipo de parqueadero o parqueaderos que tenía el inmueble de la muestra, era la cantidad de vehículos que podían parquear, ya que por ejemplo un inmueble podía tener al mismo tiempo un parqueadero lineal doble y uno sencillo, para una cantidad de 3 vehículos para el inmueble:

Tabla 25. Caracterización de los datos recolectados. Cantidad de celdas de parqueo privadas para vehículos

Cantidad de vehículos	Cantidad Datos	% respecto al total
0	6	11,32%
1	30	56,60%
2	17	32,08%
Total general	53	100,00%

Tabla 26. Caracterización de los datos recolectados. Acabado de pisos.

Acabado pisos	Cantidad Datos	% respecto al total
Cerámica	29	54,72%
Mármol	4	7,55%

Porcelanato	20	37,74%
Total general	53	100,00%

Tabla 27. Caracterización de los datos recolectados. Tipo de cocina

Cocina	Cantidad Datos	% respecto al total
Integral	52	98,11%
Sencilla	1	1,89%
Total general	53	100,00%

En este caso se observa que no es significativo evaluar el efecto de la variable cocina en el valor, cuando el 98.11% de los inmuebles cuentan con cocinas de características similares.

Tabla 28. Caracterización de los datos recolectados. Estado de conservación de Fitto y Corvini

Fitto y Corvini	Cantidad de datos	% respecto al total
1	12	22,64%
1,5	20	37,74%
2	16	30,19%
2,5	5	9,43%
Total general	53	100,00%

Tabla 29. Caracterización de los datos recolectados. Portería

Portería	Cantidad de datos	% respecto al total
Portería 24 horas	30	56,60%
Portería solo diurna	1	1,89%

Sin Portería	22	41,51%
Total general	53	100,00%

Tabla 30. Caracterización de los datos recolectados. Zonas comunes

Zona común	Cantidad de datos	% respecto al total
Sin zona común	28	52,83%
Zona común completa	19	35,85%
Zona común sencilla	6	11,32%
Total general	53	100,00%

13.3.2 Estadística descriptiva variables preseleccionadas.

Tomando los 53 datos correspondientes a los inmuebles multifamiliares en propiedad horizontal, se obtuvieron los siguientes datos estadísticos:

Tabla 31. Estadística descriptiva variables preseleccionadas

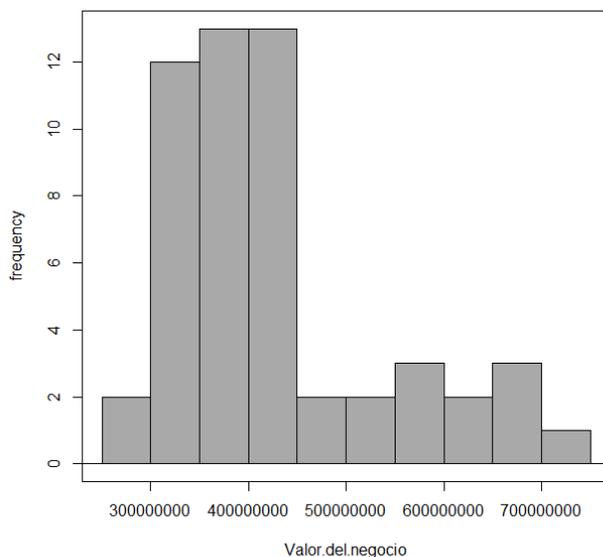
Variable	Promedio	Máximo	Mínimo	Desv. Est.	Coef. Variación	Coef. Asimetría	curtosis
Valor.del.negocio	429862000	714686000	250000000	112543932	0,26	1,09	0,4260
Aconstruida	111,00	200,0	62,0	29,35	0,26	0,91	0,8245
Ascensor	0,68	1,0	0,0	0,47	0,69	-0,79	-1,4304
Carros	1,21	2,0	0,0	0,63	0,52	-0,19	-0,5272
CuartoUtil	0,79	2,0	0,0	0,49	0,62	-0,43	0,2291
Distancia.a.Pie..m	809,96	1413,0	286,0	279,55	0,35	-0,35	-0,7483
Distancia.Car..m.	1340,36	2410,0	661,0	449,45	0,34	0,78	0,0039
Distancia.Lineal.m	581,14	998,4	66,7	258,82	0,45	-0,44	-1,1351

Estrato	4,43	5,0	3,0	0,80	0,18	-0,96	-0,7175
FittoCorvini	1,63	2,5	1,0	0,46	0,28	0,20	-0,7776
Habitaciones	3,23	6,0	2,0	0,72	0,22	0,89	2,9506
NivelPiso	5,43	17,0	1,0	3,43	0,63	1,26	1,6906
Poniente	0,23	1,0	0,0	0,42	1,87	1,35	-0,1977
Porteria	1,15	2,0	0,0	0,99	0,86	-0,31	-1,9551
ZonaComun	0,83	2,0	0,0	0,94	1,13	0,35	-1,8013

13.3.3 Verificación de la distribución normal de las muestras.

En primer lugar, construimos con el programa estadístico R el histograma de frecuencias para la variable dependiente, valor del negocio, cuyo resultado se presenta a continuación:

Gráfico 10. Histograma de frecuencias Valor del negocio

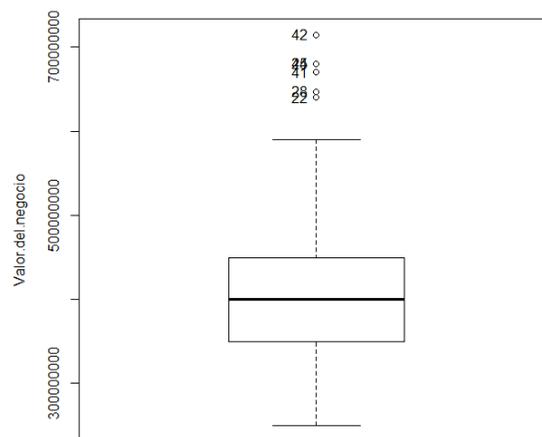


Para evaluar si los datos siguen una distribución normal se utilizó inicialmente la prueba modificada Lilliefors (Kolmogorov-Smirnov) para la variables dependiente, valor del negocio, dado que se cuentan con más de 50 datos. El resultado obtenido del estadístico $D = 0.21649$, y un $p\text{-value} = 0.000001447$, el cual es inferior al nivel de significancia del

0.05, lo que quiere decir en este caso que la variable valor del negocio no sigue una distribución normal, corroborando lo que se observa gráficamente en el histograma de frecuencia.

Procedemos entonces a verificar la presencia de datos atípicos con la aplicación de diagramas de caja como se observa a continuación:

Gráfico 11. Diagrama de caja Valor del negocio



Con el diagrama de la caja observamos que hay unas muestras que estadísticamente corresponden a outliers o datos atípicos del mercado obtenidos en el muestreo (muestras 28, 30, 31, 33, 36, 51, 53, 54 y 56), y por lo tanto luego de su identificación se procedió a analizarlos y eliminarlos ya que pueden incidir en el resultado de la investigación.

El histograma de frecuencias con las exclusiones realizadas luego del análisis del mercado, se presenta a continuación:

Gráfico 12. Histograma de frecuencias Valor del negocio excluyendo outliers

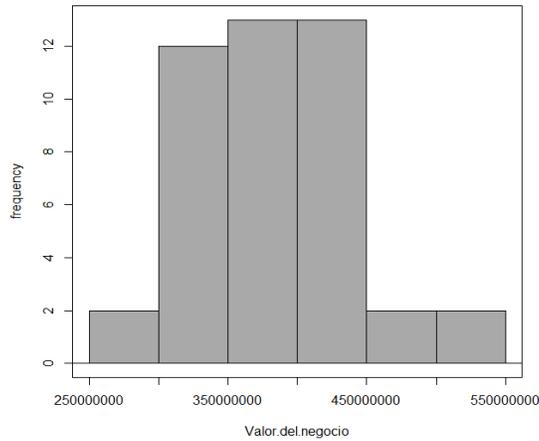
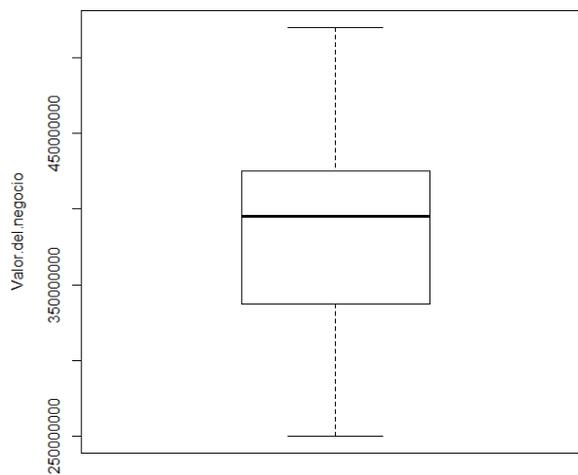


Gráfico 13. Diagrama de caja Valor del negocio excluyendo outliers



Una vez eliminados los datos atípicos se procedió a realizar el test de Shapiro-Wilk para evaluar la normalidad de los datos de la variable valor del negocio, obteniendo los siguiente

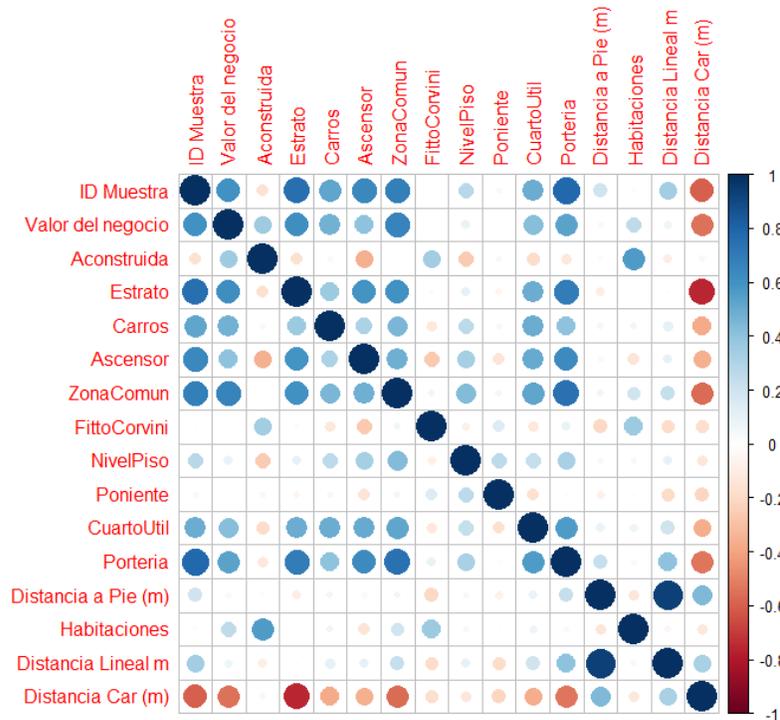
resultados : estadístico $W = 0.97766$, $p\text{-value} = 0.5428$. Por lo tanto, como $p\text{-value} (0.5428)$ es mayor que 0.05 , se acepta la hipótesis nula H_0 , lo que significa que se acepta que la variable valor del negocio sigue una distribución normal.

13.4 Modelaciones excluyendo datos atípicos

En estas modelaciones no se tuvieron en cuenta las muestras que contenían datos atípicos al ser evaluadas respecto al valor del negocio, por lo tanto, la cantidad de muestras empleadas para la elaboración del modelo fueron 44.

Para comenzar se realizó un gráfico de la matriz de correlación entre las diferentes posibles variables analizadas para el modelo, las cuales se muestran en la gráfica a continuación. El color azul significa una relación directa entre las variables y el rojo una relación inversa. Así mismo, el tamaño e intensidad del color, muestra el grado de correlación entre las variables.

Gráfico 14. Matriz de correlación excluyendo outliers (44 muestras)



13.4.1 Modelaciones con distancia a pie.

Partimos de la aplicación del modelo lineal con el software estadístico R, ingresando todas las variables preseleccionadas, utilizando en este caso la variable distancia a pie, sin tener en cuenta los datos atípicos respecto al valor del negocio, ni las variables distancia lineal euclidiana, y tampoco la distancia en carro.

Los resultados de la modelación se presentan a continuación:

Tabla 32. Modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), incluyendo todas las variables

Residuals:
 Min 1Q Median 3Q Max

-59479556 -20088035 -2463066 18587602 80166282

Coefficients:

	Estimate	Std.Error	t value	Pr(> t)	
(Intercept)	107715632	70079953	1,537	0,134429	
Aconstruida	1451967	329553	4,406	0,000117	***
Ascensor	20796683	17038913	1,221	0,231463	
Carros	16292971	10922890	1,492	0,145902	
CuartoUtil	6197466	15419764	0,402	0,690503	
Distancia.a.Pie..m.	12038	23442	0,514	0,611223	
Estrato	27757139	10323918	2,689	0,011441	*
FittoCorvini	-14397867	15012919	-0,959	0,344963	
Habitaciones	-5073890	11540776	-0,440	0,663241	
NivelPiso	-972816	2166969	-0,449	0,656605	
Poniente	3698716	12478282	0,296	0,768892	
Porteria	-8287784	11204330	-0,740	0,465050	
ZonaComun	23537485	9463679	2,487	0,018467	*

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

Residual standard error: 33300000 on 31 degrees of freedom

Multiple R-squared: 0.7638,

Adjusted R-squared: 0.6724

F-statistic: 8.356 on 12 and 31 DF,

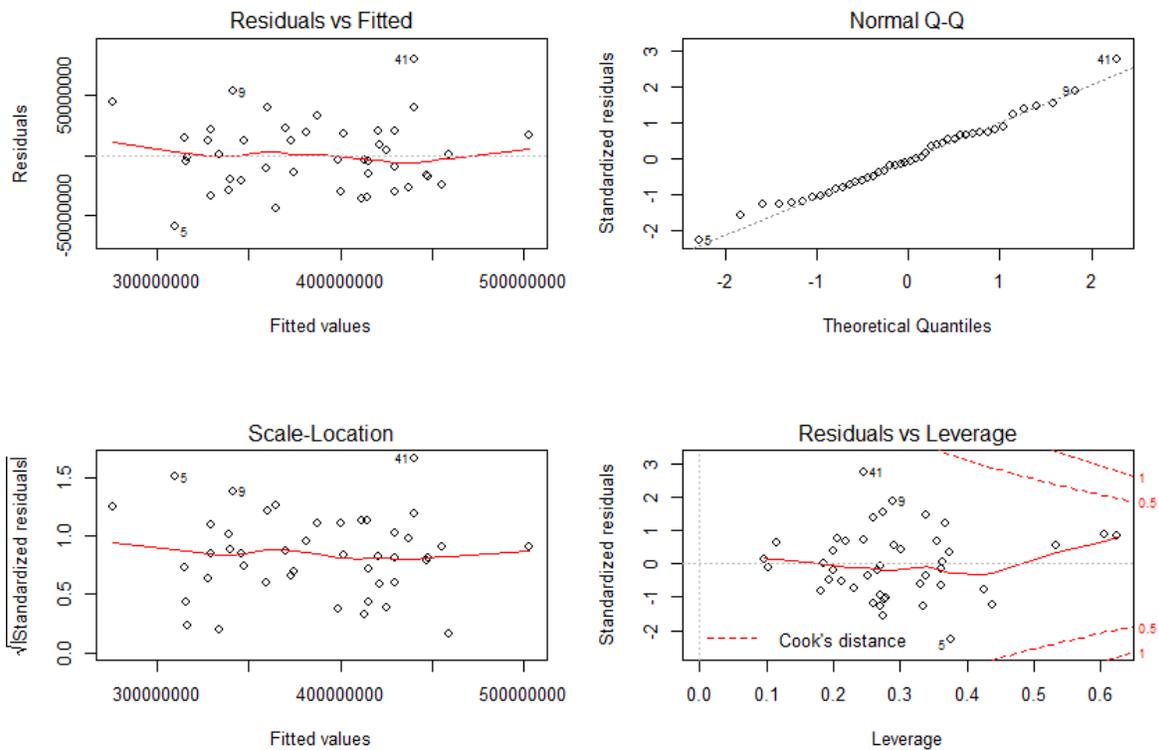
p-value: 0.0000009807

En los resultados obtenidos con la modelación en R, se observa que solo algunos de los coeficientes de las variables área construida, estrato, y zona común son estadísticamente significativos, dado que se obtienen probabilidades (p-value) que permiten rechazar la hipótesis nula, al ser inferiores a 0,05. El coeficiente de determinación R² ajustado del modelo indica en forma porcentual que el 67.24% del valor del inmueble, es explicado por las variables independientes analizadas, con un p-crítico del modelo de 9.807E-07 que aceptaría el modelo. Sin embargo, se observa un resultado del estadístico F inferior a 10. Por lo tanto el modelo no tiene una buena capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

Gráfico 15. Comportamiento de los residuos del modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), incluyendo todas las variables

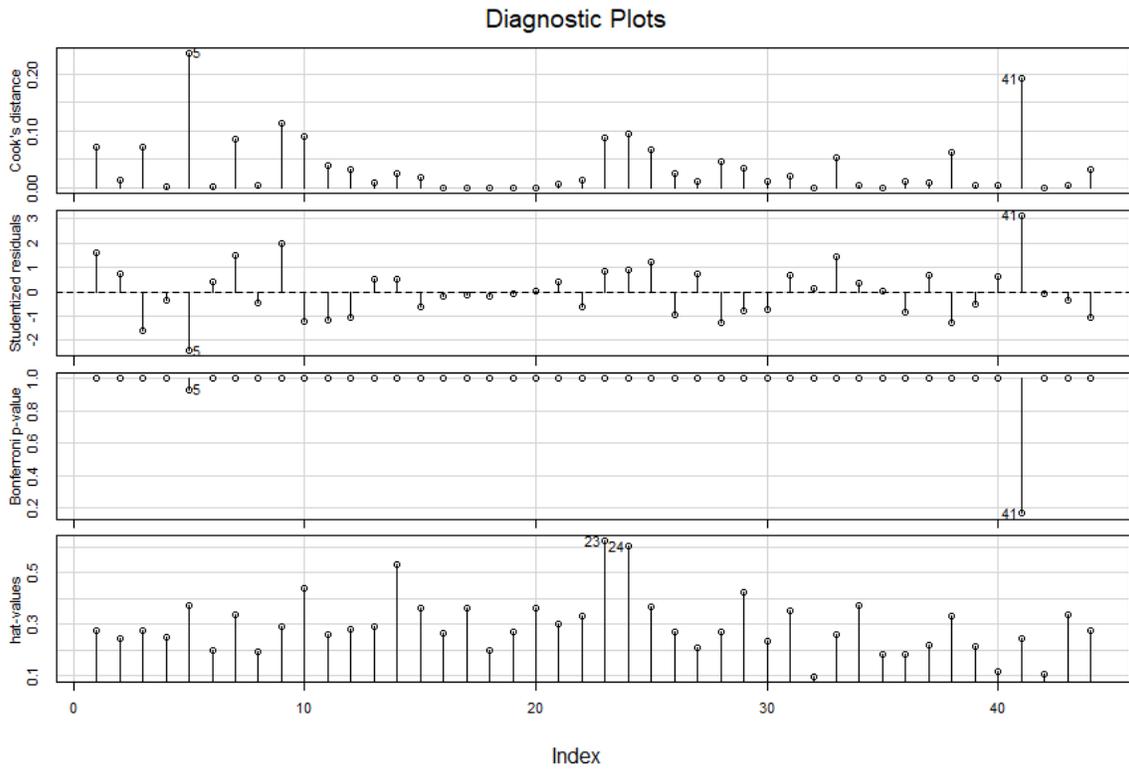
lm(Valor.del.negocio ~ Aconstruida + Ascensor + Carros + CuartoUtil + Dista ...



En la gráfica de los residuos vs los residuos ajustados, se observa que los residuos siguen patrones lineales, y están distribuidos uniformemente a lado y lado de la línea, que se puede deducir visualmente que cumplen con el criterio de la regresión lineal. En la gráfica normal Q-Q se puede apreciar que los residuos se distribuyen normalmente al encontrarse en general alineados. En el diagrama de scale-location, se observa la homocedasticidad de los residuos. En la gráfica de Residuals vs Leverage, que permite encontrar casos influyentes o que no se ajustan bien a la tendencia de la mayoría de los casos, no se observan casos que se encuentren fuera de la distancia de Cook.

En la gráfica de índices de influencia a continuación, se observa la influencia de cada una de las muestras del estudio utilizadas para el modelo:

Gráfico 16. Gráfica de índices de influencia del modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), incluyendo todas las variables



A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos:

Inflación de la varianza

Aconstruida	Ascensor	Carros	CuartoUtil
1.935643	2.731728	1.496813	1.964272
Distancia.a.Pie..m.	Estrato	FittoCorvini	Habitaciones
1.543049	2.840361	1.685259	1.777428

NivelPiso	Poniente	Porteria	ZonaComun
1.811230	1.225694	4.866361	3.089906

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 0.19186, df = 1, p\text{-value} = 0.6614$$

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por ende se asume la homocedasticidad.

Durbin-Watson test

$$DW = 1.927, p\text{-value} = 0.2898$$

De acuerdo con el resultado de la prueba de Durbin – Watson, el estadístico es cercano a 2, y dado que el p-value de la prueba es mayor a 0.05 no se rechaza la hipótesis nula de no autocorrelación y por consiguiente se concluye que no hay autocorrelación serial de primer orden en el modelo.

RESET test

$$RESET = 0.51182, df1 = 24, df2 = 7, p\text{-value} = 0.8953$$

El resultado de p-value obtenido en la prueba de RESET indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

Verificación de Outliers Test Bonferroni

Se aplicó el test de Bonferroni, con el fin de identificar la presencia de posibles datos atípicos que pudieran incidir en el resultado de las pruebas y en el R2 del modelo. Estos

datos atípicos se observan a su vez en la gráfica de índices de influencia anteriormente mostrada.

rstudent	unadjusted	p-value	Bonferroni p
41	3.141797	0.0037609	0.16548

13.4.1.1 Modelaciones metodología paso a paso (Step Wise).

A partir de este primer resultado del modelo, con el software R se realizó el análisis paso a paso, con la metodología atrás/adelante y utilizando el criterio de Akaike, con el fin de determinar las variables que mejor explicarían el modelo, y cuyos resultados se presentan a continuación:

Tabla 33. Modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), metodología paso a paso

Residuals:				
Min	1Q	Median	3Q	Max
-59167221	-18586140	-2361063	18889713	76034525
Coeficientes:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	123165644	42376141	2.906	0.00607 **
Aconstruida	1348401	242079	5.570	0.00000221 ***
Carros	16759180	9574104	1.750	0.08811 .
Estrato	30450056	7455610	4.084	0.00022 ***
FittoCorvini	-24105809	11743636	-2.053	0.04704 *
ZonaComun	20871685	6757838	3.089	0.00375 **

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

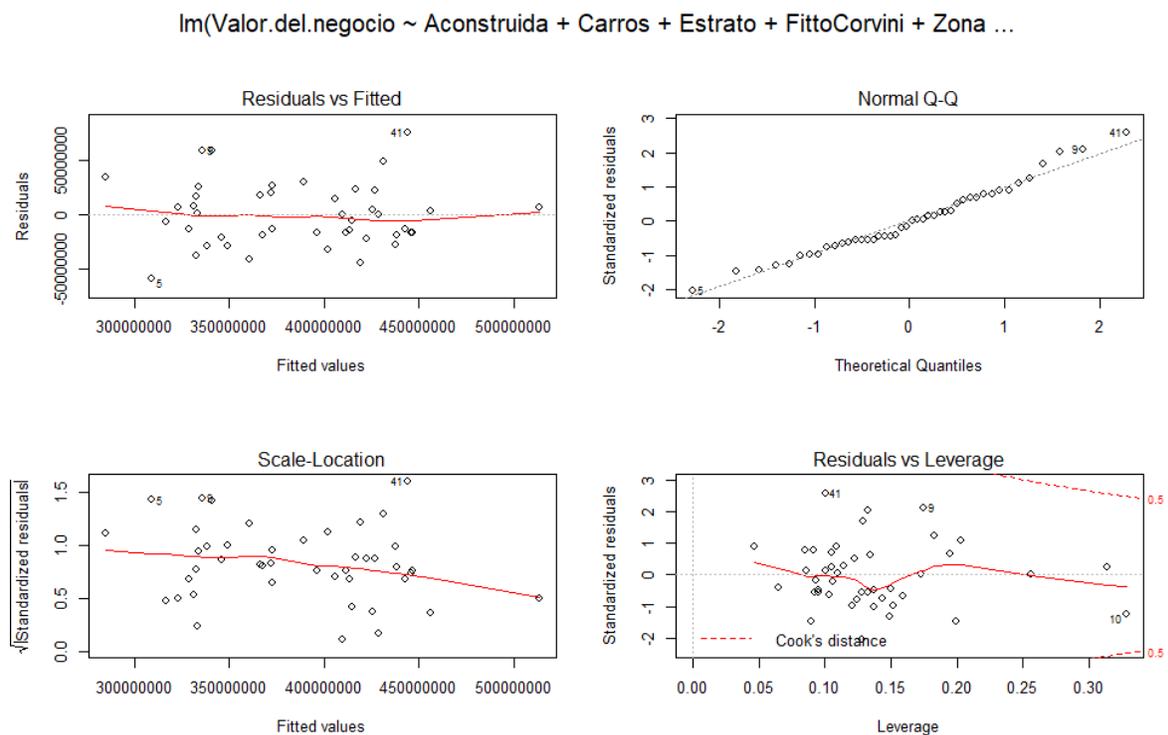
Residual standard error: 31080000 on 38 degrees of freedom
Multiple R-squared: 0.7478,
Adjusted R-squared: 0.7146
F-statistic: 22.54 on 5 and 38 DF,
p-value: 1.95e-10

En los resultados obtenidos con la modelación en R, se observa que las variables área construida, cantidad de carros, estrato, estado de conservación y la zona común son las más significativas estadísticamente, dado los resultados obtenidos de probabilidades (p-value).

El coeficiente de determinación R^2 ajustado del modelo indica en forma porcentual que el 71.46% del valor del inmueble, es explicado por las variables independientes analizadas, con un p-crítico del modelo de $1.95E-10$, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es mayor a 10 y por lo tanto el modelo tiene una muy buena capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

Gráfico 17. Comportamiento de los residuos del modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), metodología paso a paso



En la gráfica de los residuos vs los residuos ajustados, se observa que los residuos siguen patrones lineales, y están distribuidos uniformemente a lado y lado de la línea, de lo cual se puede deducir visualmente que cumplen con el criterio de la regresión lineal. En la gráfica normal Q-Q se puede observar que los residuos se distribuyen normalmente al encontrarse en general bien alineados. En el diagrama de scale-location se observa la homocedasticidad de los residuos. En la gráfica de Residuals vs Leverage, que permite

encontrar casos influyentes, o que no se ajustan bien a la tendencia de la mayoría de los casos, no se observan casos que se encuentren fuera de la distancia de Cook.

A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos:

Inflación de la varianza

Aconstruida	Carros	Estrato	FittoCorvini	ZonaComun
1.198953	1.320091	1.700457	1.183740	1.808655

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 0.69269, df = 1, p\text{-value} = 0.4053$$

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por ende se asume la homocedasticidad.

Durbin-Watson test

$$DW = 1.8663, p\text{-value} = 0.2603$$

Por lo tanto, el resultado que es muy cercano a 2 indica que hay una correlación que es entre moderada y leve entre las variables y se podría decir que ninguna de las variables tiene una correlación sustancial.

RESET test

$$RESET = 0.14557, df1 = 2, df2 = 36, p\text{-value} = 0.865$$

El resultado de p-value obtenido en la prueba de RESET, mayor a 0.05, indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

13.4.1.2 Modelo 1 - Modelación incluyendo variable distancia a pie.

Como la investigación se va a verificar la incidencia de la variable distancia, se introduce la variable distancia a pie en el modelo obtenido anteriormente, logrando los siguientes resultados:

Tabla 34. Modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia a pie

Residuals:				
Min	1Q	Median	3Q	Max
-59393981	-18890149	-2738423	18620414	76257334
Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	117808603	48727724	2.418	0.020662 *
Aconstruida	1347432	245185	5.496	0.00000303 ***
Carros	16718293	9697168	1.724	0.093048 .
Distancia.a.Pie..m.	4310	18569	0.232	0.817753
Estrato	30749669	7659771	4.014	0.000279 ***
FittoCorvini	-23539693	12140206	-1.939	0.060158 .
ZonaComun	20628831	6923110	2.980	0.005073 **

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

Residual standard error: 31470000 on 37 degrees of freedom

Multiple R-squared: 0.7482, Adjusted R-squared: 0.7074

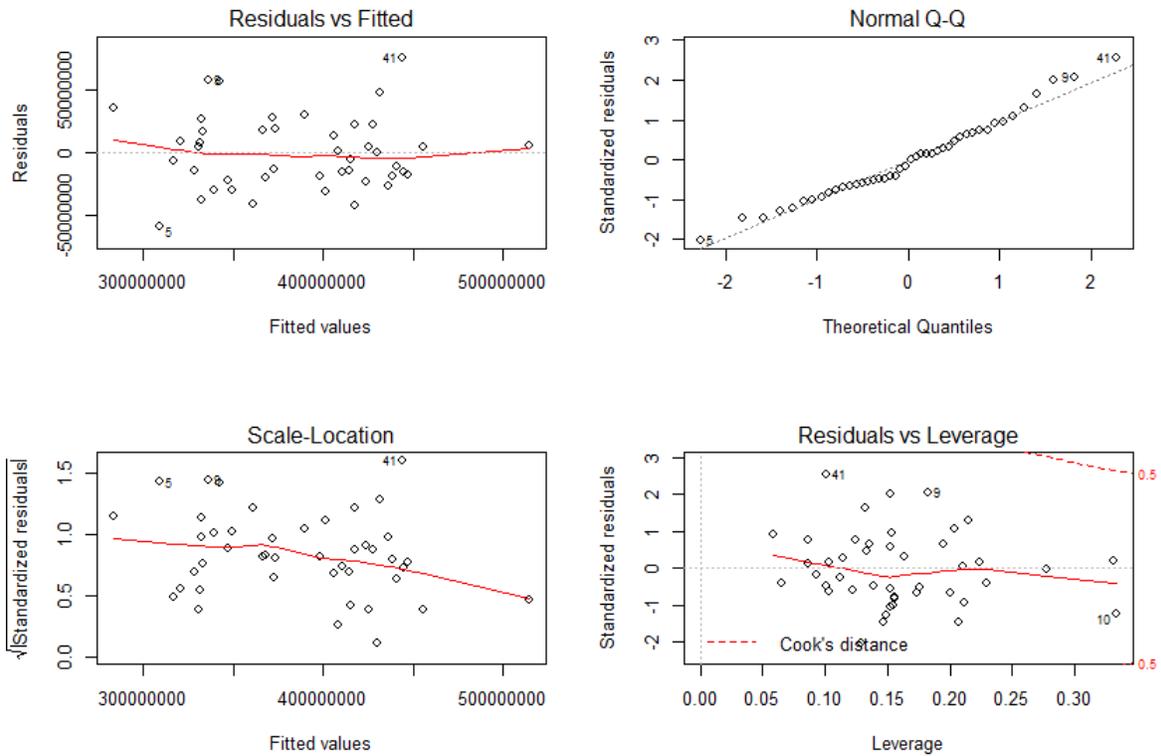
F-statistic: 18.32 on 6 and 37 DF, p-value: 9.63e-10

En los resultados obtenidos con la modelación en R, se observa que las variables área construida, estrato y zona común siguen siendo significativas estadísticamente, dado los resultados de probabilidades (p-value). La variable distancia presenta un valor de probabilidad asociado alto comparado con las otras variables utilizadas y por lo tanto explicaría en menor magnitud el valor del inmueble. El coeficiente de determinación R² ajustado del modelo indica en forma porcentual que el 70.74% del valor del inmueble, es explicado por las variables independientes analizadas, con un p-crítico del modelo de 1.95E-10, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es mayor a 10 y por lo tanto el modelo tiene una buena capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal múltiple obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

Gráfico 18. Comportamiento de los residuos del modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia a pie

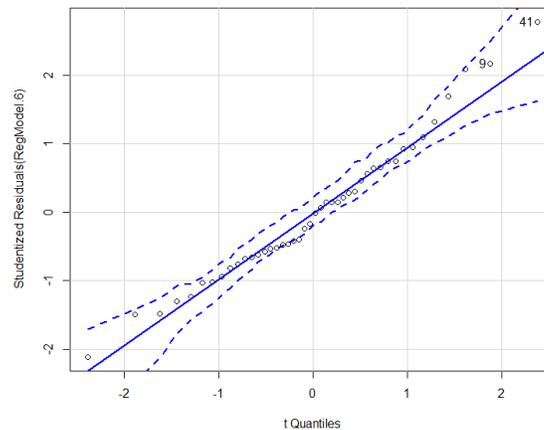
lm(Valor.del.negocio ~ Aconstruida + Carros + Distancia.a.Pie..m. + Estrato ...



De acuerdo con el resultado observado en el gráfico Residuals vs Fitted se puede decir que es un buen modelo que cumple con los supuestos de linealidad.

Así mismo, se observa en el gráfico Normal Q-Q la normalidad de los residuos, cumpliéndose de igual forma con el supuesto del modelo. En el gráfico de cuantiles se observa que los residuos studentizados se encuentran dentro de los intervalos, lo cual indica que siguen una distribución normal.

Gráfico 19. Gráfico de cuantiles del modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia a pie

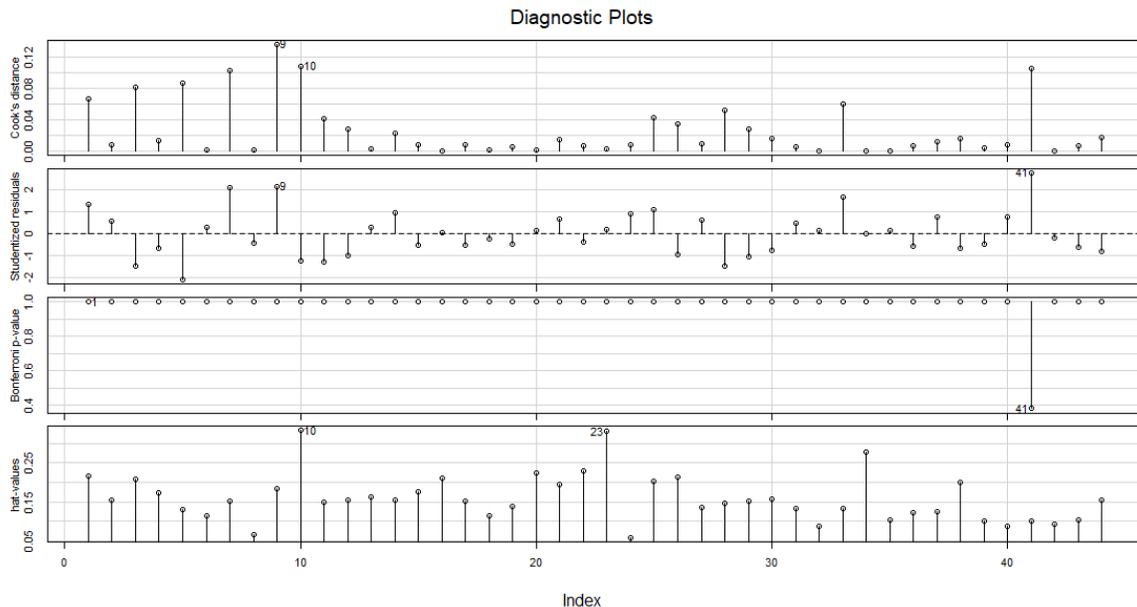


En la gráfica Scale-Location se observa que los residuos están distribuidos equitativamente a lo largo de los rangos de los predictores con lo cual se puede comprobar el supuesto de la varianza igual u homocedasticidad. Este supuesto también es verificado a través del test Breusch-Pagan.

Finalmente, en la gráfica Residuals vs Leverage la cual ayuda a identificar los casos que son más influyentes en la regresión lineal (sean atípicos o no). En todos los casos se observa que los datos se encuentran dentro de la distancia Cook, aun cuando el programa muestra que hay unos datos que si se eliminaran podrían mejorar los resultados del coeficiente de determinación.

En la gráfica de índices de influencia a continuación, se observa la influencia de cada una de las muestras del estudio utilizadas para el modelo:

Gráfico 20. Gráfico de índices de influencia con distancia a pie excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia a pie



A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos:

Inflación de la varianza

Aconstruida	Carros	Distancia.a.Pie..m.	Estrato
1.199301	1.320527	1.083728	1.750172
	FittoCorvini	ZonaComun	
	1.233540	1.850941	

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 3.4768, df = 6, p\text{-value} = 0.7471$$

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por ende se asume la homocedasticidad.

Durbin-Watson test

DW = 1.8372, p-value = 0.2188

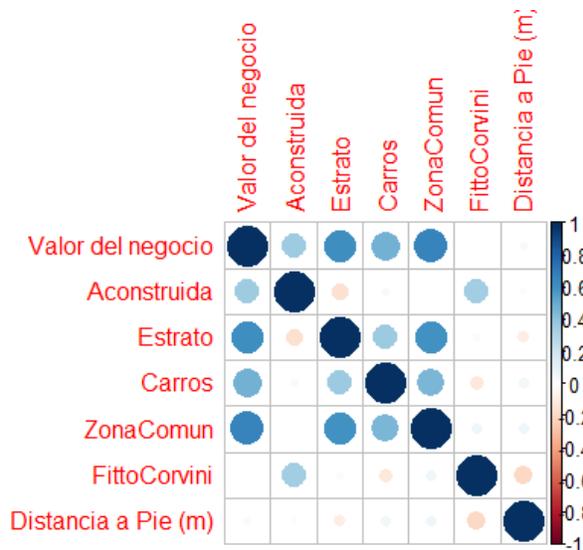
Por lo tanto, el resultado que es muy cercano a 2 indica que hay una correlación que es entre moderada y leve entre las variables y se podría decir que ninguna de las variables tiene una correlación sustancial.

Tabla 35. Tabla de correlación entre variables. Modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia a pie

	Aconstruida	Carros	Distancia. a.Pie..m.	Estrato	FittoCorvini	Valor.del. negocio	ZonaComun
Aconstruida	1	-0,035	-0,022	-0,164	0,349	0,358	0,005
Carros	-0,035	1	0,057	0,363	-0,121	0,480	0,455
Distancia.a.Pie..m.	-0,022	0,057	1	-0,095	-0,204	0,034	0,063
Estrato	-0,164	0,363	-0,095	1	0,028	0,613	0,610
FittoCorvini	0,349	-0,121	-0,204	0,028	1	0,005	0,063
Valor.del.negocio	0,358	0,480	0,034	0,613	0,005	1	0,668
Zona Común	0,005	0,455	0,063	0,610	0,063	0,668	1

Visualmente la correlación entre las variables se observa a continuación, en la cual el color azul significa una correlación directa y el rojo una correlación inversa, y la intensidad del color y su tamaño, la magnitud de la correlación.

Gráfico 21. Correlación del modelo de regresión lineal con distancia a pie excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia a pie



Se observa una correlación moderada entre la variable estrato y la variable zona común. Pero como ya se mostró anteriormente el resultado de la prueba Durbin-Watson el modelo cumple con el supuesto respecto a esta característica.

RESET test

$$\text{RESET} = 1.9928, \text{df1} = 12, \text{df2} = 25, \text{p-value} = 0.07063$$

El resultado de p-value obtenido en la prueba de RESET, mayor a 0.05, indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

Verificación de Outliers Test Bonferroni

Se aplicó el test de Bonferroni, con el fin de identificar la presencia de posibles datos atípicos que pudieran incidir en el resultado de las pruebas y en el R2 del modelo. Estos datos atípicos se observan a su vez en la gráfica de índices de influencia anteriormente mostrada.

rstudent	unadjusted	p-value	Bonferroni p
41	2.77794	0.0086343	0.37991

De acuerdo con el test de valores atípicos de Bonferroni, la muestra ubicada en la fila 41 representa un valor atípico que como se observa en las gráficas puede influir en el resultado del coeficiente de determinación R^2 del modelo.

13.4.2 Modelaciones con distancia en carro.

Partimos de la aplicación del modelo lineal con el software estadístico R, para todas las variables preseleccionadas, utilizando en este caso la variable distancia en carro, sin tener en cuenta los datos atípicos respecto al valor del negocio, ni las variables distancia lineal euclidiana, ni la distancia a pie.

Tabla 36. Modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras) incluyendo todas las variables

Residuals:

Min	1Q	Median	3Q	Max
-58640462	-20035227	-4660601	15670647	80214798

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	158171342	85465246	1.851	0.07376 .
Aconstruida	1465522	331009	4.427	0.00011 ***
Ascensor	18670011	16400806	1.138	0.26369
Carros	15531735	11058337	1.405	0.17010
CuartoUtil	6265896	15438188	0.406	0.68763
Distancia.Car..m.	-8133	18681	-0.435	0.66632
Estrato	22679833	12437541	1.823	0.07788 .
FittoCorvini	-18109019	14317468	-1.265	0.21536
Habitaciones	-5957247	11573792	-0.515	0.61040
NivelPiso	-1054860	2154539	-0.490	0.62786
Poniente	1456207	13273428	0.110	0.91335
Porteria	-4903491	9725987	-0.504	0.61771
ZonaComun	22565403	9643515	2.340	0.02590 *

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

Residual standard error: 33340000 on 31 degrees of freedom

Multiple R-squared: 0.7633, Adjusted R-squared: 0.6717

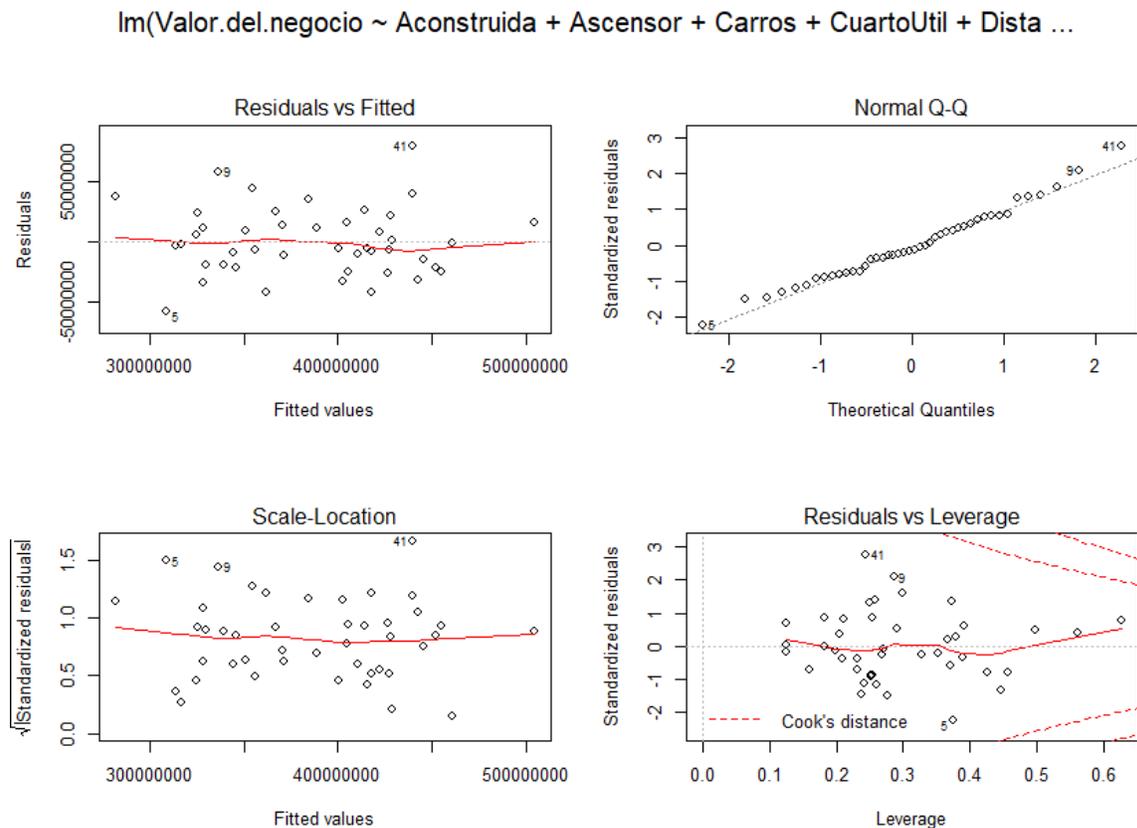
F-statistic: 8.33 on 12 and 31 DF, p-value: 0.000001014

En los resultados obtenidos con la modelación en R, se observa que las variables área construida, estrato y la zona común son las significativas estadísticamente, dado los resultados de probabilidades (p-value). El coeficiente de determinación R^2 ajustado del modelo indica en forma porcentual que el 67.17% del valor del inmueble, es explicado por

las variables independientes analizadas, con un p-crítico del modelo de 1.014E-6, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es menor a 10 y por lo tanto el modelo no tiene una buena capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal múltiple obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

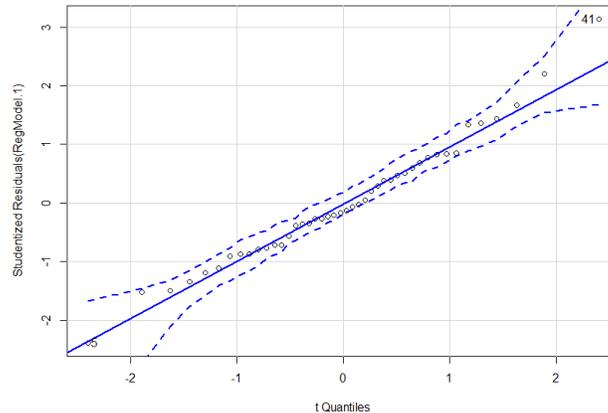
Gráfico 22. Comportamiento de los residuos del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras) incluyendo todas las variables



De acuerdo con el resultado observado en el gráfico Residuals vs Fitted se puede decir que es un buen modelo que cumple con los supuestos de linealidad.

Así mismo, se observa en el gráfico Normal Q-Q la normalidad de los residuos, cumpliéndose de igual forma con el supuesto del modelo. En el gráfico de cuantiles se observa que los residuos studentizados se encuentran dentro de los intervalos, lo cual indica que siguen una distribución normal.

Gráfico 23. Cuantiles del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), incluyendo todas las variables

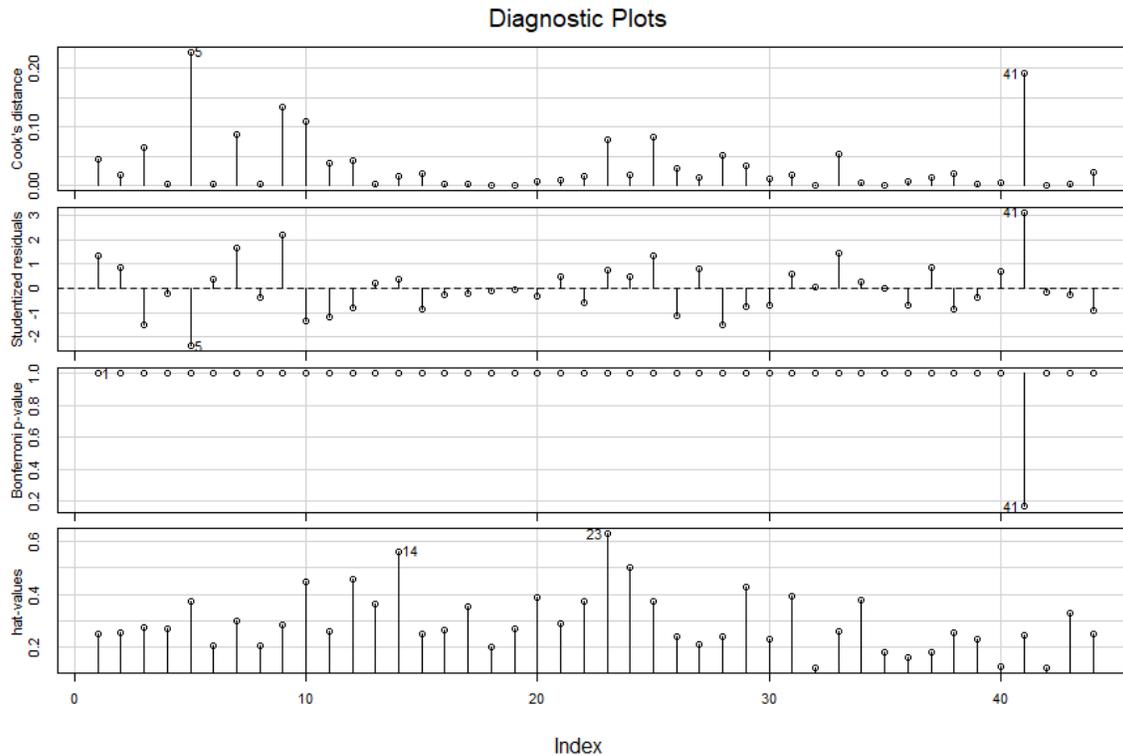


En la gráfica Scale-Location se observa que los residuos están distribuidos equitativamente a lo largo de los rangos de los predictores con lo cual se puede comprobar el supuesto de la varianza igual u homocedasticidad.

Finalmente, en la gráfica Residuals vs Leverage la cual ayuda a identificar los casos que son más influyentes en la regresión lineal (sean atípicos o no). En todos los casos se observa que los datos se encuentran dentro de la distancia Cook, aun cuando el programa muestra que hay unos datos que si se eliminaran podrían mejorar los resultados del coeficiente de determinación.

En la gráfica de índices de influencia a continuación, se observa la influencia de cada una de las muestras del estudio utilizadas para el modelo:

Gráfico 24. Índices de influencia del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), incluyendo todas las variables



A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos:

Inflación de la varianza

Aconstruida	Ascensor	Carros	CuartoUtil
1.948152	2.524948	1.530525	1.964297
Distancia.Car..m.	Estrato	FittoCorvini	Habitaciones
3.113050	4.112651	1.529104	1.783371
NivelPiso	Poniente	Porteria	ZonaComun
1.786264	1.383588	3.658207	3.200842

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 0.090289, df = 1, p\text{-value} = 0.7638$$

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por ende se asume la homocedasticidad.

Durbin-Watson test

$$DW = 2.0681, p\text{-value} = 0.4846$$

Por lo tanto, el resultado que es muy cercano a 2 indica que hay una correlación que es entre moderada y leve entre las variables y se podría decir que ninguna de las variables tiene una correlación sustancial.

RESET test

$$RESET = 0.1124, df1 = 2, df2 = 29, p\text{-value} = 0.8941$$

El resultado de p-value obtenido en la prueba de RESET, mayor a 0.05, indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

13.4.2.1 Modelaciones metodología paso a paso

A partir de este primer resultado del modelo, con el software R se realizó el análisis paso a paso, con la metodología atrás/adelante y utilizando el criterio de Akaike, con el fin de determinar las variables que mejor explicarían el modelo, y cuyos resultados se presentan a continuación:

Tabla 37. Modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), metodología paso a paso

Residuals:				
Min	1Q	Median	3Q	Max
-59167221	-18586140	-2361063	18889713	76034525
Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	123165644	42376141	2.906	0.00607 **
Aconstruida	1348401	242079	5.570	0.00000221 ***
Carros	16759180	9574104	1.750	0.08811 .
Estrato	30450056	7455610	4.084	0.00022 ***
FittoCorvini	-24105809	11743636	-2.053	0.04704 *
ZonaComun	20871685	6757838	3.089	0.00375 **

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

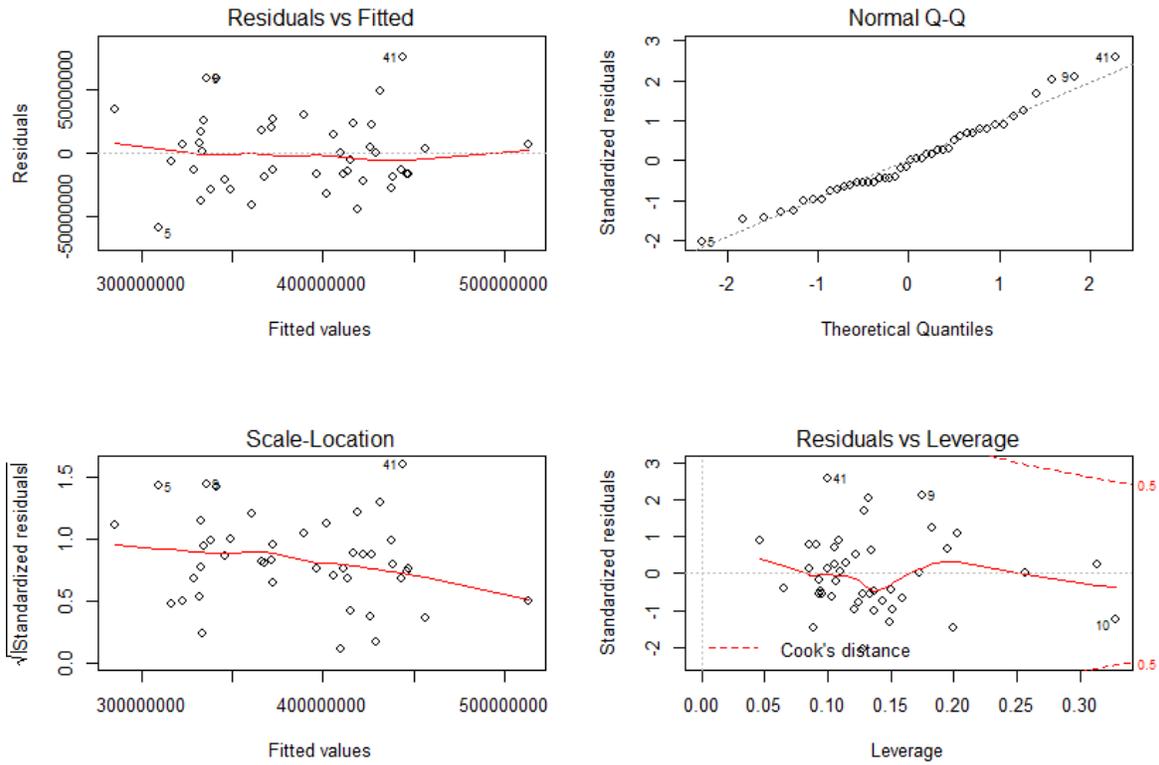
Residual standard error: 31080000 on 38 degrees of freedom
 Multiple R-squared: 0.7478, Adjusted R-squared: 0.7146
 F-statistic: 22.54 on 5 and 38 DF, p-value: 1.95e-10

En los resultados obtenidos con la modelación en R, se observa que las variables área construida, cantidad de carros, estrato, estado de conservación y la zona común siguen siendo las variables más significativas estadísticamente, a pesar de haber introducido la variable distancia en carro, dado los resultados obtenidos de probabilidades (p-value). El coeficiente de determinación R² ajustado del modelo indica en forma porcentual que el 71.46% del valor del inmueble, es explicado por las variables independientes analizadas, con un p-crítico del modelo de 1.95E-10, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es mayor a 10 y por lo tanto el modelo tiene una muy buena capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

Gráfico 25. Comportamiento de los residuos del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), metodología paso a paso

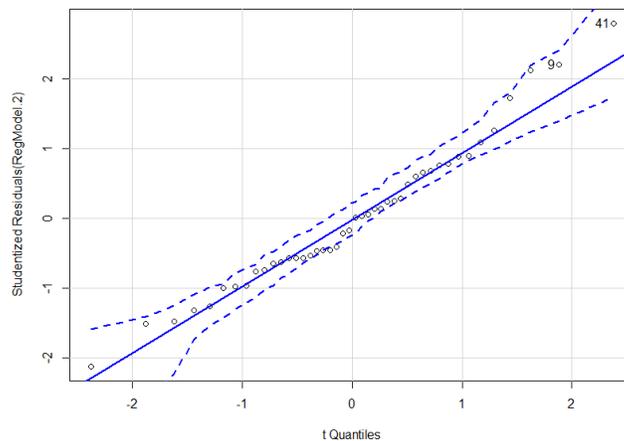
lm(Valor.del.negocio ~ Aconstruida + Carros + Estrato + FittoCorvini + Zona ...



De acuerdo con el resultado observado en el gráfico Residuals vs Fitted se puede decir que es un buen modelo que cumple con los supuestos de linealidad.

Así mismo, se observa en el gráfico Normal Q-Q la normalidad de los residuos, cumpliéndose de igual forma con el supuesto del modelo. En el gráfico de cuantiles se observa que los residuos studentizados se encuentran dentro de los intervalos, lo cual indica que siguen una distribución normal.

Gráfico 26. Cuantiles del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), metodología paso a paso

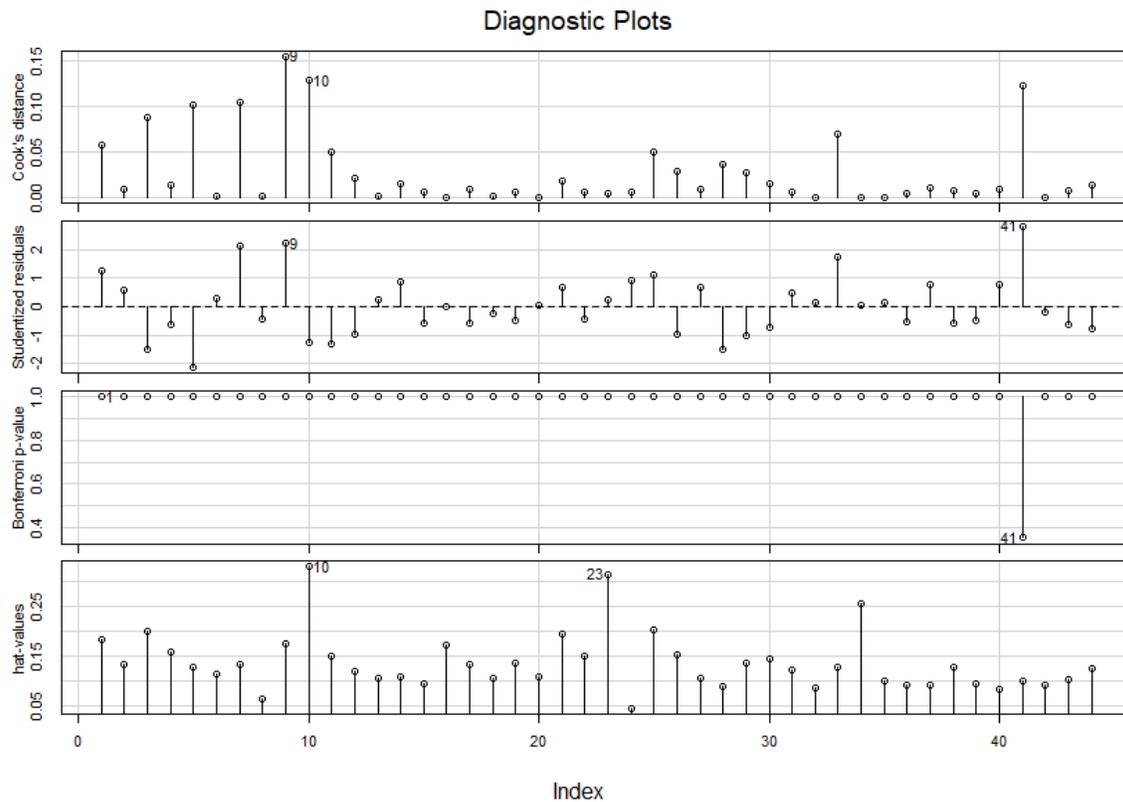


En la gráfica Scale-Location se observa que los residuos están distribuidos equitativamente a lo largo de los rangos de los predictores con lo cual se puede comprobar el supuesto de la varianza igual u homocedasticidad. Este supuesto también es verificado a través del test Breusch-Pagan.

Finalmente, en la gráfica Residuals vs Leverage la cual ayuda a identificar los casos que son más influyentes en la regresión lineal (sean atípicos o no). En todos los casos se observa que los datos se encuentran dentro de la distancia Cook, aun cuando el programa muestra que hay unos datos que si se eliminaran podrían mejorar los resultados del coeficiente de determinación.

En la gráfica de índices de influencia a continuación, se observa la influencia de cada una de las muestras del estudio utilizadas para el modelo:

Gráfico 27. Índices de influencia del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), metodología paso a paso



A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos:

Inflación de la varianza

Aconstruida	Carros	Estrato	FittoCorvini	ZonaComun
1.198953	1.320091	1.700457	1.183740	1.808655

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 3.0917, df = 5, p\text{-value} = 0.6858$$

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por ende se asume la homocedasticidad del modelo.

Durbin-Watson test

$$DW = 1.8663, p\text{-value} = 0.2603$$

De acuerdo con el resultado de la prueba de Durbin – Watson, el estadístico es cercano a 2, y dado que el p-value de la prueba es mayor a 0.05 no se rechaza la hipótesis nula de no autocorrelación. Por consiguiente, se concluye que no hay autocorrelación serial de primer orden.

RESET test

$$RESET = 0.14557, df1 = 2, df2 = 36, p\text{-value} = 0.865$$

El resultado de p-value obtenido en la prueba de RESET indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

13.4.2.2 Modelaciones incluyendo variable distancia en carro.

Como la investigación se va a verificar la incidencia de la variable distancia, se introduce la variable distancia en carro en el modelo obtenido anteriormente, logrando los siguientes resultados:

Tabla 38. Modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia en carro

Residuals:				
Min	1Q	Median	3Q	Max
-60317882	-18292647	-3721024	19668765	75016420
Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	142584873	69740760	2.044	0.04807 *
Aconstruida	1348196	244917	5.505	0.00000294 ***
Carros	16253684	9791579	1.660	0.10537
Distancia.Car..m.	-5713	16182	-0.353	0.72606
Estrato	28342145	9620151	2.946	0.00554 **
FittoCorvini	-25134858	12233610	-2.055	0.04703 *
ZonaComun	20513152	6912060	2.968	0.00523 **

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

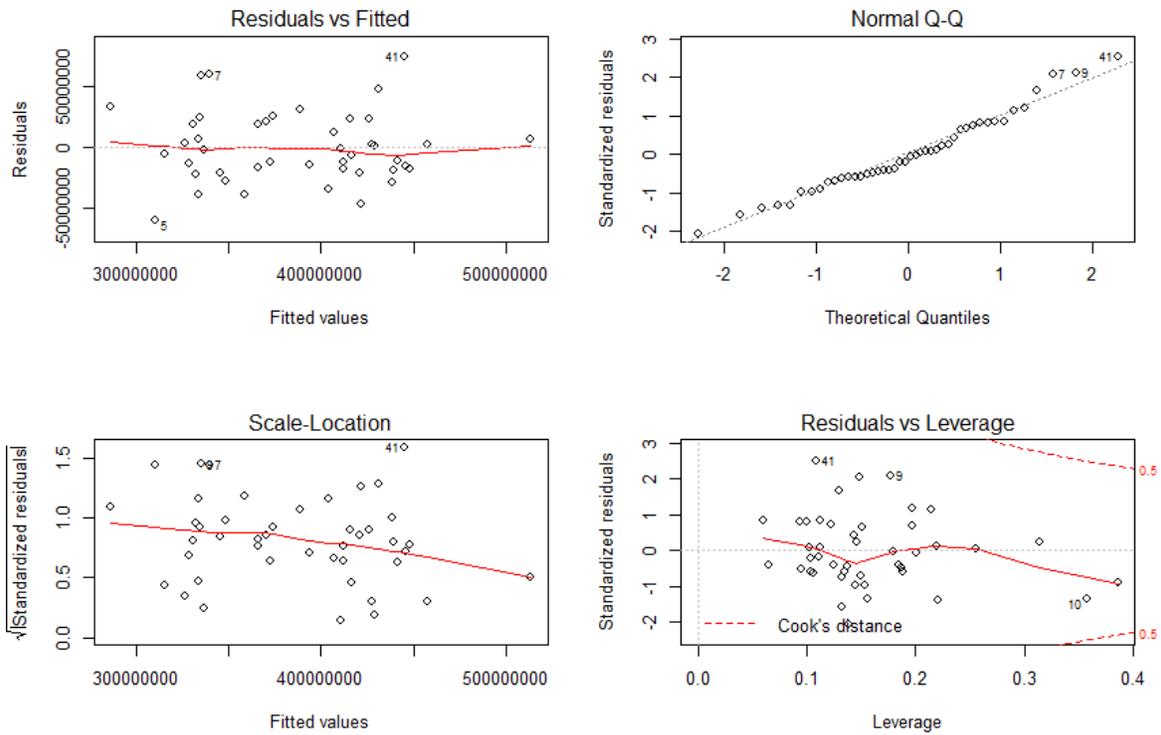
Residual standard error: 31440000 on 37 degrees of freedom
 Multiple R-squared: 0.7487, Adjusted R-squared: 0.7079
 F-statistic: 18.37 on 6 and 37 DF, p-value: 9.307e-10

En los resultados obtenidos con la modelación en R, se observa que las variables área construida, cantidad de carros, estrato, estado de conservación y la zona común siguen siendo las variables más significativas estadísticamente, a pesar de haber introducido la variable distancia en carro, dado los resultados obtenidos de probabilidades (p-value). El coeficiente de determinación R² ajustado del modelo indica en forma porcentual que el 70.79% del valor del inmueble, es explicado por las variables independientes analizadas, con un p-crítico del modelo de 9.307E-10, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es mayor a 18 y por lo tanto el modelo tiene una buena capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

Gráfico 28. Comportamiento de los residuos del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia en carro

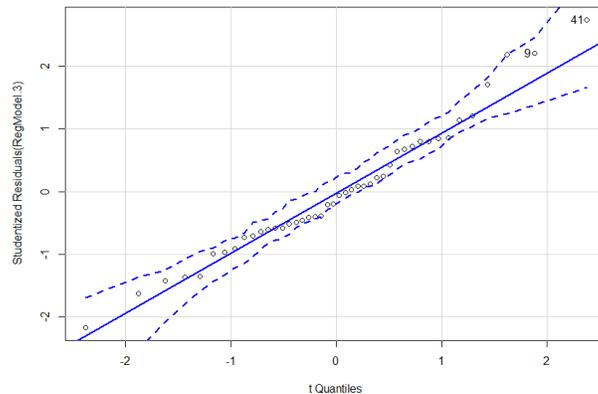
lm(Valor.del.negocio ~ Aconstruida + Carros + Distancia.Car..m. + Estrato + ...)



De acuerdo con el resultado observado en el gráfico Residuals vs Fitted se puede decir que el modelo que cumple con los supuestos de linealidad.

Así mismo, se observa en el gráfico Normal Q-Q la normalidad de los residuos, cumpliéndose de igual forma con el supuesto del modelo. En el gráfico de cuantiles se observa que los residuos studentizados se encuentran dentro de los intervalos, lo cual indica que siguen una distribución normal.

Gráfico 29. Cuantiles del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia en carro

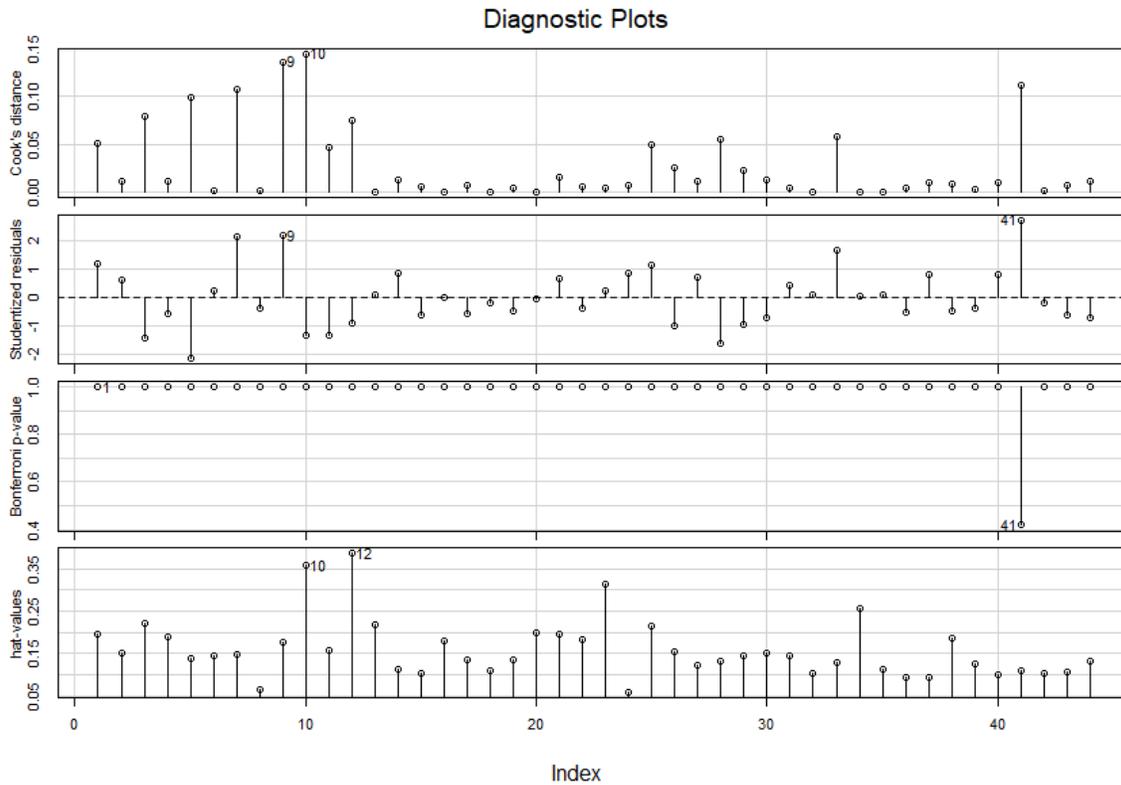


En la gráfica Scale-Location se observa que los residuos están distribuidos equitativamente a lo largo de los rangos de los predictores con lo cual se puede comprobar el supuesto de la varianza igual u homocedasticidad.

Finalmente, en la gráfica Residuals vs Leverage la cual ayuda a identificar los casos que son más influyentes en la regresión lineal (sean atípicos o no). En todos los casos se observa que los datos se encuentran dentro de la distancia Cook, aun cuando el programa muestra que hay unos datos que si se eliminaran podrían mejorar los resultados del coeficiente de determinación.

En la gráfica de índices de influencia a continuación, se observa la influencia de cada una de las muestras del estudio utilizadas para el modelo:

Gráfico 30. Índices de influencia del modelo de regresión lineal con distancia en carro excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia en carro



A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos:

Inflación de la varianza

Aconstruida	Carros	Distancia.Car..m.	Estrato
1.198960	1.348937	2.625801	2.765933
	FittoCorvini	ZonaComun	
	1.254987	1.848561	

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 0.64214, df = 1, p\text{-value} = 0.4229$$

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por ende se asume la homocedasticidad.

Durbin-Watson test

$$DW = 1.9034, p\text{-value} = 0.3109$$

De acuerdo con el resultado de la prueba de Durbin – Watson, el estadístico es cercano a 2, y dado que el p-value de la prueba es mayor a 0.05 no se rechaza la hipótesis nula de no autocorrelación. Por consiguiente, se concluye que no hay autocorrelación serial de primer orden.

RESET test

$$RESET = 0.096605, df1 = 2, df2 = 35, p\text{-value} = 0.9082$$

El resultado de p-value obtenido en la prueba de RESET indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

13.4.3 Modelaciones con distancia euclidiana (lineal).

Partimos de la aplicación del modelo lineal con el software estadístico R, para todas las variables preseleccionadas, utilizando en este caso la variable distancia euclidiana (lineal), sin tener en cuenta los datos atípicos respecto al valor del negocio, ni las variables distancia en carro, ni distancia a pie.

Tabla 39. Modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras) incluyendo todas las variables

Residuals:

Min 1Q Median 3Q Max
 -58975652 -20602501 -2594003 18032357 79812809

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	113146861	71266458	1.588	0.122512
Aconstruida	1467963	332362	4.417	0.000113 ***
Ascensor	20183434	17113878	1.179	0.247227
Carros	16417323	10954010	1.499	0.144054
CuartoUtil	6314861	15452320	0.409	0.685593
Distancia.Lineal.m	10994	29546	0.372	0.712355
Estrato	27694027	10754505	2.575	0.015015 *
FittoCorvini	-15074030	15091251	-0.999	0.325596
Habitaciones	-5735812	11552629	-0.496	0.623048
NivelPiso	-1067193	2156045	-0.495	0.624107
Poniente	4456756	12797309	0.348	0.730001
Porteria	-8017475	12011254	-0.667	0.509394
ZonaComun	23234805	9482539	2.450	0.020119 *

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

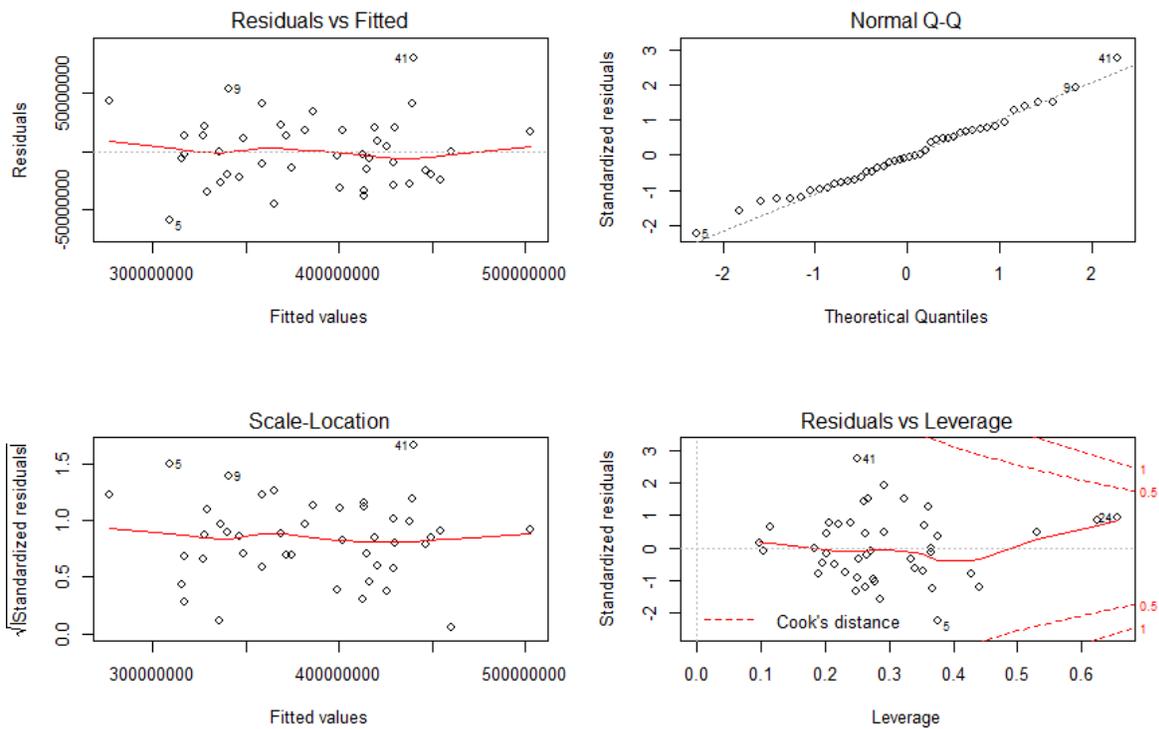
Residual standard error: 33360000 on 31 degrees of freedom
 Multiple R-squared: 0.7629, Adjusted R-squared: 0.6711
 F-statistic: 8.312 on 12 and 31 DF, p-value: 0.000001038

En los resultados obtenidos con la modelación en R, se observa que las variables área construida, estrato y la zona común son las significativas estadísticamente, dado los resultados de probabilidades (p-value). El coeficiente de determinación R² ajustado del modelo indica en forma porcentual que el 67.11% del valor del inmueble, es explicado por las variables independientes analizadas, con un p-crítico del modelo de 1.038E-6, casi cero, que aceptaría el modelo. El estadístico F es menor a 10 y por lo tanto el modelo no tiene una buena capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal múltiple obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

Gráfico 31. Comportamiento de los residuos del modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras) incluyendo todas las variables

lm(Valor.del.negocio ~ Aconstruida + Ascensor + Carros + CuartoUtil + Dista ...

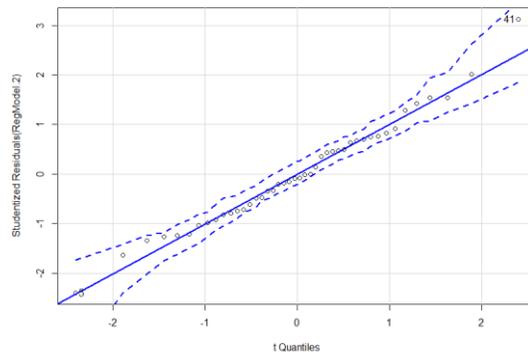


De acuerdo con el resultado observado en el gráfico Residuals vs Fitted se puede decir que es un buen modelo que cumple con los supuestos de linealidad.

En la gráfica Scale-Location se observa que los residuos están distribuidos equitativamente a lo largo de los rangos de los predictores con lo cual se puede comprobar el supuesto de la varianza igual u homocedasticidad.

Así mismo, se observa en el gráfico Normal Q-Q la normalidad de los residuos, cumpliéndose de igual forma con el supuesto del modelo. En el gráfico de cuantiles se observa que los residuos studentizados se encuentran dentro de los intervalos, lo cual indica que siguen una distribución normal.

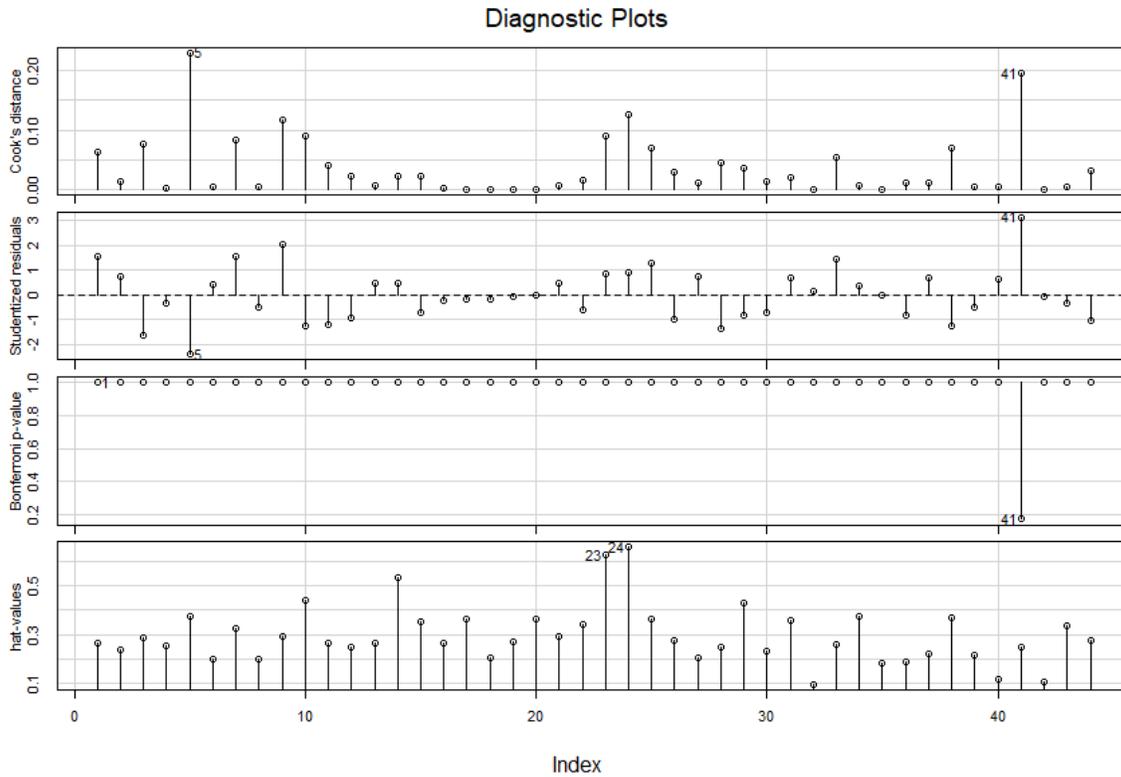
Gráfico 32. Cuantiles del modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras) incluyendo todas las variables



Finalmente, en la gráfica Residuals vs Leverage la cual ayuda a identificar los casos que son más influyentes en la regresión lineal (sean atípicos o no). En todos los casos se observa que los datos se encuentran dentro de la distancia Cook, aun cuando el programa muestra que hay unos datos que si se eliminaran podrían mejorar los resultados del coeficiente de determinación.

En la gráfica de índices de influencia a continuación, se observa la influencia de cada una de las muestras del estudio utilizadas para el modelo:

Gráfico 33. Índices de influencia del modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras) incluyendo todas las variables



A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos:

Inflación de la varianza

Aconstruida	Ascensor	Carros	CuartoUtil
1.960890	2.744777	1.499323	1.964672
Distancia.Lineal.m	Estrato	FittoCorvini	Habitaciones
1.904406	3.069883	1.696068	1.773945
NivelPiso	Poniente	Porteria	ZonaComun
1.785831	1.284003	5.570136	3.089804

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 0.16148, df = 1, p\text{-value} = 0.6878$$

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por ende se asume la homocedasticidad.

Durbin-Watson test

$$DW = 1.949, p\text{-value} = 0.3103$$

De acuerdo con el resultado de la prueba de Durbin – Watson, el estadístico es cercano a 2, y dado que el p-value de la prueba es mayor a 0.05 no se rechaza la hipótesis nula de no autocorrelación. Por consiguiente, se concluye que no hay autocorrelación serial de primer orden.

RESET test

$$RESET = 0.33968, df1 = 2, df2 = 29, p\text{-value} = 0.7148$$

El resultado de p-value obtenido en la prueba de RESET indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

13.4.3.1 Modelaciones metodología paso a paso.

A partir de este primer resultado del modelo, con el software R se realizó el análisis paso a paso, con la metodología atrás/adelante y utilizando el criterio de Akaike, con el fin de determinar las variables que mejor explicarían el modelo, y cuyos resultados se presentan a continuación:

Tabla 40. Modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras), metodología paso a paso

Residuals:

Min	1Q	Median	3Q	Max
-59167221	-18586140	-2361063	18889713	76034525

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	123165644	42376141	2.906	0.00607 **
Aconstruida	1348401	242079	5.570	0.00000221 ***
Carros	16759180	9574104	1.750	0.08811 .
Estrato	30450056	7455610	4.084	0.00022 ***
FittoCorvini	-24105809	11743636	-2.053	0.04704 *
ZonaComun	20871685	6757838	3.089	0.00375 **

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

Residual standard error: 31080000 on 38 degrees of freedom

Multiple R-squared: 0.7478, Adjusted R-squared: 0.7146

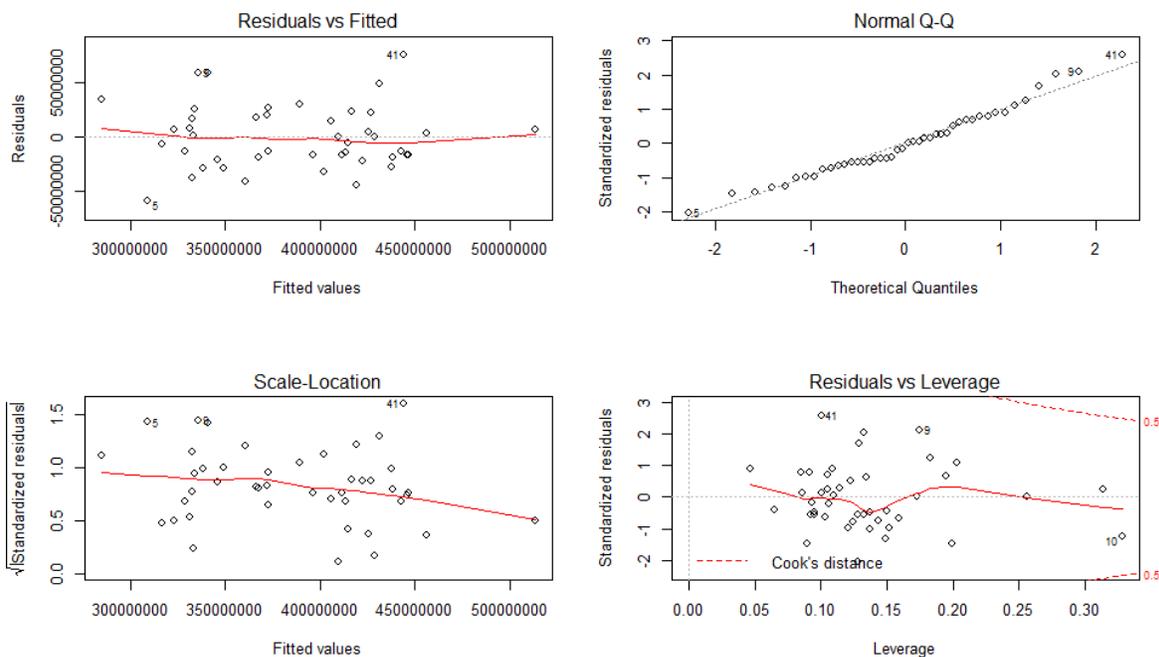
F-statistic: 22.54 on 5 and 38 DF, p-value: 1.95e-10

En los resultados obtenidos con la modelación en R, se observa que las variables área construida, cantidad de carros, estrato, estado de conservación y la zona común siguen siendo las variables más significativas estadísticamente, a pesar de haber introducido la variable distancia Euclidiana (lineal), dados los resultados obtenidos de probabilidades (p-value). El coeficiente de determinación R2 ajustado del modelo indica en forma porcentual que el 71.46% del valor del inmueble, es explicado por las variables independientes analizadas, con un p-crítico del modelo de 1.95E-10, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es mayor a 10 y por lo tanto el modelo tiene una muy buena capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

Gráfico 34. Comportamiento de los residuos del modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras), metodología paso a paso

lm(Valor.del.negocio ~ Aconstruida + Carros + Estrato + FittoCorvini + Zona ...



En la gráfica de los residuos vs los residuos ajustados, se observa que los residuos siguen patrones lineales, y están distribuidos uniformemente a lado y lado de la línea, de lo cual se puede deducir visualmente que cumplen con el criterio de la regresión lineal. En la gráfica normal Q-Q se puede observar que los residuos se distribuyen normalmente al encontrarse en general bien alineados. En el diagrama de scale-location se observa la homocedasticidad de los residuos. En la gráfica de Residuals vs Leverage, que permite encontrar casos influyentes, o que no se ajustan bien a la tendencia de la mayoría de los casos, no se observan casos que se encuentren fuera de la distancia de Cook.

A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos:

Inflación de la varianza

Aconstruida Carros Estrato FittoCorvini ZonaComun

1.198953 1.320091 1.700457 1.183740 1.808655

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

BP = 0.69269, df = 1, p-value = 0.4053

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por ende se asume la homocedasticidad del modelo.

Durbin-Watson test

DW = 1.8663, p-value = 0.2603

De acuerdo con el resultado de la prueba de Durbin – Watson, el estadístico es cercano a 2, y dado que el p-value de la prueba es mayor a 0.05 no se rechaza la hipótesis nula de no autocorrelación. Por consiguiente, se concluye que no hay autocorrelación serial de primer orden.

RESET test

RESET = 0.14557, df1 = 2, df2 = 36, p-value = 0.865

El resultado de p-value obtenido en la prueba de RESET indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

13.4.3.2 Modelaciones incluyendo la variable distancia lineal Euclidiana.

Como el objetivo de la investigación es determinar la incidencia de la variable distancia, se introduce la variable distancia lineal Euclidiana en el modelo obtenido anteriormente, logrando los siguientes resultados:

Tabla 41. Modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia euclidiana

Residuals:				
Min	1Q	Median	3Q	Max
-59134555	-18604533	-2321095	18886852	76043146
Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	122822938.2	49037393.6	2.505	0.016790 *
Aconstruida	1348683.9	246102.3	5.480	0.00000318 ***
Carros	16759758.6	9702676.5	1.727	0.092448 .
Distancia.Lineal.m	315.4	21786.8	0.014	0.988527
Estrato	30474372.6	7740129.6	3.937	0.000351 ***
FittoCorvini	-24075505.9	12083917.1	-1.992	0.053749 .
ZonaComun	20839032.6	7210390.4	2.890	0.006407 **

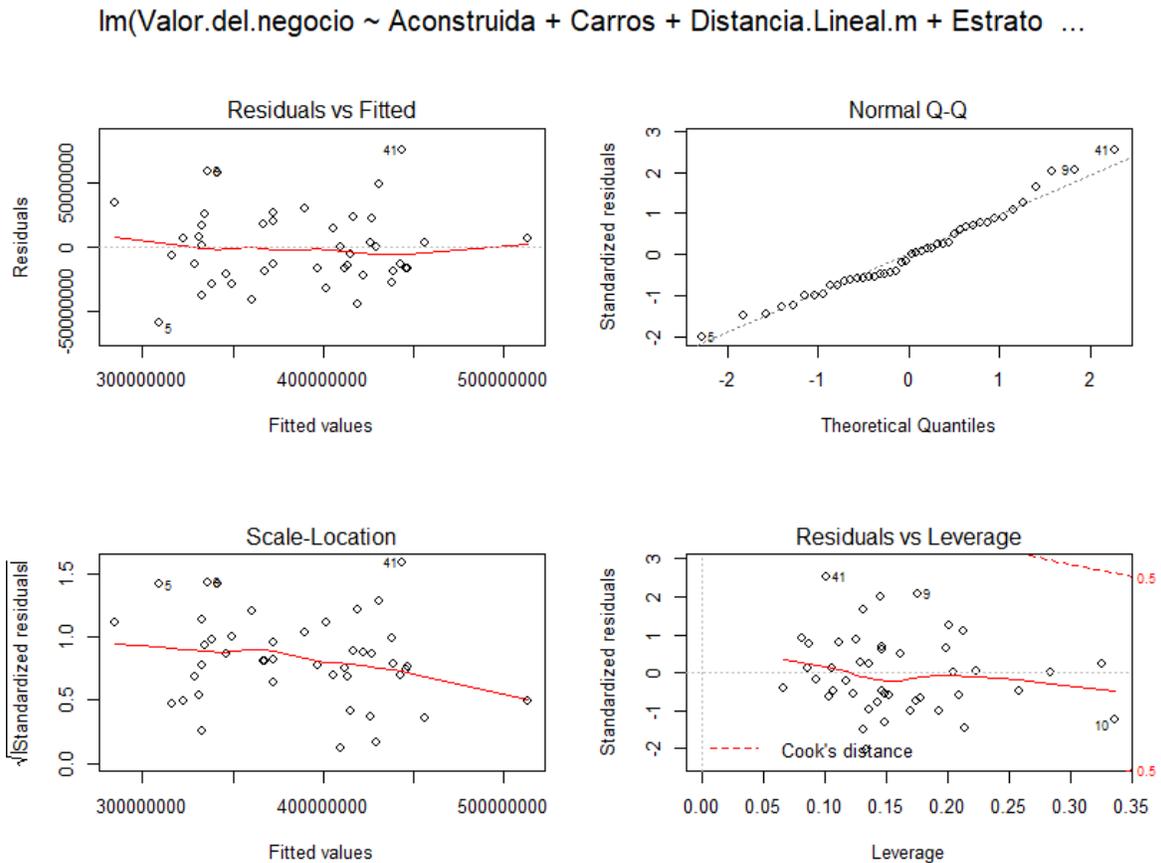
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

Residual standard error: 31490000 on 37 degrees of freedom
 Multiple R-squared: 0.7478, Adjusted R-squared: 0.7069
 F-statistic: 18.29 on 6 and 37 DF, p-value: 9.882e-10

En los resultados obtenidos con la modelación en R, se observa que las variables área construida, cantidad de carros, estrato, estado de conservación y la zona común siguen siendo las variables más significativas estadísticamente, a pesar de haber introducido la variable distancia en carro, dado los resultados obtenidos de probabilidades (p-value). El coeficiente de determinación R² ajustado del modelo indica en forma porcentual que el 70.69% del valor del inmueble, es explicado por las variables independientes analizadas, con un p-crítico del modelo de 9.88E-10, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es levemente mayor a 18 y por lo tanto el modelo tiene una aceptable capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

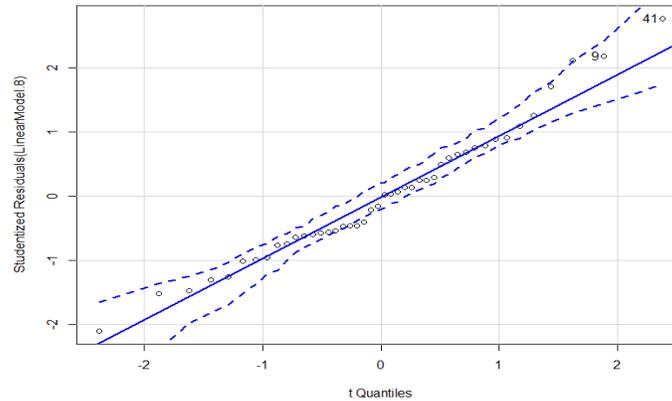
Gráfico 35. Comportamiento de los residuos del modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia euclidiana



De acuerdo con el resultado observado en el gráfico Residuals vs Fitted se puede decir que el modelo que cumple con los supuestos de linealidad.

Así mismo, se observa en el gráfico Normal Q-Q la normalidad de los residuos, cumpliéndose de igual forma con el supuesto del modelo. En el gráfico de cuantiles se observa que los residuos studentizados se encuentran dentro de los intervalos, lo cual indica que siguen una distribución normal.

Gráfico 36. Cuantiles del modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia euclidiana

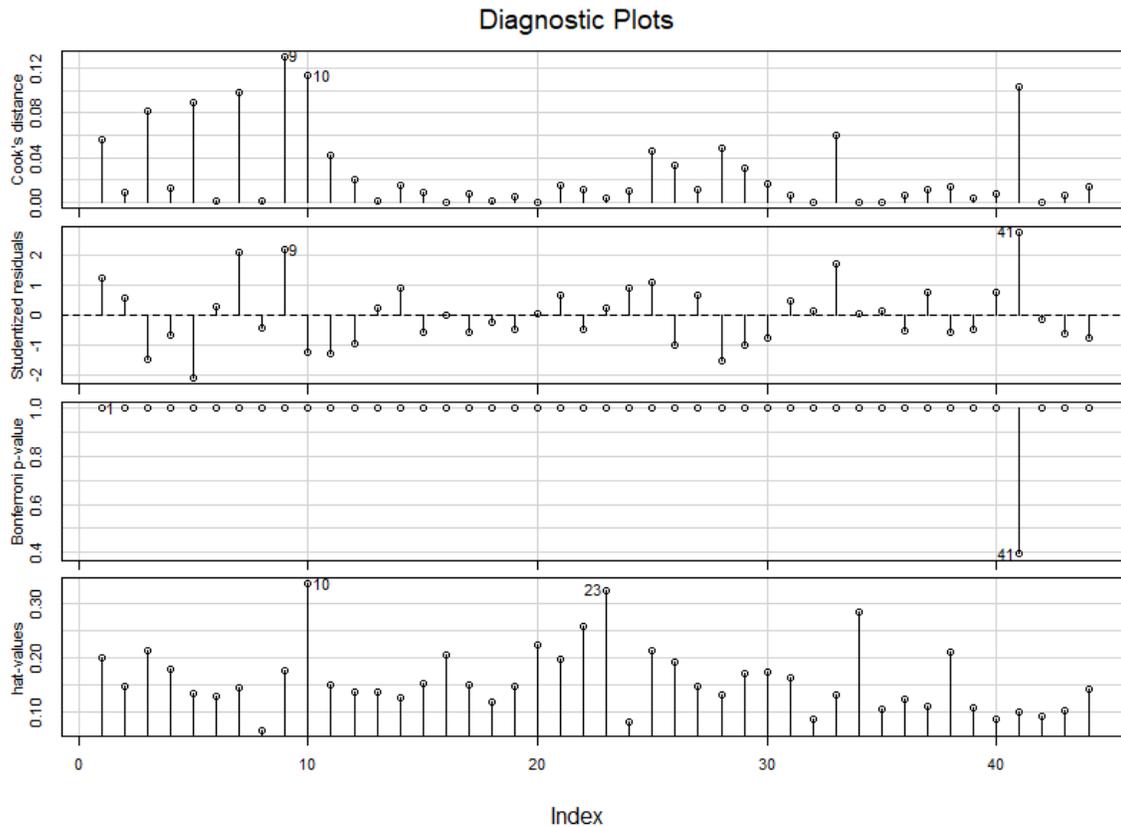


En la gráfica Scale-Location se observa que los residuos están distribuidos equitativamente a lo largo de los rangos de los predictores con lo cual se puede comprobar el supuesto de la varianza igual u homocedasticidad.

Finalmente, en la gráfica Residuals vs Leverage la cual ayuda a identificar los casos que son más influyentes en la regresión lineal (sean atípicos o no). En todos los casos se observa que los datos se encuentran dentro de la distancia Cook, aun cuando el programa muestra que hay unos datos que si se eliminaran podrían mejorar los resultados del coeficiente de determinación.

En la gráfica de índices de influencia a continuación, se observa la influencia de cada una de las muestras del estudio utilizadas para el modelo:

Gráfico 37. Índices de influencia del modelo de regresión lineal con distancia euclidiana excluyendo outliers (44 muestras), incluyendo al mejor modelo la distancia euclidiana



A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos:

Inflación de la varianza

Aconstruida	Carros	Distancia.Lineal.m	Estrato
1.206538	1.320113	1.162042	1.784500
	FittoCorvini	ZonaComun	
	1.220358	2.004833	

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 0.70038, df = 1, p\text{-value} = 0.4027$$

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por ende se asume la homocedasticidad.

Durbin-Watson test

$$DW = 1.8643, p\text{-value} = 0.2475$$

De acuerdo con el resultado de la prueba de Durbin – Watson, el estadístico es cercano a 2, y dado que el p-value de la prueba es mayor a 0.05 no se rechaza la hipótesis nula de no autocorrelación. Por consiguiente, se concluye que no hay autocorrelación serial de primer orden.

RESET test

$$RESET = 0.15465, df1 = 2, df2 = 35, p\text{-value} = 0.8573$$

El resultado de p-value obtenido en la prueba de RESET indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

13.5 Modelaciones con totalidad de los datos

En estas modelaciones tuvieron en cuenta la totalidad de las muestras sin excluir datos atípicos, ya que si bien son considerados como tal, puede que no tengan influencia en el resultado, como si lo podrían tener los datos influyentes detectados en los análisis. Por lo tanto, la cantidad de muestras inicial empleadas para la elaboración del modelo fueron 53.

13.5.1 Modelaciones con distancia a pie.

Partimos de la aplicación del modelo lineal con el software estadístico R, para todas las variables preseleccionadas, incluyendo en la evaluación la variable distancia a pie:

Tabla 42. Modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), incluyendo todas las variables

Coefficients:	Estimate	Std. Error	t	value Pr(> t)	
(Intercept)	-59234205	77384847	-0.765	0.4485	
Aconstruida	2550352	361583	7.053	0.0000000158	***
Ascensor	36571168	24055527	1.520	0.1363	
Carros	39387608	14921666	2.640	0.0118	*
CuartoUtil	9425461	20288743	0.465	0.6448	
Distancia.a.Pie..m.	31294	28293	1.106	0.2753	
Estrato	39002568	14182452	2.750	0.0089	**
FittoCorvini	-22806792	17820023	-1.280	0.2080	
Habitaciones	-11122218	15760520	-0.706	0.4845	
NivelPiso	983442	2591957	0.379	0.7064	
Poniente	-6819915	17909509	-0.381	0.7054	
Porteria	-8954660	13846267	-0.647	0.5215	
ZonaComun	8977483	11986141	0.749	0.4582	

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

Residual standard error: 48820000 on 40 degrees of freedom

Multiple R-squared: 0.8552

Adjusted R-squared: 0.8118

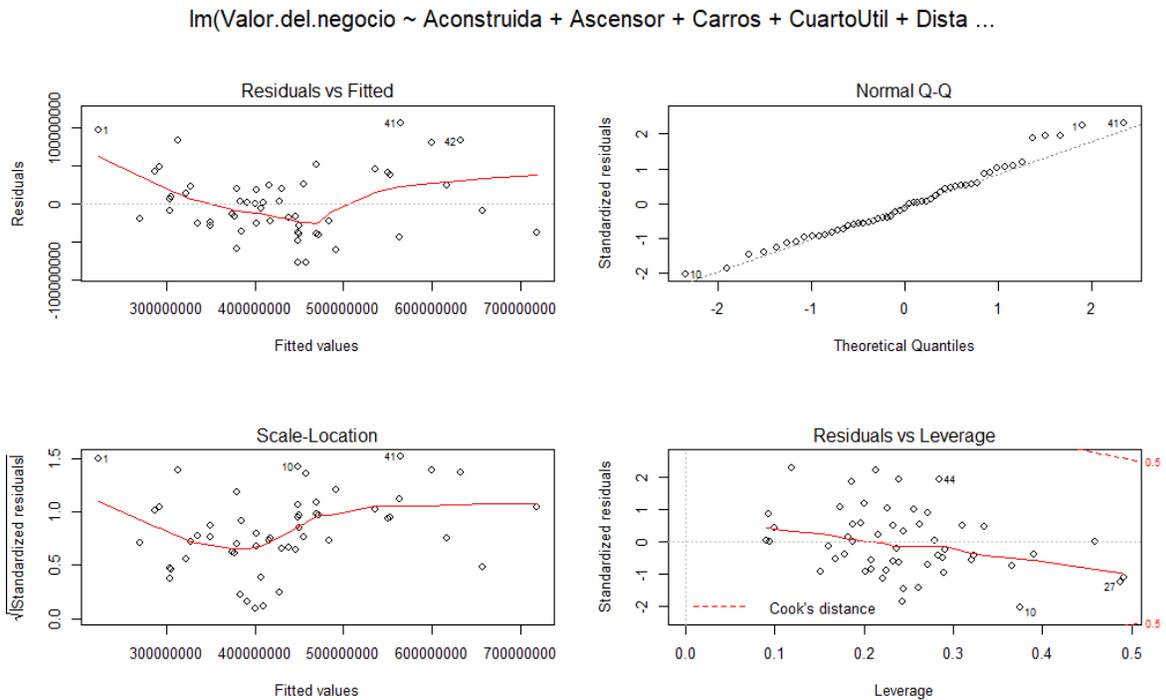
F-statistic: 19.69

p-value: 4.132e-13

En los resultados obtenidos con la modelación en R, se observa que las variables área construida, cantidad de carros, estrato son las variables más significativas estadísticamente, a pesar de haber introducido la variable distancia Euclidiana (lineal), dados los resultados obtenidos de probabilidades (p-value). El coeficiente de determinación R² ajustado del modelo indica en forma porcentual que el 81.18% del valor del inmueble es explicado por las variables independientes analizadas, con un p-crítico del modelo de 4.132E-13, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es mayor a 10 y por lo tanto el modelo tiene una aceptable capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

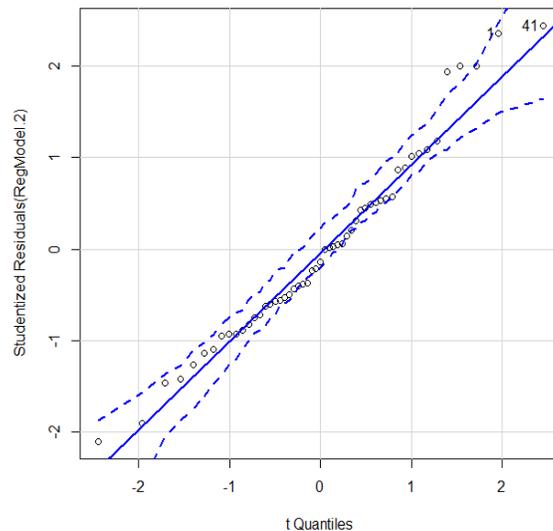
Gráfico 38. Comportamiento de los residuos del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), incluyendo todas las variables



De acuerdo con el resultado observado en el gráfico Residuals vs Fitted se puede decir que visualmente el modelo que cumple con los supuestos de linealidad, pero se debe confirmar a través de las pruebas de hipótesis.

Así mismo, se observa en el gráfico Normal Q-Q la normalidad de los residuos, cumpliéndose en general con el supuesto del modelo. En el gráfico de cuantiles se observa que los residuos studentizados se encuentran en general dentro de los intervalos, con algunos puntos por fuera lo cual indica que en general siguen una distribución normal.

Gráfico 39. Cuantiles del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), incluyendo todas las variables

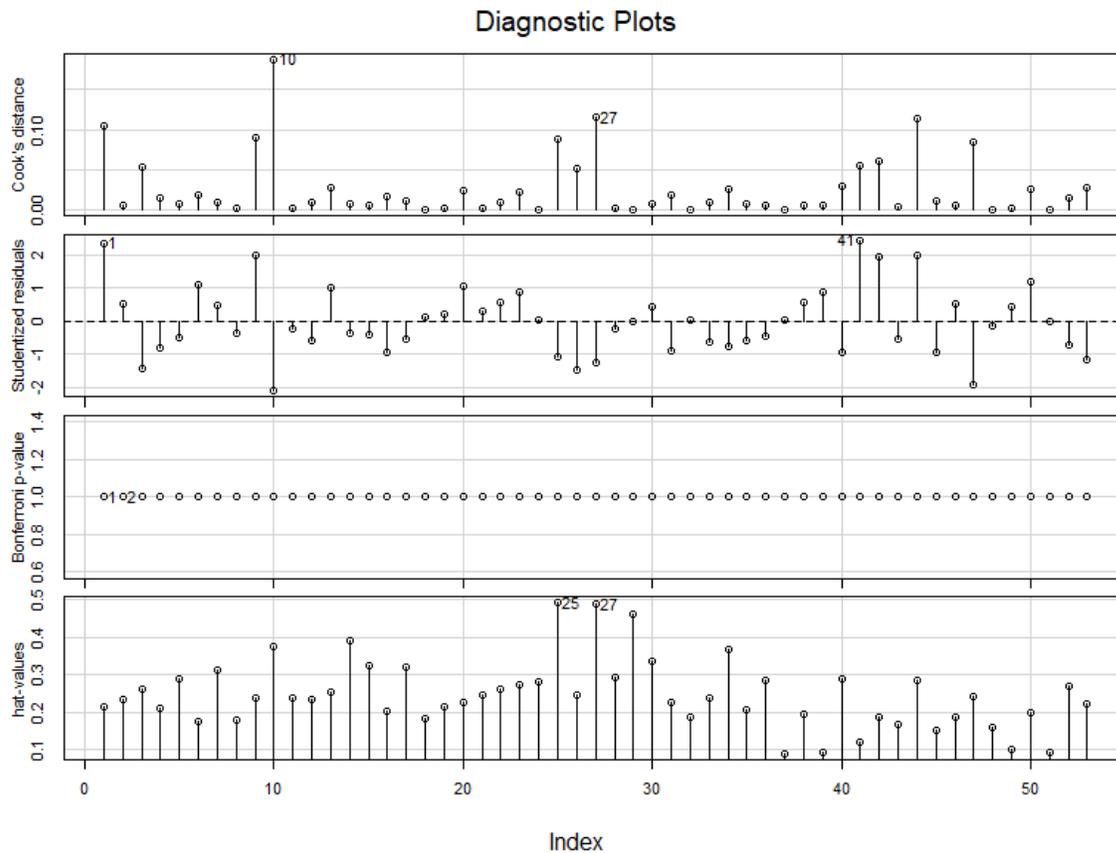


En la gráfica Scale-Location se observa que los residuos están distribuidos equitativamente a lo largo de los rangos de los predictores con lo cual se puede comprobar el supuesto de la varianza igual u homocedasticidad.

Finalmente, en la gráfica Residuals vs Leverage la cual ayuda a identificar los casos que son más influyentes en la regresión lineal (sean atípicos o no). En todos los casos se observa que los datos se encuentran dentro de la distancia Cook, aun cuando el programa muestra que hay unos datos que si se eliminaran podrían mejorar los resultados del coeficiente de determinación.

En la gráfica de índices de influencia a continuación, se observa la influencia de cada una de las muestras del estudio utilizadas para el modelo:

Gráfico 40. Índices de influencia del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), incluyendo todas las variables



A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos:

Inflación de la varianza

Aconstruida	Ascensor	Carros	CuartoUtil
2.457601	2.803406	1.935271	2.196215
Distancia.a.Pie..m.	Estrato	FittoCorvini	Habitaciones
1.364779	2.786413	1.475569	2.843244
NivelPiso	Poniente	Porteria	ZonaComun
1.727865	1.249215	4.085500	2.740820

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 1.4297, df = 1, p\text{-value} = 0.2318$$

Durbin-Watson test

$$DW = 1.8291, p\text{-value} = 0.18$$

De acuerdo con el resultado de la prueba de Durbin – Watson, el estadístico es cercano a 2, y dado que el p-value de la prueba es mayor a 0.05 no se rechaza la hipótesis nula de no autocorrelación. Por consiguiente, se concluye que no hay autocorrelación serial de primer orden.

RESET test

$$RESET = 10.483, df1 = 2, df2 = 38, p\text{-value} = 0.0002368$$

El resultado de p-value obtenido en la prueba de RESET indica que la forma funcional (linealidad) no es correcta, de tal manera que hay un problema de especificación.

De acuerdo con el resultado obtenido en la prueba de hipótesis de linealidad, se realizó un análisis de influencia de los datos y de datos atípicos a través de la prueba de Bonferroni en el software R, con el fin de determinar las muestras que están influyendo en mayor magnitud en el resultado, obteniendo que la muestra 53 ubicada en las fila 41 (como dato atípico), la muestra 2 (localizada en la fila 1), la muestra 16 (localizada en la fila 10) y la muestra 56 (localizada en la fila 44) son las que tienen mayor influencia.

Al realizar nuevamente la modelación sin dichas muestras obtenemos el siguiente resultado de modelación con 49 datos:

Tabla 43. Modelo de regresión lineal con distancia a pie (49 muestras luego de aplicar Bonferroni), incluyendo todas las variables

Residuals:				
Min	1Q	Median	3Q	Max
-64026714	-26150489	-3084342	25904732	71212749
Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-63017120	62280373	-1.012	0.31856
Aconstruida	2275447	289335	7.864	0.0000000304 ***
Ascensor	40779810	18688722	2.182	0.03591 *
Carros	49203114	12848820	3.829	0.00051 ***
CuartoUtil	-4739439	17004285	-0.279	0.78210
Distancia.a.Pie..m.	27899	22422	1.244	0.22167
Estrato	36758989	11373349	3.232	0.00268 **
FittoCorvini	-17418462	15073929	-1.156	0.25570
Habitaciones	-1381212	12484797	-0.111	0.91254
NivelPiso	768052	2065061	0.372	0.71219
Poniente	8362145	14531402	0.575	0.56867
Porteria	-12807196	10968544	-1.168	0.25085
ZonaComun	8388043	9192708	0.912	0.36777

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

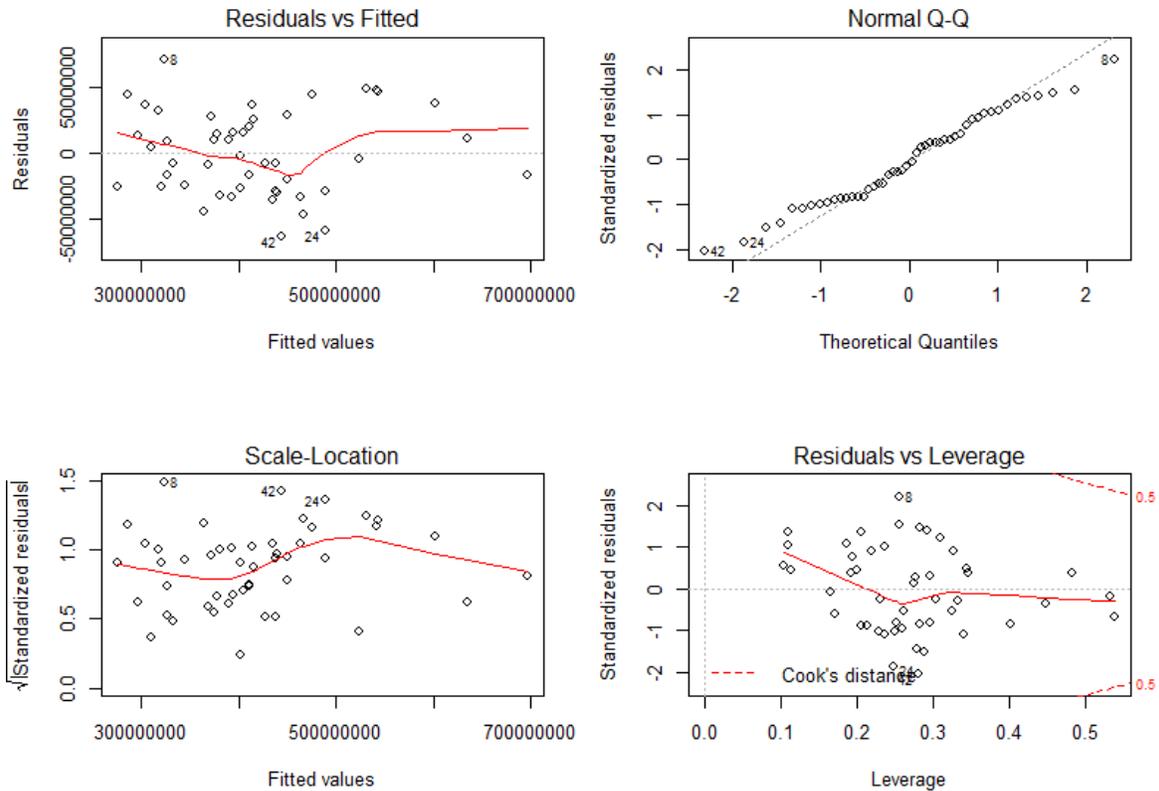
Residual standard error: 37080000 on 35 degrees of freedom
 Multiple R-squared: 0.8891, Adjusted R-squared: 0.8511
 F-statistic: 23.39 on 12 and 35 DF, p-value: 3.309e-13

Una vez realizados los cambios, en los resultados obtenidos con la modelación en R, se observa que las variables área construida, ascensor, cantidad de carros y estrato son las variables más significativas estadísticamente, dados los resultados obtenidos de probabilidades (p-value). El coeficiente de determinación R² ajustado del modelo indica en forma porcentual que el 85.11% del valor del inmueble es explicado por las variables independientes analizadas, con un p-crítico del modelo de 3.309E-13, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es mayor a 10 y por lo tanto el modelo tiene una aceptable capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

Gráfico 41. Comportamiento de los residuos del modelo de regresión lineal con distancia a pie (49 muestras luego de aplicar Bonferroni), incluyendo todas las variables

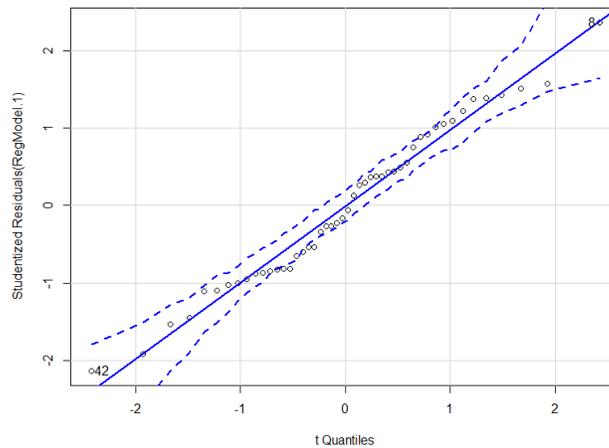
lm(Valor.del.negocio ~ Aconstruida + Ascensor + Carros + CuartoUtil + Dista ...



De acuerdo con el resultado observado en el gráfico Residuals vs Fitted, mejoró el resultado obtenido previamente, y se puede decir que el modelo que cumple con los supuestos de linealidad.

Así mismo, se observa en el gráfico Normal Q-Q la normalidad de los residuos, cumpliéndose de igual forma con el supuesto del modelo. En el gráfico de cuantiles se observa que los residuos studentizados se encuentran dentro de los intervalos, lo cual indica que siguen una distribución normal.

Gráfico 42. Cuantiles del modelo de regresión lineal con distancia a pie (49 muestras luego de aplicar Bonferroni), incluyendo todas las variables

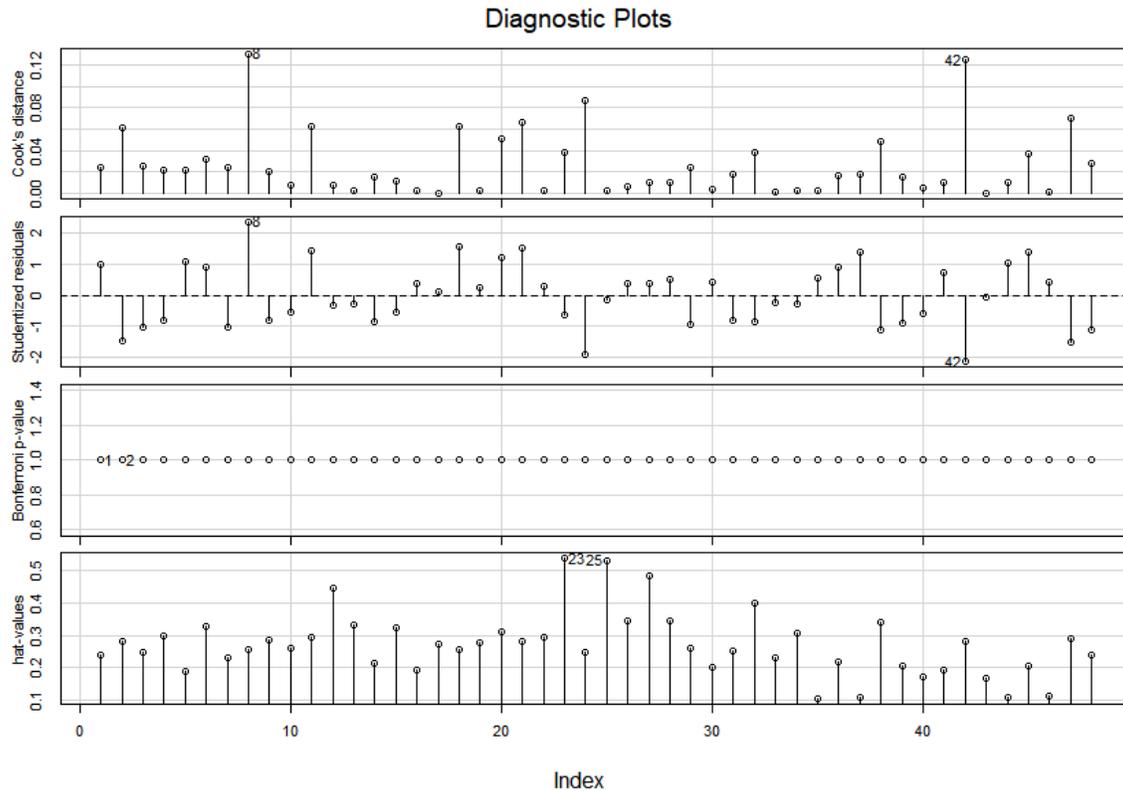


En la gráfica Scale-Location se observa que los residuos están distribuidos equitativamente a lo largo de los rangos de los predictores con lo cual se puede comprobar el supuesto de la varianza igual u homocedasticidad.

Finalmente, en la gráfica Residuals vs Leverage la cual ayuda a identificar los casos que son más influyentes en la regresión lineal (sean atípicos o no). En todos los casos se observa que los datos se encuentran dentro de la distancia Cook, aun cuando el programa muestra que hay unos datos que si se eliminaran podrían mejorar los resultados del coeficiente de determinación.

En la gráfica de índices de influencia a continuación, se observa la influencia de cada una de las muestras del estudio utilizadas para el modelo:

Gráfico 43. Índices de influencia de regresión lineal con distancia a pie (49 muestras luego de aplicar Bonferroni), incluyendo todas las variables



A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos:

Shapiro-Wilk normality test (Normalidad de los residuos)

$$W = 0.97672, p\text{-value} = 0.4505$$

En este caso, para verificar la normalidad de los residuos, una vez fueron excluidas las muestras influyentes se aplicó la prueba de Shapiro – Wilk a los mismos. El resultado de p-value de la prueba mayor a 0.05 confirma la normalidad de los residuos.

Inflación de la varianza

Aconstruida	Ascensor	Carros	CuartoUtil
2.288233	2.619455	2.001073	2.085308
Distancia.a.Pie..m	Estrato	FittoCorvini	Habitaciones
1.402352	2.626201	1.528560	2.869365
NivelPiso	Poniente	Porteria	ZonaComun
1.879366	1.302135	4.022970	2.622185

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 0.33124, df = 1, p\text{-value} = 0.5649$$

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por ende se asume la homocedasticidad.

Durbin-Watson test

$$DW = 1.8413, p\text{-value} = 0.1995$$

De acuerdo con el resultado de la prueba de Durbin – Watson, el estadístico es cercano a 2, y dado que el p-value de la prueba es mayor a 0.05 no se rechaza la hipótesis nula de no autocorrelación. Por consiguiente, se concluye que no hay autocorrelación serial de primer orden.

RESET test

$$RESET = 2.9562, df1 = 2, df2 = 33, p\text{-value} = 0.06592$$

El resultado de p-value obtenido en la prueba de RESET indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

13.5.1.1 Modelación metodología Paso a Paso.

A partir de este resultado del modelo con todas las variables, con el software R se realizó el análisis paso a paso, con la metodología atrás/adelante y utilizando el criterio de Akaike, con el fin de determinar las variables que mejor explicarían el modelo, y cuyos resultados se presentan a continuación:

Tabla 44. Modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), metodología paso a paso

Residuals:				
Min	1Q	Median	3Q	Max
-59509046	-25028336	-2320333	21746977	83186079
Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-11156363	38098598	-0.293	0.77109
Aconstruida	2203784	197859	11.138	4.08e-14 ***
Ascensor	32531986	14665066	2.218	0.03199 *
Carros	50160358	10650788	4.710	2.72e-05 ***
Estrato	31931861	9091157	3.512	0.00108 **
FittoCorvini	-20632382	12813914	-1.610	0.11486

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

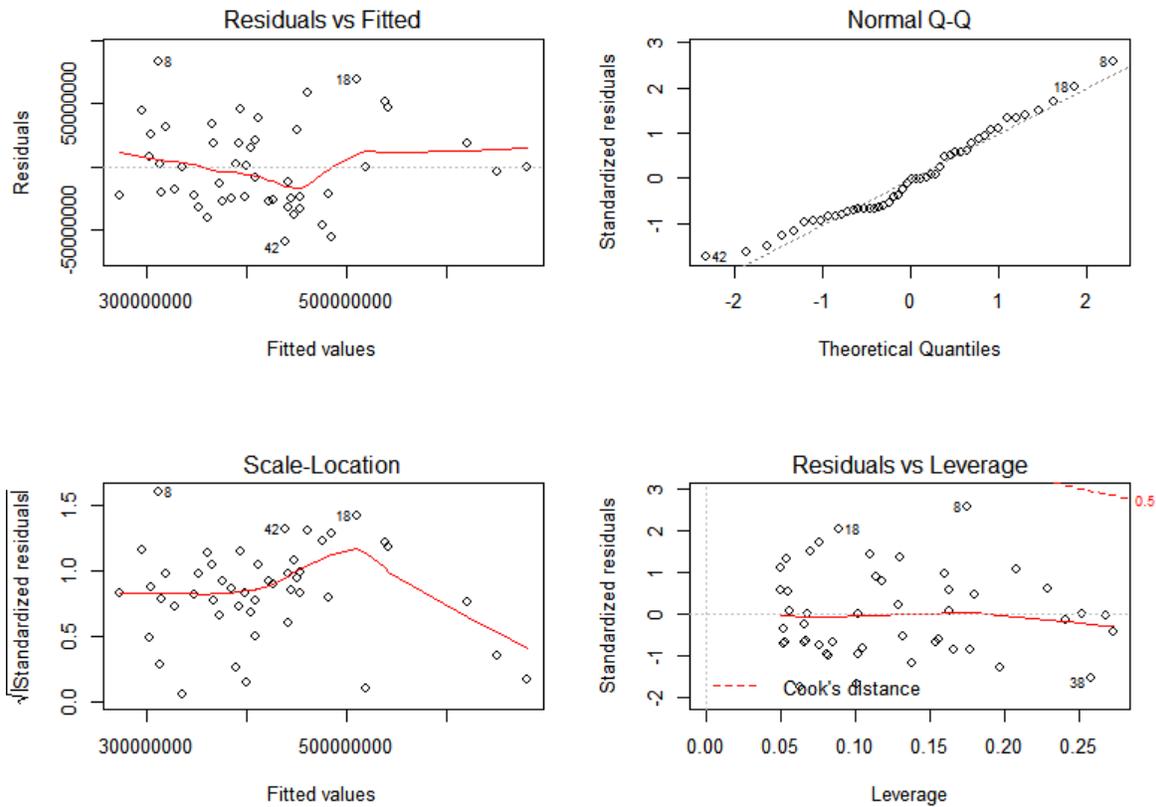
Residual standard error: 35520000 on 42 degrees of freedom
 Multiple R-squared: 0.8779, Adjusted R-squared: 0.8634
 F-statistic: 60.41 on 5 and 42 DF, p-value: < 2.2e-16

Una vez realizados los cambios, en los resultados obtenidos con la modelación en R, se observa que las variables área construida, ascensor, cantidad de carros y estrato son las variables más significativas estadísticamente, dados los resultados obtenidos de probabilidades (p-value). El coeficiente de determinación R2 ajustado del modelo indica en forma porcentual que el 86.34%% del valor del inmueble es explicado por las variables independientes analizadas, con un p-crítico del modelo de 2.2E-16, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es mucho mayor a 10, y por lo tanto, el modelo tiene una muy buena capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

Gráfico 44. Comportamiento de los residuos del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), metodología paso a paso

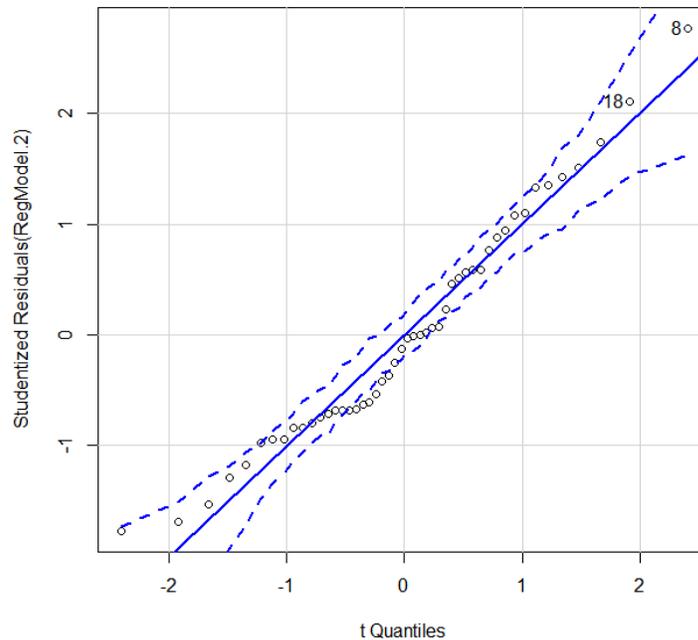
lm(Valor.del.negocio ~ Aconstruida + Ascensor + Carros + Estrato + FittoCor ...



De acuerdo con el resultado observado en el gráfico Residuals vs Fitted, mejoró el resultado obtenido previamente, y se puede decir que el modelo que cumple con los supuestos de linealidad.

Así mismo, se observa en el gráfico Normal Q-Q la normalidad de los residuos, cumpliéndose de igual forma con el supuesto del modelo. En el gráfico de cuantiles se observa que los residuos studentizados se encuentran en general dentro de los intervalos, lo cual indica que siguen una distribución normal.

Gráfico 45. Cuantiles del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), metodología paso a paso

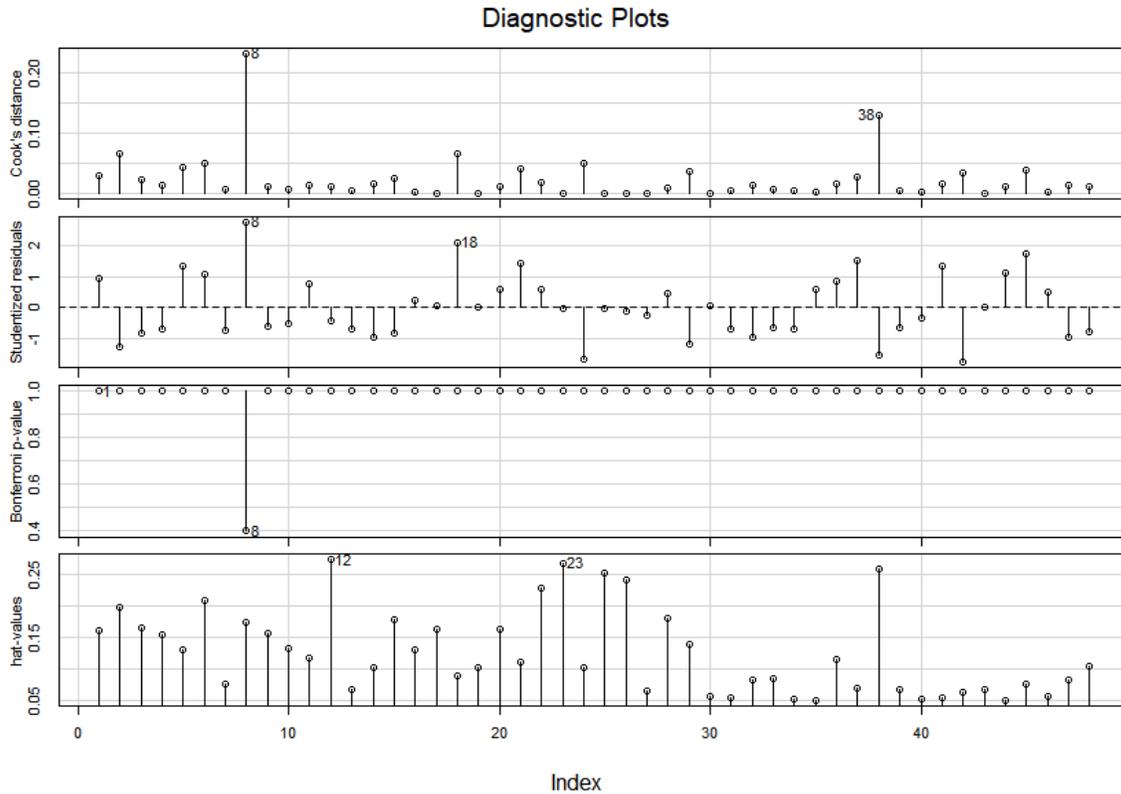


En la gráfica Scale-Location se observa que los residuos están distribuidos equitativamente a lo largo de los rangos de los predictores con lo cual se puede comprobar el supuesto de la varianza igual u homocedasticidad.

Finalmente, en la gráfica Residuals vs Leverage la cual ayuda a identificar los casos que son más influyentes en la regresión lineal (sean atípicos o no). En todos los casos se observa que los datos se encuentran dentro de la distancia Cook, aun cuando el programa muestra que hay unos datos que si se eliminaran podrían mejorar los resultados del coeficiente de determinación.

En la gráfica de índices de influencia a continuación, se observa la influencia de cada una de las muestras del estudio utilizadas para el modelo:

Gráfico 46. Índices de influencia del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), metodología paso a paso



A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos:

Inflación de la varianza

Aconstruida	Ascensor	Carros	Estrato	FittoCorvini
1.166516	1.758334	1.498931	1.829242	1.204134

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 0.043028, df = 1, p\text{-value} = 0.8357$$

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por ende se asume la homocedasticidad.

Durbin-Watson test

$$DW = 1.9071, p\text{-value} = 0.3137$$

De acuerdo con el resultado de la prueba de Durbin – Watson, el estadístico es cercano a 2, y dado que el p-value de la prueba es mayor a 0.05 no se rechaza la hipótesis nula de no autocorrelación. Por consiguiente, se concluye que no hay autocorrelación serial de primer orden.

RESET test

$$RESET = 2.0446, df1 = 2, df2 = 40, p\text{-value} = 0.1427$$

El resultado de p-value obtenido en la prueba de RESET indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

Shapiro-Wilk normality test

$$W = 0.96529, p\text{-value} = 0.165$$

El resultado de p-value de la prueba de Shapiro – Wilk mayor a 0.05 confirma la normalidad de los residuos.

13.5.1.2 Modelación incluyendo variable distancia a pie.

Como el objetivo de la investigación es la incidencia de la variable distancia, se introduce la variable distancia a pie en el modelo obtenido anteriormente, consiguiendo los siguientes resultados:

Tabla 45. Modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia a pie

Residuals:				
Min	1Q	Median	3Q	Max
-66181472	-24578973	-4791262	23597841	82734139
Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-31890285	43591736	-0.732	0.468594
Aconstruida	2219611	198608	11.176	5.09e-14 ***
Ascensor	32762549	14673781	2.233	0.031089 *
Carros	49117014	10708773	4.587	4.19e-05 ***
Distancia.a.Pie..m.	18122	18487	0.980	0.332707
Estrato	33235341	9192079	3.616	0.000812 ***
FittoCorvini	-20853125	12821859	-1.626	0.111533

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

Residual standard error: 35530000 on 41 degrees of freedom

Multiple R-squared: 0.8807, Adjusted R-squared: 0.8633

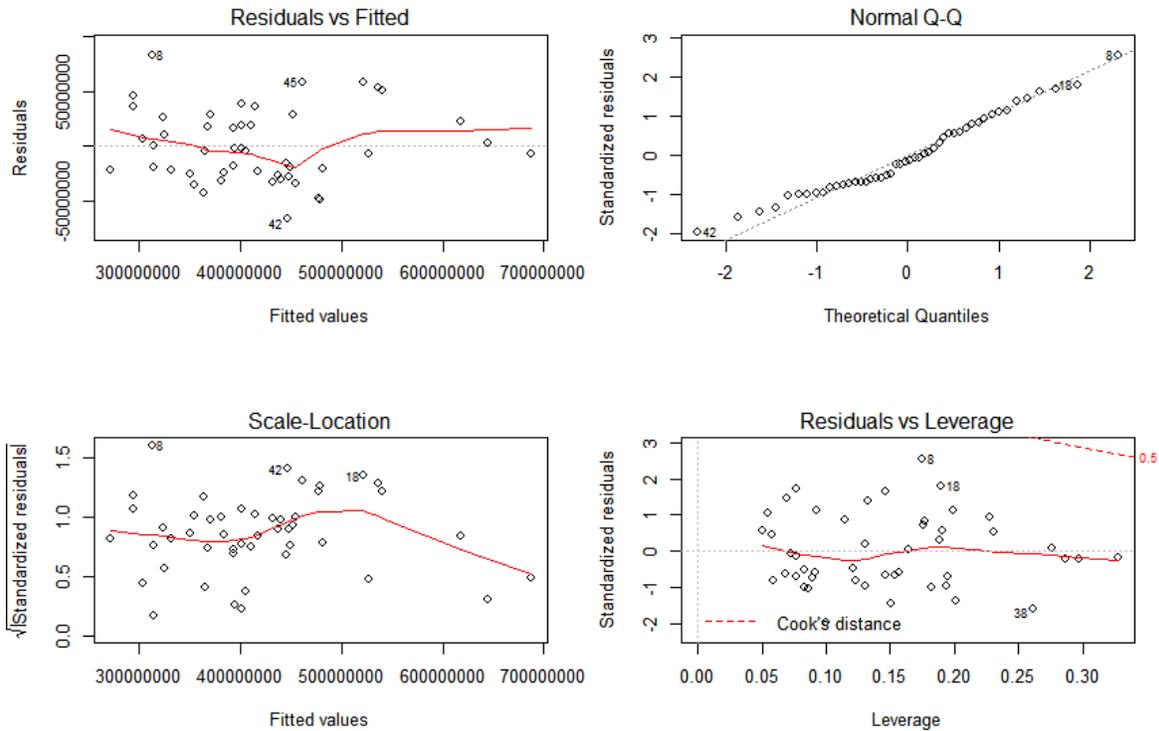
F-statistic: 50.46 on 6 and 41 DF, p-value: < 2.2e-16

Una vez realizados los cambios, en los resultados obtenidos con la modelación en R, se observa que las variables área construida, ascensor, cantidad de carros y estrato son las variables más significativas estadísticamente, dados los resultados obtenidos de probabilidades (p-value). El coeficiente de determinación R² ajustado del modelo indica en forma porcentual que el 86.34% del valor del inmueble es explicado por las variables independientes analizadas, con un p-crítico del modelo de 2.2E-16, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es mucho mayor a 10, y por lo tanto, el modelo tiene una muy buena capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

Gráfico 47. Comportamiento de los residuos del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia a pie

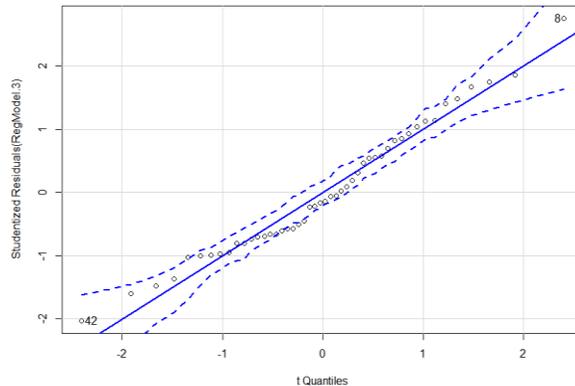
lm(Valor.del.negocio ~ Aconstruida + Ascensor + Carros + Distancia.a.Pie..m ...



De acuerdo con el resultado observado en el gráfico Residuals vs Fitted, se puede decir que visualmente el modelo que cumple con los supuestos de linealidad.

Así mismo, se observa en el gráfico Normal Q-Q la normalidad de los residuos, cumpliéndose de igual forma con el supuesto del modelo. En el gráfico de cuantiles se observa que los residuos studentizados se encuentran en general dentro de los intervalos, lo cual indica que siguen una distribución normal.

Gráfico 48. Cuantiles del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia a pie

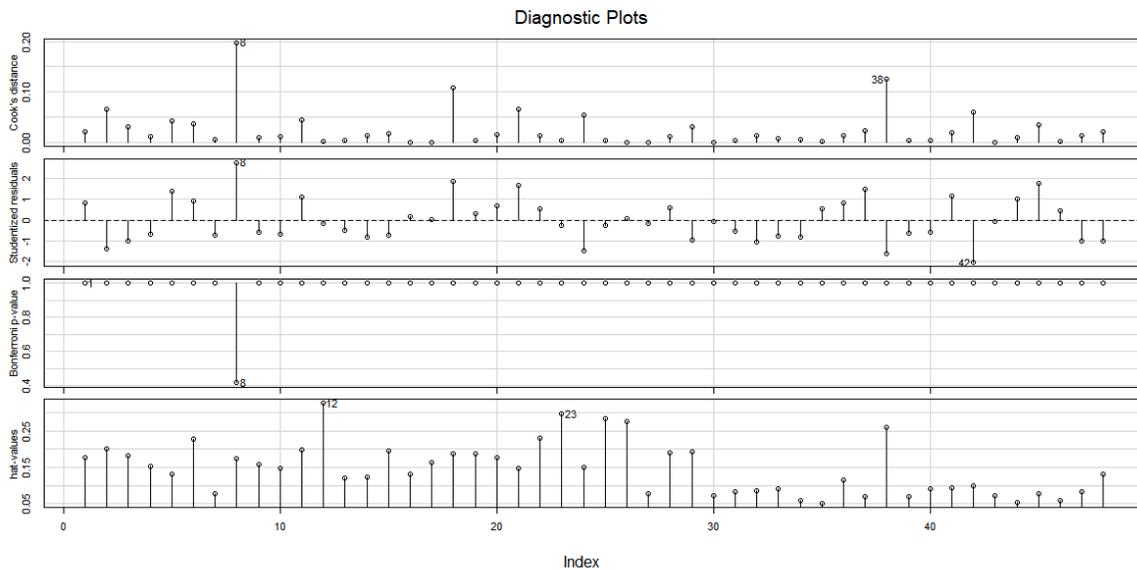


En la gráfica Scale-Location se observa que los residuos están distribuidos equitativamente a lo largo de los rangos de los predictores con lo cual se puede comprobar el supuesto de la varianza igual u homocedasticidad.

Finalmente, en la gráfica Residuals vs Leverage la cual ayuda a identificar los casos que son más influyentes en la regresión lineal (sean atípicos o no). En todos los casos se observa que los datos se encuentran dentro de la distancia Cook, aun cuando el programa muestra que hay unos datos que si se eliminaran podrían mejorar los resultados del coeficiente de determinación.

En la gráfica de índices de influencia a continuación, se observa la influencia de cada una de las muestras del estudio utilizadas para el modelo:

Gráfico 49. Índices de influencia del modelo de regresión lineal con distancia a pie incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia a pie



A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos:

Inflación de la varianza

Aconstruida	Ascensor	Carros	Distancia.a.Pie..m.
1.174276	1.758786	1.513886	1.038261
	Estrato	FittoCorvini	
	1.868340	1.204505	

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 0.019487, df = 1, p\text{-value} = 0.889$$

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por consiguiente se asume la homocedasticidad.

Durbin-Watson test

$$DW = 1.8375, p\text{-value} = 0.223$$

De acuerdo con el resultado de la prueba de Durbin – Watson, el estadístico es cercano a 2, y dado que el p-value de la prueba es mayor a 0.05 no se rechaza la hipótesis nula de no autocorrelación. Por consiguiente, se concluye que no hay autocorrelación serial de primer orden.

RESET test

$$RESET = 2.7425, df1 = 2, df2 = 39, p\text{-value} = 0.07684$$

El resultado de p-value obtenido en la prueba de RESET indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

13.5.2 Modelaciones con distancia lineal.

Partimos de la aplicación del modelo lineal con el software estadístico R, para todas las variables preseleccionadas, incluyendo en la evaluación la variable distancia a pie:

Tabla 46. Modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), incluyendo todas las variables

Residuals:				
Min	1Q	Median	3Q	Max
-60480621	-28367684	-2617606	29097997	73571759
Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-46832426	67300530	-0.696	0.491105
Aconstruida	2300841	307443	7.484	9.17e-09 ***
Ascensor	38928653	19095459	2.039	0.049097 *
Carros	49783008	13074703	3.808	0.000543 ***
CuartoUtil	-4497722	17249726	-0.261	0.795820
Distancia.Lineal.m	20793	29110	0.714	0.479793
Estrato	35679007	12071574	2.956	0.005553 **
FittoCorvini	-18581656	15336530	-1.212	0.233784
Habitaciones	-1963863	12826612	-0.153	0.879192
NivelPiso	651564	2091886	0.311	0.757290
Poniente	9006571	15105905	0.596	0.554859
Porteria	-11522315	11818354	-0.975	0.336275
ZonaComun	7956199	9475304	0.840	0.406788

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

Residual standard error: 37620000 on 35 degrees of freedom

Multiple R-squared: 0.8859, Adjusted R-squared: 0.8467

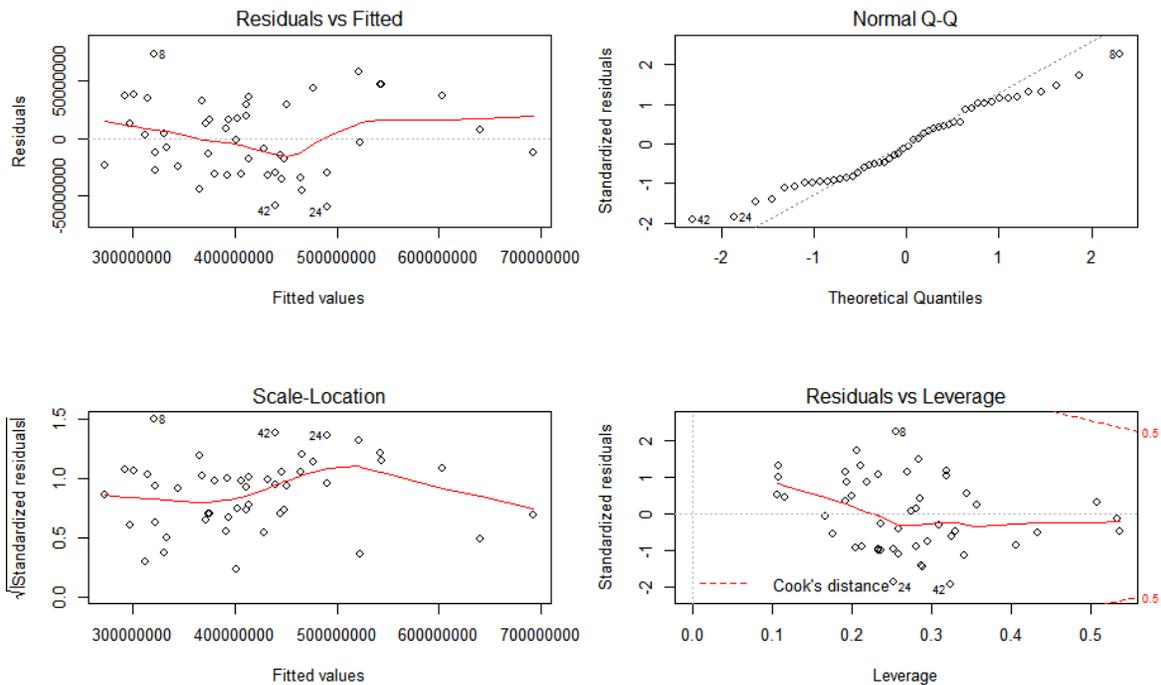
F-statistic: 22.64 on 12 and 35 DF, p-value: 5.386e-13

En los resultados obtenidos con la modelación en R, se observa que las variables área construida, ascensor, cantidad de carros y estrato son las variables más significativas estadísticamente, acorde con los resultados obtenidos de probabilidades (p-value). El coeficiente de determinación R² ajustado del modelo indica en forma porcentual que el 84.67% del valor del inmueble es explicado por las variables independientes analizadas, con un p-crítico del modelo de 5.38E-13, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es mayor a 10, y por lo tanto, el modelo tiene una muy buena capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

Gráfico 50. Comportamiento de los residuos del modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), incluyendo todas las variables

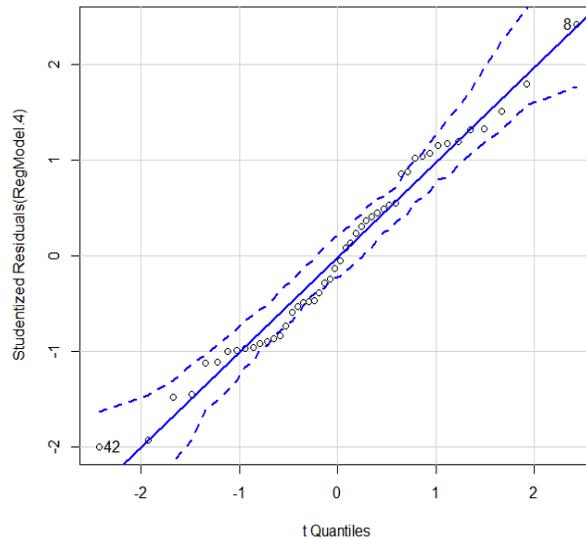
lm(Valor.del.negocio ~ Aconstruida + Ascensor + Carros + CuartoUtil + Dista ...



De acuerdo con el resultado observado en el gráfico Residuals vs Fitted, se puede decir que visualmente el modelo que cumple con los supuestos de linealidad.

Así mismo, se observa en el gráfico Normal Q-Q la normalidad de los residuos, cumpliéndose de igual forma con el supuesto del modelo. En el gráfico de cuantiles se observa que los residuos studentizados se encuentran en general dentro de los intervalos, lo cual indica que siguen una distribución normal.

Gráfico 51. Cuantiles del modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), incluyendo todas las variables



En la gráfica Scale-Location se observa que los residuos están distribuidos equitativamente a lo largo de los rangos de los predictores con lo cual se puede comprobar el supuesto de la varianza igual u homocedasticidad.

Finalmente, en la gráfica Residuals vs Leverage la cual ayuda a identificar los casos que son más influyentes en la regresión lineal (sean atípicos o no). En todos los casos se observa que los datos se encuentran dentro de la distancia Cook, aun cuando el programa muestra que hay unos datos que si se eliminaran podrían mejorar los resultados del coeficiente de determinación.

A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos:

Inflación de la varianza

Aconstruida	Ascensor	Carros	CuartoUtil
2.510233	2.657043	2.013199	2.084992
Distancia.Lineal.m	Estrato	FittoCorvini	Habitaciones
1.914777	2.874522	1.537341	2.942614
NivelPiso	Poniente	Porteria	ZonaComun
1.873735	1.367165	4.537843	2.706758

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 0.31387, df = 1, p\text{-value} = 0.5753$$

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por consiguiente se asume la homocedasticidad.

Durbin-Watson test

$$DW = 1.8915, p\text{-value} = 0.2519$$

De acuerdo con el resultado de la prueba de Durbin – Watson, el estadístico es cercano a 2, y dado que el p-value de la prueba es mayor a 0.05 no se rechaza la hipótesis nula de no autocorrelación. Por consiguiente, se concluye que no hay autocorrelación serial de primer orden.

RESET test

$$RESET = 2.7345, df1 = 2, df2 = 33, p\text{-value} = 0.07965$$

El resultado de p-value obtenido en la prueba de RESET indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

Shapiro-Wilk normality test

$$W = 0.97421, p\text{-value} = 0.366$$

El resultado de p-value de la prueba de Shapiro – Wilk mayor a 0.05 confirma la normalidad de los residuos.

13.5.2.1 Modelación metodología Paso a Paso.

A partir de este resultado del modelo con todas las variables, con el software R se realizó el análisis paso a paso, con la metodología atrás/adelante y utilizando el criterio de Akaike, con el fin de determinar las variables que mejor explicarían el modelo, y cuyos resultados se presentan a continuación:

Tabla 47. Modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), metodología paso a paso

Residuals:				
Min	1Q	Median	3Q	Max
-59509046	-25028336	-2320333	21746977	83186079
Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-11156363	38098598	-0.293	0.77109
Aconstruida	2203784	197859	11.138	4.08e-14 ***
Ascensor	32531986	14665066	2.218	0.03199 *
Carros	50160358	10650788	4.710	2.72e-05 ***
Estrato	31931861	9091157	3.512	0.00108 **
FittoCorvini	-20632382	12813914	-1.610	0.11486

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

Residual standard error: 35520000 on 42 degrees of freedom

Multiple R-squared: 0.8779, Adjusted R-squared: 0.8634

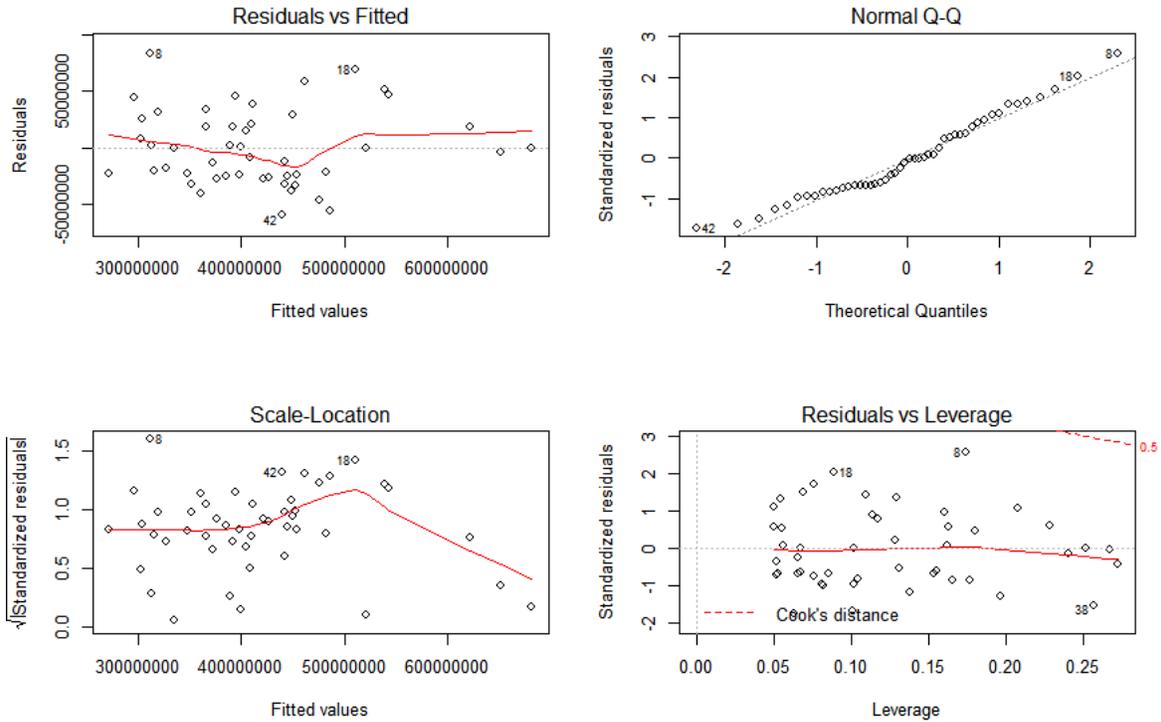
F-statistic: 60.41 on 5 and 42 DF, p-value: < 2.2e-16

En los resultados obtenidos con la modelación en R, se observa que nuevamente las variables área construida, ascensor, cantidad de carros y estrato son las variables más significativas estadísticamente, acorde con los resultados obtenidos de probabilidades (p-value). El coeficiente de determinación R² ajustado del modelo indica en forma porcentual que el 86.34% del valor del inmueble es explicado por las variables independientes analizadas, con un p-crítico del modelo de 2.2E-16, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es mucho mayor a 10, y por lo tanto, el modelo tiene una muy alta capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

Gráfico 52. Comportamiento de los residuos del modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), metodología paso a paso

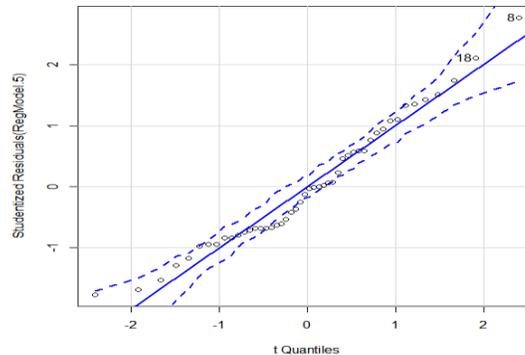
lm(Valor.del.negocio ~ Aconstruida + Ascensor + Carros + Estrato + FittoCor ...



De acuerdo con el resultado observado en el gráfico Residuals vs Fitted, se puede decir que visualmente el modelo que cumple con los supuestos de linealidad.

Así mismo, se observa en el gráfico Normal Q-Q la normalidad de los residuos, cumpliéndose de igual forma con el supuesto del modelo. En el gráfico de cuantiles se observa que los residuos studentizados se encuentran en general dentro de los intervalos, lo cual indica que siguen una distribución normal.

Gráfico 53. Cuantiles del modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), metodología paso a paso

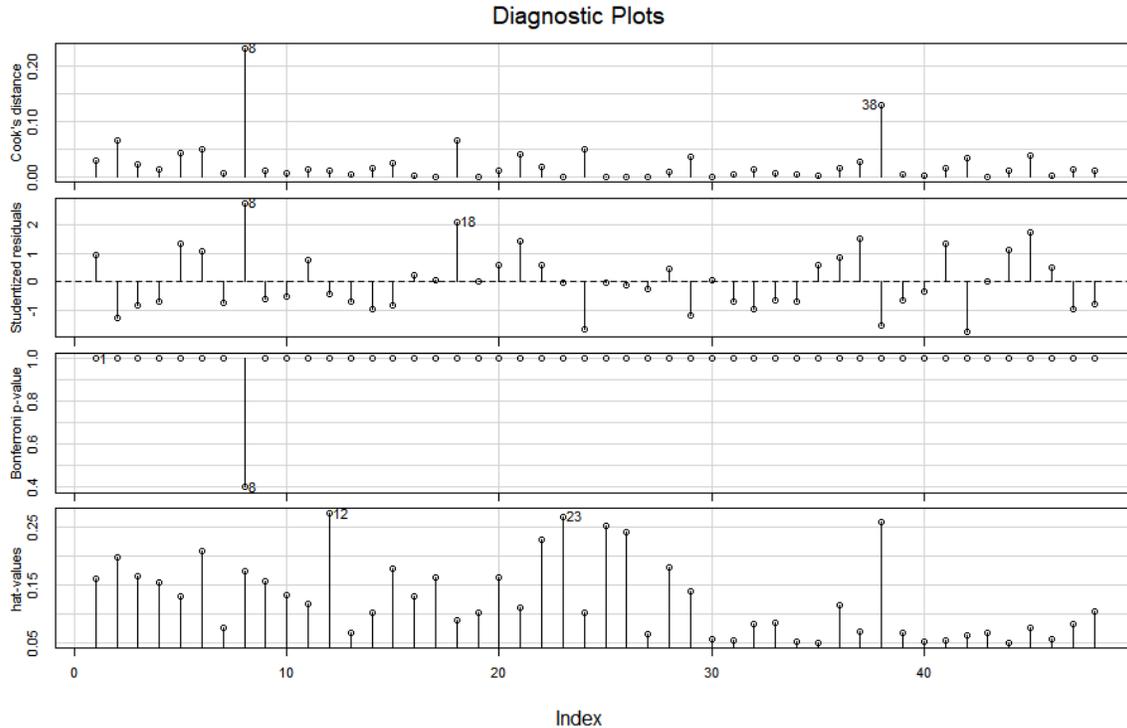


En la gráfica Scale-Location se observa que los residuos están distribuidos equitativamente a lo largo de los rangos de los predictores con lo cual se puede comprobar el supuesto de la varianza igual u homocedasticidad.

Finalmente, en la gráfica Residuals vs Leverage la cual ayuda a identificar los casos que son más influyentes en la regresión lineal (sean atípicos o no). En todos los casos se observa que los datos se encuentran dentro de la distancia Cook, aun cuando el programa muestra que hay unos datos que si se eliminaran podrían mejorar los resultados del coeficiente de determinación.

En la gráfica de índices de influencia a continuación, se observa la influencia de cada una de las muestras del estudio utilizadas para el modelo:

Gráfico 54. Índices de influencia del modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), metodología paso a paso



A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos.

Inflación de la Varianza

Aconstruida	Ascensor	Carros	Estrato	FittoCorvini
1.166516	1.758334	1.498931	1.829242	1.204134

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 0.043028, df = 1, p\text{-value} = 0.8357$$

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por consiguiente se asume la homocedasticidad.

Durbin-Watson test

$$DW = 1.9071, p\text{-value} = 0.3137$$

De acuerdo con el resultado de la prueba de Durbin – Watson, el estadístico es cercano a 2, y dado que el p-value de la prueba es mayor a 0.05 no se rechaza la hipótesis nula de no autocorrelación. Por consiguiente, se concluye que no hay autocorrelación serial de primer orden.

RESET test

$$RESET = 2.0446, df1 = 2, df2 = 40, p\text{-value} = 0.1427$$

El resultado de p-value obtenido en la prueba de RESET indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

Shapiro-Wilk normality test

$$W = 0.96529, p\text{-value} = 0.165$$

El resultado de p-value de la prueba de Shapiro – Wilk mayor a 0.05 confirma la normalidad de los residuos.

13.5.2.2 Modelación incluyendo variable distancia lineal.

Como el objetivo de la investigación es determinar la incidencia de la variable distancia, se introduce la variable distancia lineal Euclidiana en el modelo obtenido anteriormente, logrando los siguientes resultados:

Tabla 48. Modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia lineal

Residuals:

Min	1Q	Median	3Q	Max
-62095345	-24465894	-3990574	23529658	83904474

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-19944415	42651650	-0.468	0.6425
Aconstruida	2221151	203007	10.941	9.81e-14 ***
Ascensor	32035251	14838634	2.159	0.0368 *
Carros	49642590	10805018	4.594	4.09e-05 ***
Distancia.Lineal.m	9824	20625	0.476	0.6364
Estrato	32616764	9287999	3.512	0.0011 **
FittoCorvini	-21146739	12978516	-1.629	0.1109

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

Residual standard error: 35850000 on 41 degrees of freedom

Multiple R-squared: 0.8786, Adjusted R-squared: 0.8608

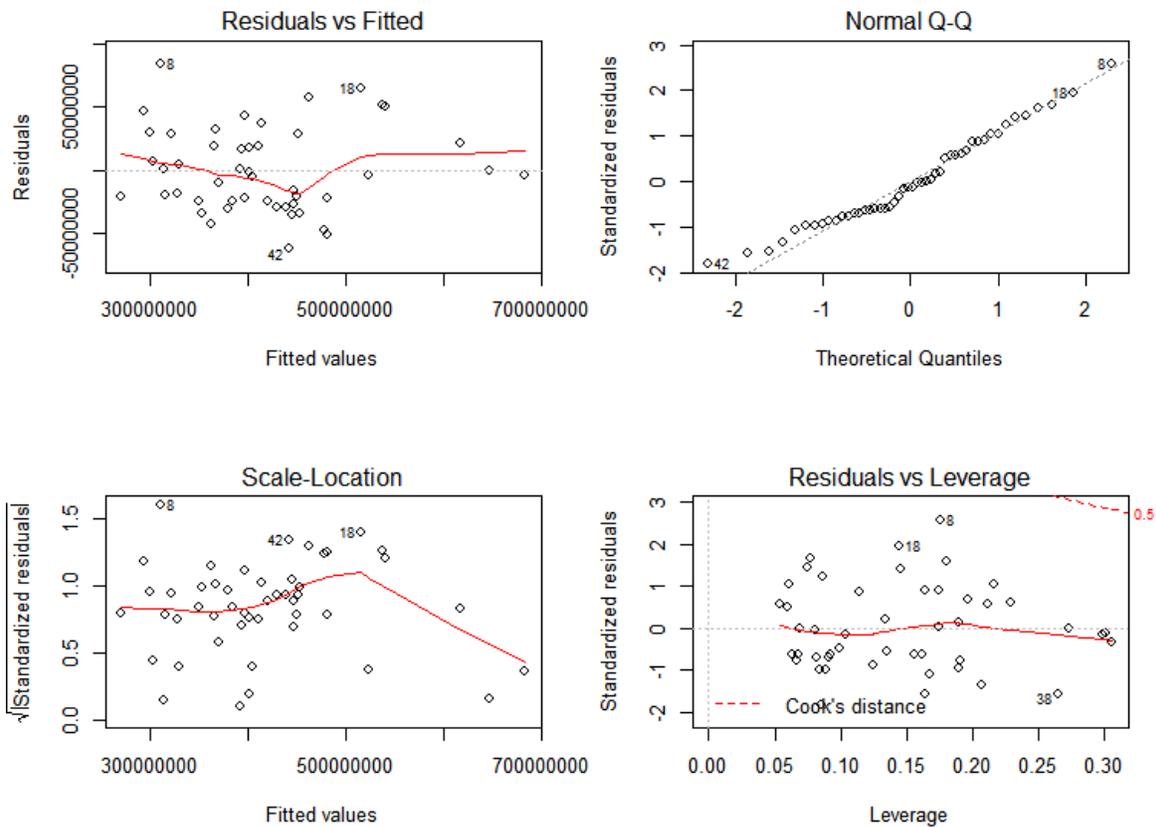
F-statistic: 49.46 on 6 and 41 DF, p-value: < 2.2e-16

En los resultados obtenidos con la modelación en R, se observa que nuevamente las variables área construida, ascensor, cantidad de carros y estrato son las variables más significativas estadísticamente, acorde con los resultados obtenidos de probabilidades (p-value). El coeficiente de determinación R² ajustado del modelo indica en forma porcentual que el 86.08% del valor del inmueble es explicado por las variables independientes analizadas, con un p-crítico del modelo de 2.2E-16, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es mucho mayor a 10, y por lo tanto, el modelo tiene una alta capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

Gráfico 55. Comportamiento de los residuos del modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia lineal

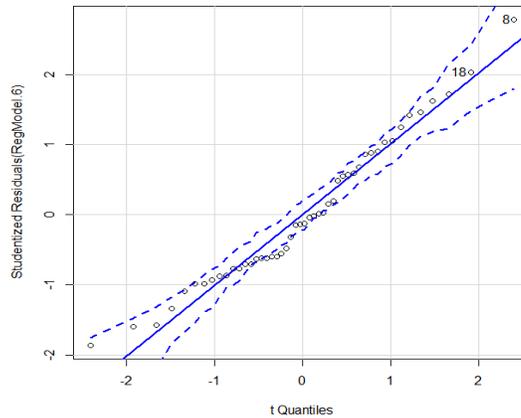
lm(Valor.del.negocio ~ Aconstruida + Ascensor + Carros + Distancia.Lineal.m ...



De acuerdo con el resultado observado en el gráfico Residuals vs Fitted, se puede decir que visualmente el modelo que cumple con los supuestos de linealidad.

Así mismo, se observa en el gráfico Normal Q-Q la normalidad de los residuos, cumpliéndose de igual forma con el supuesto del modelo. En el gráfico de cuantiles se observa que los residuos studentizados se encuentran en general dentro de los intervalos, lo cual indica que siguen una distribución normal.

Gráfico 56. Cuantiles de regresión lineal con distancia lineal incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia lineal

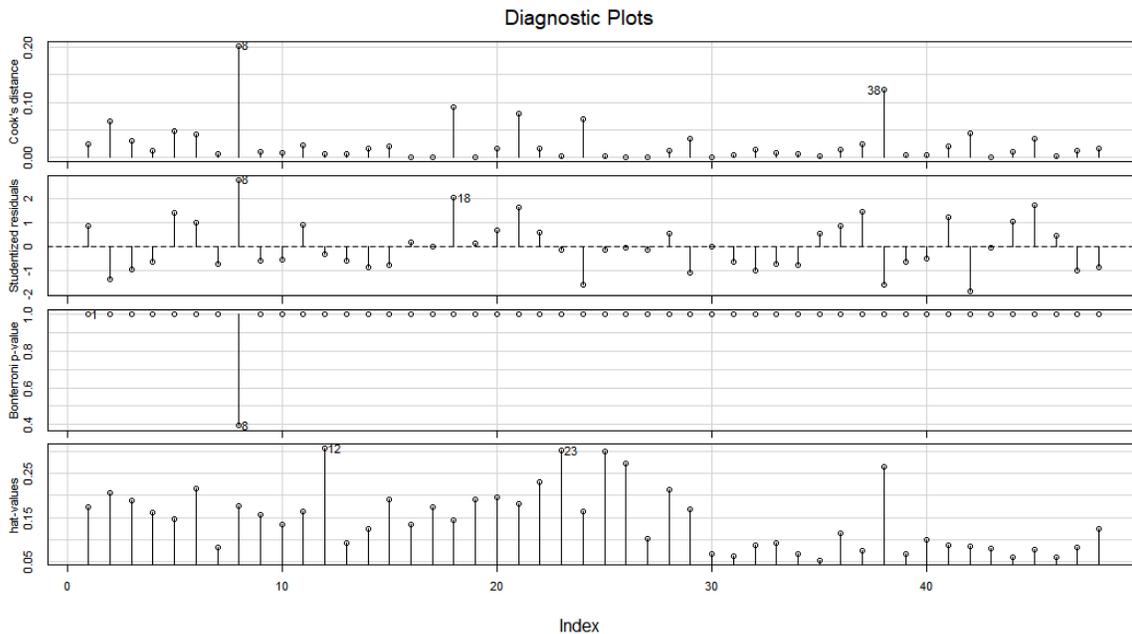


En la gráfica Scale-Location se observa que los residuos están distribuidos equitativamente a lo largo de los rangos de los predictores con lo cual se puede comprobar el supuesto de la varianza igual u homocedasticidad.

Finalmente, en la gráfica Residuals vs Leverage la cual ayuda a identificar los casos que son más influyentes en la regresión lineal (sean atípicos o no). En todos los casos se observa que los datos se encuentran dentro de la distancia Cook, aun cuando el programa muestra que hay unos datos que si se eliminaran podrían mejorar los resultados del coeficiente de determinación.

En la gráfica de índices de influencia a continuación, se observa la influencia de cada una de las muestras del estudio utilizadas para el modelo:

Gráfico 57. Índices de influencia del modelo de regresión lineal con distancia lineal incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia lineal



A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos.

Inflación de la Varianza

Aconstruida	Ascensor	Carros	Distancia.Lineal.m
1.205402	1.767063	1.514258	1.058637
	Estrato	FittoCorvini	
	1.874166	1.212529	

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 0.026654, df = 1, p\text{-value} = 0.8703$$

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por consiguiente se asume la homocedasticidad.

Durbin-Watson test

$$DW = 1.866, p\text{-value} = 0.2484$$

De acuerdo con el resultado de la prueba de Durbin – Watson, el estadístico es cercano a 2, y dado que el p-value de la prueba es mayor a 0.05 no se rechaza la hipótesis nula de no autocorrelación. Por consiguiente, se concluye que no hay autocorrelación serial de primer orden.

RESET test

$$RESET = 2.5178, df1 = 2, df2 = 39, p\text{-value} = 0.09366$$

El resultado de p-value obtenido en la prueba de RESET indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

Shapiro-Wilk normality test

$$W = 0.9673, p\text{-value} = 0.1984$$

El resultado de p-value de la prueba de Shapiro – Wilk mayor a 0.05 confirma la normalidad de los residuos.

13.5.3 Modelaciones con distancia en carro.

Partimos de la aplicación del modelo lineal con el software estadístico R, para todas las variables preseleccionadas, incluyendo en la evaluación la variable distancia en carro, obteniendo el siguiente resultado:

Tabla 49. Modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), incluyendo todas las variables

Residuals:				
Min	1Q	Median	3Q	Max
-61509148	-27785764	-3434829	28714411	77539551
Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-991961	88983390	-0.011	0.991169
Aconstruida	2234465	294155	7.596	0.00000000661 ***
Ascensor	35557224	18594502	1.912	0.064058 .
Carros	48866254	13169395	3.711	0.000715 ***
CuartoUtil	-4707432	17420273	-0.270	0.788573
Distancia.Car..m.	-3548	20789	-0.171	0.865463
Estrato	30251038	14431396	2.096	0.043363 *
FittoCorvini	-20986037	15318660	-1.370	0.179423
Habitaciones	-497994	12737743	-0.039	0.969036
NivelPiso	597776	2108592	0.283	0.778467
Poniente	5515089	15614439	0.353	0.726054
Porteria	-6918929	10329019	-0.670	0.507347
ZonaComun	8938833	9543996	0.937	0.355386

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

Residual standard error: 37880000 on 35 degrees of freedom

Multiple R-squared: 0.8843, Adjusted R-squared: 0.8446

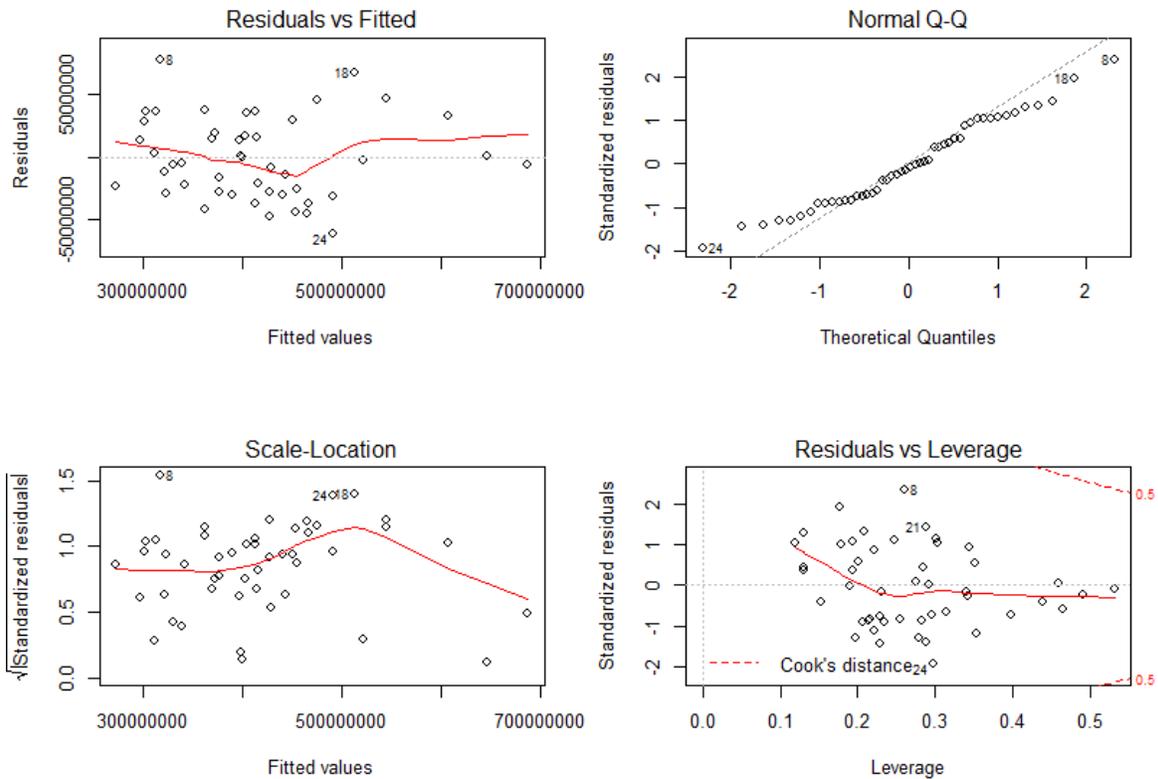
F-statistic: 22.29 on 12 and 35 DF, p-value: 6.781e-13

En los resultados obtenidos con la modelación en R, se observa que nuevamente las variables área construida, ascensor, cantidad de carros y estrato son las variables más significativas estadísticamente, acorde con los resultados obtenidos de probabilidades (p-value). El coeficiente de determinación R² ajustado del modelo indica en forma porcentual que el 84.46% del valor del inmueble es explicado por las variables independientes analizadas, con un p-crítico del modelo de 6.781E-13, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es mayor a 10, y por lo tanto, el modelo tiene una buena capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

Gráfico 58. Comportamiento de los residuos del modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), incluyendo todas las variables

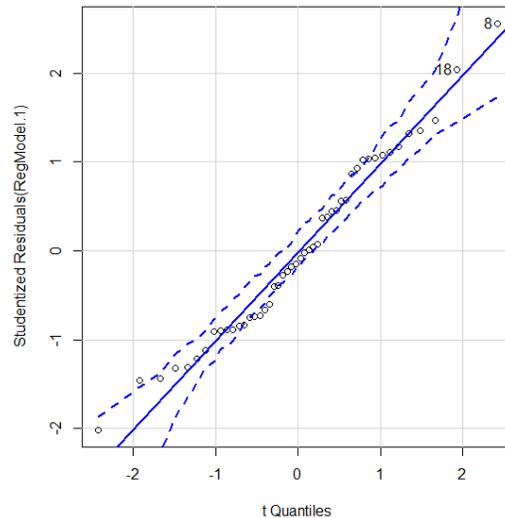
lm(Valor.del.negocio ~ Aconstruida + Ascensor + Carros + CuartoUtil + Dista ...



De acuerdo con el resultado observado en el gráfico Residuals vs Fitted, se puede decir que visualmente el modelo que cumple con los supuestos de linealidad.

Así mismo, se observa en el gráfico Normal Q-Q la normalidad de los residuos, cumpliéndose de igual forma con el supuesto del modelo. En el gráfico de cuantiles se observa que los residuos studentizados se encuentran en general dentro de los intervalos, lo cual indica que siguen una distribución normal.

Gráfico 59. Cuantiles del modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), incluyendo todas las variables

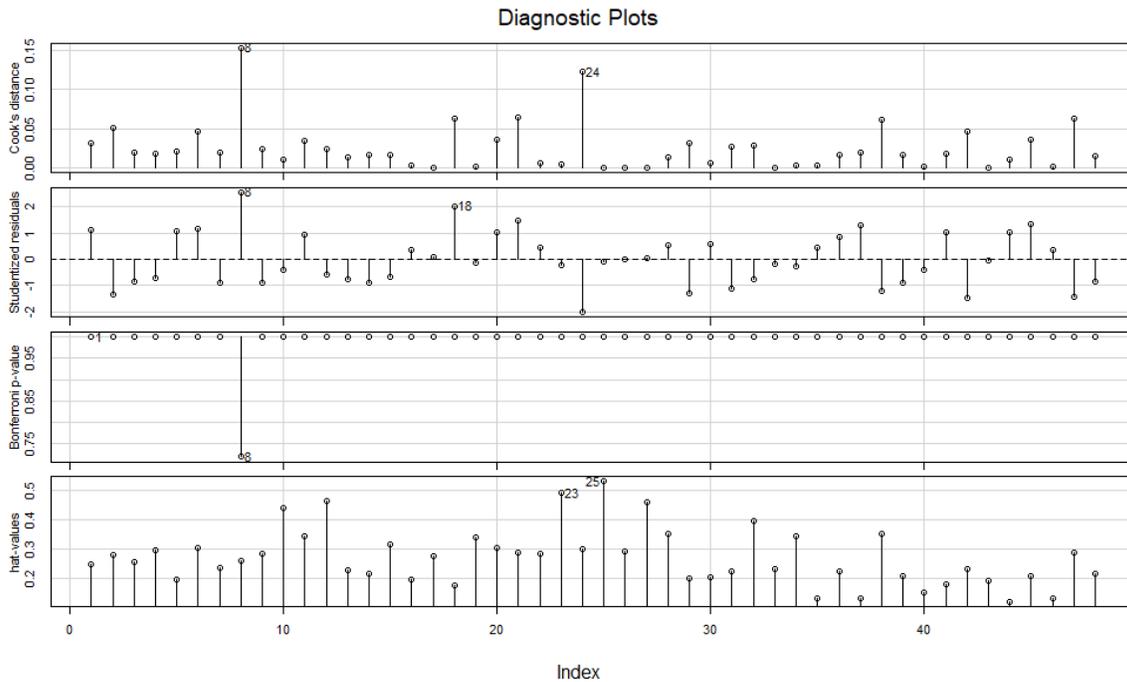


En la gráfica Scale-Location se observa que los residuos están distribuidos equitativamente a lo largo de los rangos de los predictores con lo cual se puede comprobar el supuesto de la varianza igual u homocedasticidad.

Finalmente, en la gráfica Residuals vs Leverage la cual ayuda a identificar los casos que son más influyentes en la regresión lineal (sean atípicos o no). En todos los casos se observa que los datos se encuentran dentro de la distancia Cook, aun cuando el programa muestra que hay unos datos que si se eliminaran podrían mejorar los resultados del coeficiente de determinación.

En la gráfica de índices de influencia a continuación, se observa la influencia de cada una de las muestras del estudio utilizadas para el modelo:

Gráfico 60. Índices de influencia modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), incluyendo todas las variables



A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos.

Inflación de la Varianza

Aconstruida	Ascensor	Carros	CuartoUtil
2.266808	2.485328	2.014796	2.097618
Distancia.Car..m.	Estrato	FittoCorvini	Habitaciones
3.093536	4.052574	1.512983	2.862666
NivelPiso	Poniente	Porteria	ZonaComun
1.877991	1.440976	3.419243	2.708943

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 0.28365, df = 1, p\text{-value} = 0.5943$$

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por consiguiente se asume la homocedasticidad.

Durbin-Watson test

$$DW = 1.9762, p\text{-value} = 0.3719$$

De acuerdo con el resultado de la prueba de Durbin – Watson, el estadístico es cercano a 2, y dado que el p-value de la prueba es mayor a 0.05 no se rechaza la hipótesis nula de no autocorrelación. Por consiguiente, se concluye que no hay autocorrelación serial de primer orden.

RESET test

$$RESET = 2.0978, df1 = 2, df2 = 33, p\text{-value} = 0.1388$$

El resultado de p-value obtenido en la prueba de RESET indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

Shapiro-Wilk normality test

$$W = 0.97163, p\text{-value} = 0.2927$$

El resultado de p-value de la prueba de Shapiro – Wilk mayor a 0.05 confirma la normalidad de los residuos.

13.5.3.1 Modelación metodología Paso a Paso.

A partir de este resultado del modelo con todas las variables, con el software R se realizó el análisis paso a paso, con la metodología atrás/adelante y utilizando el criterio de Akaike, con el fin de determinar las variables que mejor explicarían el modelo, y cuyos resultados se presentan a continuación:

Tabla 50. Modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), metodología paso a paso

Residuals:				
Min	1Q	Median	3Q	Max
-59509046	-25028336	-2320333	21746977	83186079
Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-11156363	38098598	-0.293	0.77109
Aconstruida	2203784	197859	11.138	4.08e-14 ***
Ascensor	32531986	14665066	2.218	0.03199 *
Carros	50160358	10650788	4.710	2.72e-05 ***
Estrato	31931861	9091157	3.512	0.00108 **
FittoCorvini	-20632382	12813914	-1.610	0.11486

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

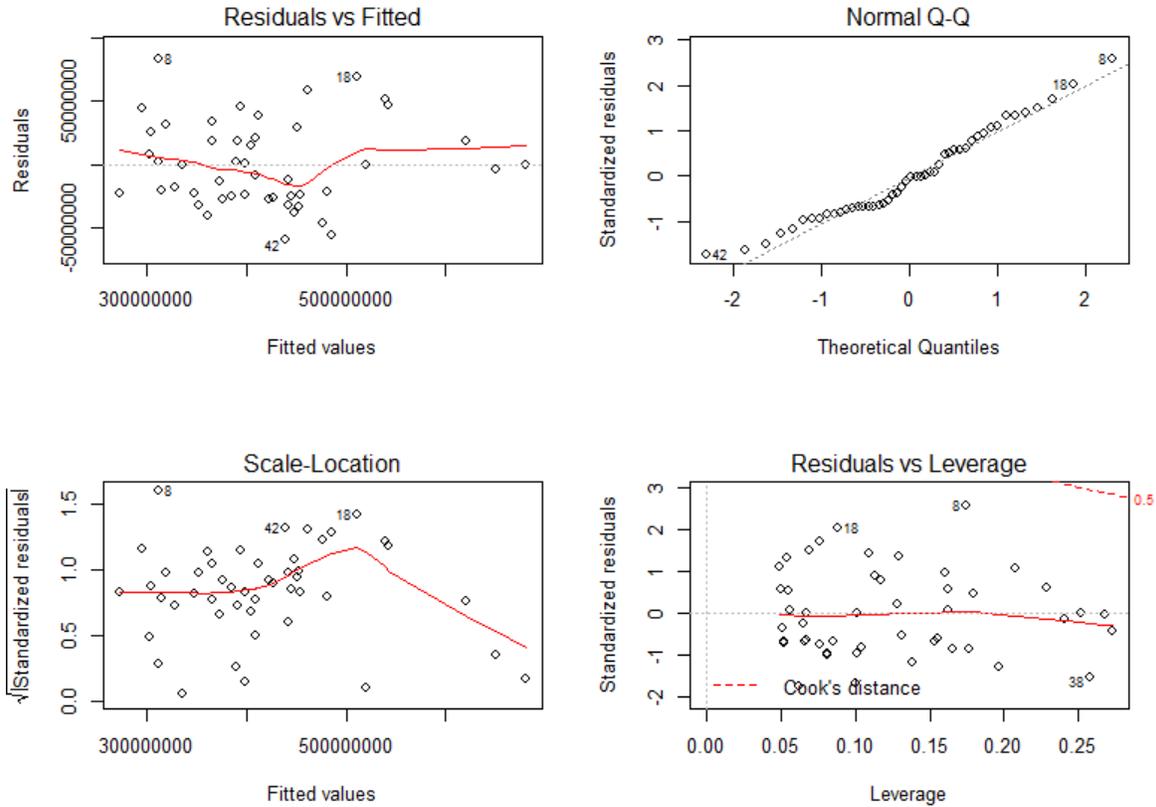
Residual standard error: 35520000 on 42 degrees of freedom
 Multiple R-squared: 0.8779, Adjusted R-squared: 0.8634
 F-statistic: 60.41 on 5 and 42 DF, p-value: < 2.2e-16

En los resultados obtenidos con la modelación en R, se observa que nuevamente las variables área construida, ascensor, cantidad de carros y estrato son las variables más significativas estadísticamente, acorde con los resultados obtenidos de probabilidades (p-value). El coeficiente de determinación R² ajustado del modelo indica en forma porcentual que el 84.46% del valor del inmueble es explicado por las variables independientes analizadas, con un p-crítico del modelo de 6.781E-13, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es mayor a 10, y por lo tanto, el modelo tiene una buena capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

Gráfico 61. Comportamiento de los residuos del modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), metodología paso a paso

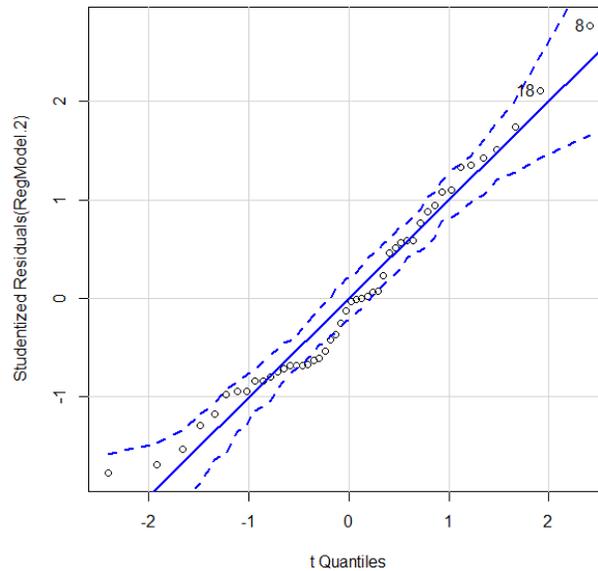
lm(Valor.del.negocio ~ Aconstruida + Ascensor + Carros + Estrato + FittoCor ...



De acuerdo con el resultado observado en el gráfico Residuals vs Fitted, se puede decir que visualmente el modelo que cumple con los supuestos de linealidad.

Así mismo, se observa en el gráfico Normal Q-Q la normalidad de los residuos, cumpliéndose de igual forma con el supuesto del modelo. En el gráfico de cuantiles se observa que los residuos studentizados se encuentran en general dentro de los intervalos, lo cual indica que siguen una distribución normal.

Gráfico 62. Cuantiles del modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), metodología paso a paso

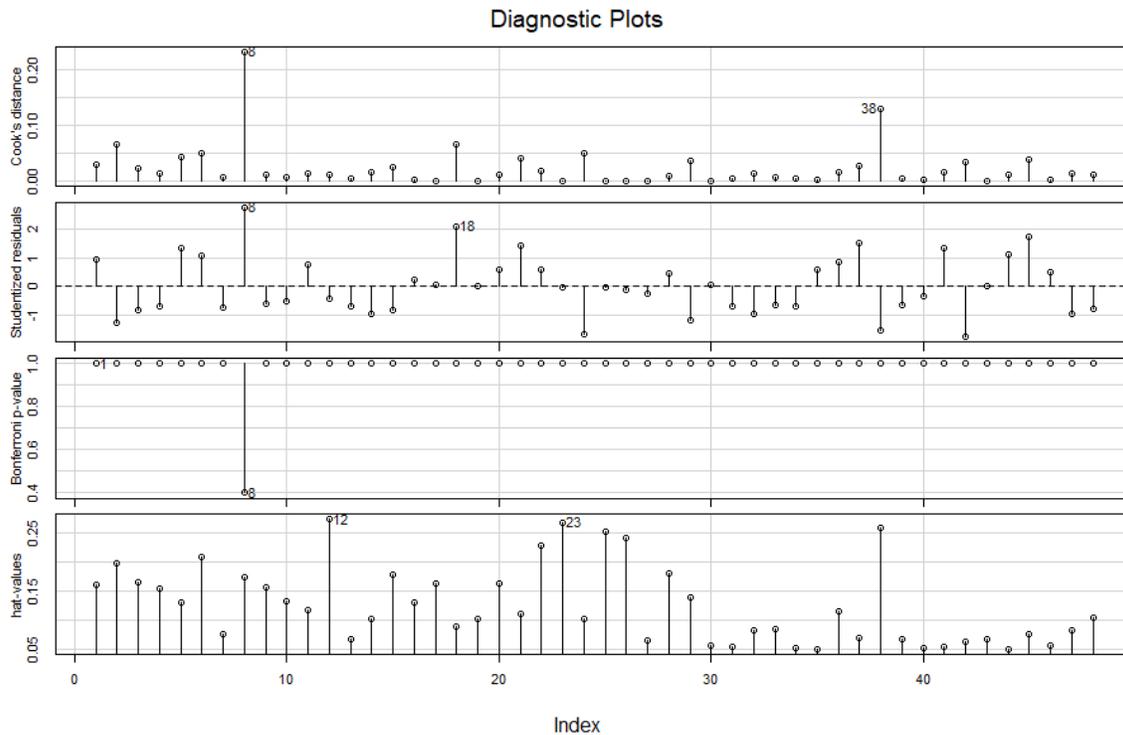


En la gráfica Scale-Location se observa que los residuos están distribuidos equitativamente a lo largo de los rangos de los predictores, con algunos puntos por fuera del intervalo, con lo cual se puede comprobar el supuesto de la varianza igual u homocedasticidad, pero se debe verificar a través de los test de hipótesis.

Finalmente, en la gráfica Residuals vs Leverage la cual ayuda a identificar los casos que son más influyentes en la regresión lineal (sean atípicos o no). En todos los casos se observa que los datos se encuentran dentro de la distancia Cook, aun cuando el programa muestra que hay unos datos que si se eliminaran podrían mejorar los resultados del coeficiente de determinación.

En la gráfica de índices de influencia a continuación, se observa la influencia de cada una de las muestras del estudio utilizadas para el modelo:

Gráfico 63. Índices de influencia del modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), metodología paso a paso



A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos.

Inflación de la Varianza

Aconstruida	Ascensor	Carros	Estrato	FittoCorvini
1.166516	1.758334	1.498931	1.829242	1.204134

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

$$BP = 0.043028, df = 1, p\text{-value} = 0.8357$$

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por consiguiente se asume la homocedasticidad.

Durbin-Watson test

$$DW = 1.9071, p\text{-value} = 0.3137$$

De acuerdo con el resultado de la prueba de Durbin – Watson, el estadístico es cercano a 2, y dado que el p-value de la prueba es mayor a 0.05 no se rechaza la hipótesis nula de no autocorrelación. Por consiguiente, se concluye que no hay autocorrelación serial de primer orden.

RESET test

$$RESET = 2.0446, df1 = 2, df2 = 40, p\text{-value} = 0.1427$$

El resultado de p-value obtenido en la prueba de RESET indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

Shapiro-Wilk normality test

$$W = 0.96529, p\text{-value} = 0.165$$

El resultado de p-value de la prueba de Shapiro – Wilk mayor a 0.05 confirma la normalidad de los residuos.

13.5.3.2 Modelación Incluyendo Variable Distancia en Carro.

Como el objetivo de la investigación es determinar la incidencia de la variable distancia, se introduce la variable distancia lineal Euclidiana en el modelo obtenido anteriormente, logrando los siguientes resultados:

Tabla 51. Modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia lineal

Residuals:				
Min	1Q	Median	3Q	Max
-54035993	-25495276	-4007616	19514031	84521931
Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	27895760	73912798	0.377	0.7078
Aconstruida	2221342	201343	11.033	7.59e-14 ***
Ascensor	33753139	14905573	2.264	0.0289 *
Carros	49402550	10799778	4.574	4.36e-05 ***
Distancia.Car..m.	-11123	17991	-0.618	0.5398
Estrato	26620316	12557602	2.120	0.0401 *
FittoCorvini	-22099463	13125493	-1.684	0.0998 .

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

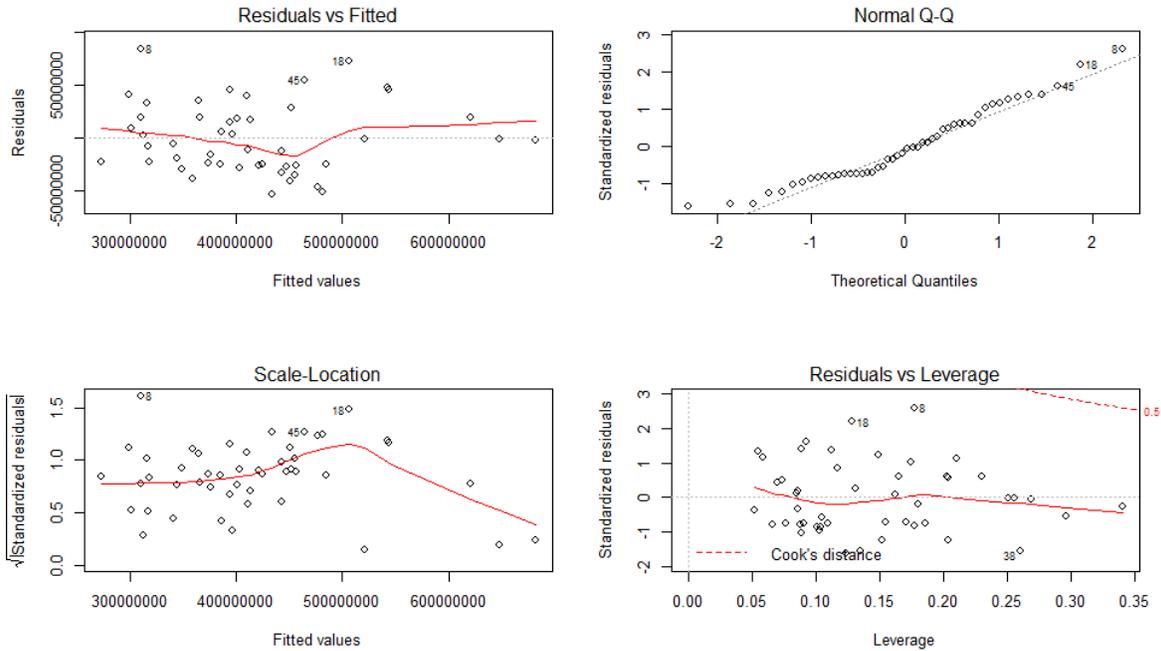
Residual standard error: 35780000 on 41 degrees of freedom
 Multiple R-squared: 0.8791, Adjusted R-squared: 0.8614
 F-statistic: 49.67 on 6 and 41 DF, p-value: < 2.2e-16

En los resultados obtenidos con la modelación en R, se observa que nuevamente las variables área construida, ascensor, cantidad de carros y estrato son las variables más significativas estadísticamente, acorde con los resultados obtenidos de probabilidades (p-value). El coeficiente de determinación R² ajustado del modelo indica en forma porcentual que el 86.14% del valor del inmueble es explicado por las variables independientes analizadas, con un p-crítico del modelo de 2.2E-16, casi cero, que aceptaría el modelo. Así mismo, el estadístico F es mayor a 10, y por lo tanto, el modelo tiene una buena capacidad explicativa.

En las gráficas a continuación se observa el comportamiento de los residuos del modelo de regresión lineal obtenido con las variables analizadas, con el fin de verificar visualmente los supuestos de los modelos de regresión lineal:

Gráfico 64. Comportamiento de los residuos del modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia lineal

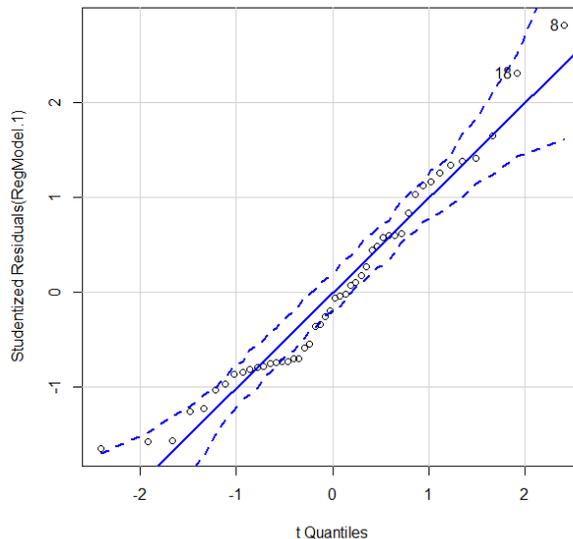
lm(Valor.del.negocio ~ Aconstruida + Ascensor + Carros + Distancia.Car..m. ...



De acuerdo con el resultado observado en el gráfico Residuals vs Fitted, se puede decir que visualmente el modelo que cumple con los supuestos de linealidad.

Así mismo, se observa en el gráfico Normal Q-Q la normalidad de los residuos, cumpliéndose de igual forma con el supuesto del modelo. En el gráfico de cuantiles se observa que los residuos studentizados se encuentran en general dentro de los intervalos, lo cual indica que siguen una distribución normal.

Gráfico 65. Cuantiles del modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia lineal

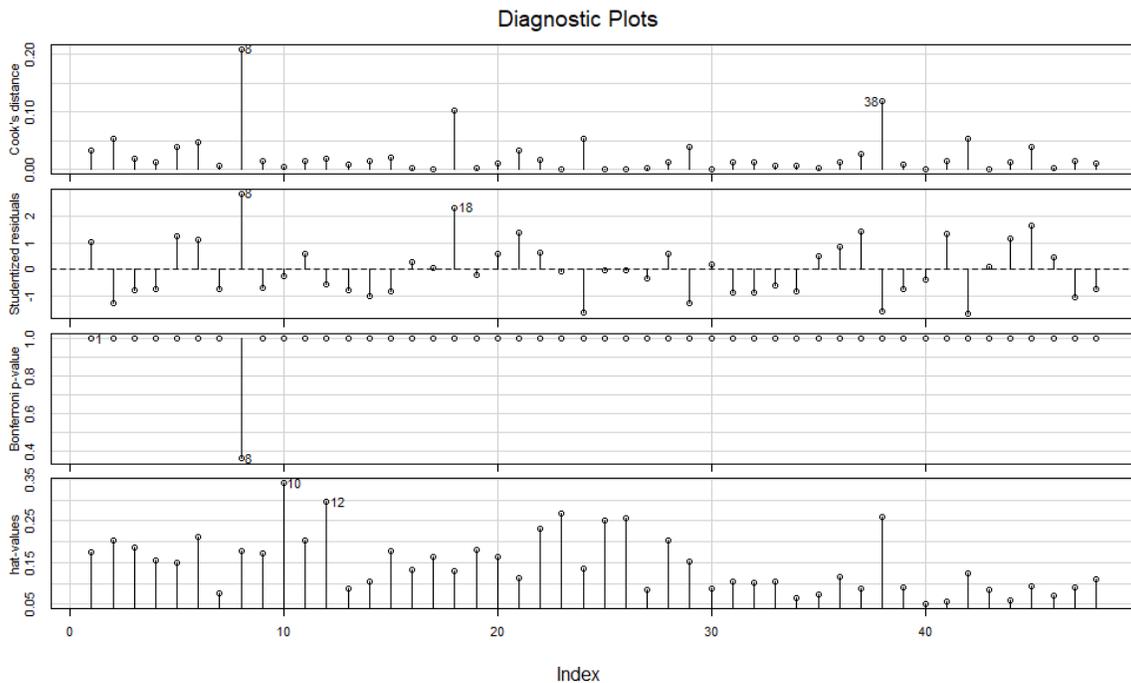


En la gráfica Scale-Location se observa que los residuos están distribuidos equitativamente a lo largo de los rangos de los predictores, con algunos puntos por fuera del intervalo, con lo cual se puede comprobar el supuesto de la varianza igual u homocedasticidad, pero se debe verificar a través de los test de hipótesis.

Finalmente, en la gráfica Residuals vs Leverage la cual ayuda a identificar los casos que son más influyentes en la regresión lineal (sean atípicos o no). En todos los casos se observa que los datos se encuentran dentro de la distancia Cook, aun cuando el programa muestra que hay unos datos que si se eliminaran podrían mejorar los resultados del coeficiente de determinación.

En la gráfica de índices de influencia a continuación, se observa la influencia de cada una de las muestras del estudio utilizadas para el modelo:

Gráfico 66. Índices de influencia de los residuos del modelo de regresión lineal con distancia en carro incluyendo outliers (53 muestras), incluyendo al mejor modelo la distancia lineal



A continuación, se realiza el análisis de los supuestos del modelo de regresión lineal a través de la aplicación de los test de hipótesis ya descritos.

Inflación de la Varianza

Aconstruida	Ascensor	Carros	Distancia.Car..m.
1.190196	1.789762	1.518491	2.596566
	Estrato	FittoCorvini	
	3.438832	1.244821	

De acuerdo con el resultado de la prueba de inflación de la varianza, los resultados obtenidos en el software R son inferiores a 5, y por lo tanto se puede concluir que no hay una fuerte colinealidad entre las variables utilizadas en el modelo.

Breusch-Pagan test

BP = 0.0465, df = 1, p-value = 0.8293

De acuerdo con el resultado de la prueba de Breusch – Pagan, el p-value de la prueba es mayor a 0.05 y por lo tanto no se rechaza la hipótesis nula de no heteroscedasticidad, y por consiguiente se asume la homocedasticidad.

Durbin-Watson test

$$DW = 1.9508, p\text{-value} = 0.3826$$

De acuerdo con el resultado de la prueba de Durbin – Watson, el estadístico es cercano a 2, y dado que el p-value de la prueba es mayor a 0.05 no se rechaza la hipótesis nula de no autocorrelación. Por consiguiente, se concluye que no hay autocorrelación serial de primer orden.

RESET test

$$RESET = 1.8695, df1 = 2, df2 = 39, p\text{-value} = 0.1678$$

El resultado de p-value obtenido en la prueba de RESET indica que la forma funcional (linealidad) es correcta, de tal manera que el modelo está bien especificado.

Shapiro-Wilk normality test

$$W = 0.95797, p\text{-value} = 0.08359$$

El resultado de p-value de la prueba de Shapiro – Wilk mayor a 0.05 confirma la normalidad de los residuos.

13.5.4 Resumen de las modelaciones.

En la tabla a continuación se resumen los resultados de los modelos definidos a través de la metodología de paso a paso (Step Wise) con el programa estadístico R:

Tabla 52. Resumen modelos paso a paso

Modelación	Coefficientes	p-value < a	Multiple R-squared	Adjusted R-squared	F-statistic:	p-value < a
	(Intercept)	123165644	0.00607 **			
	Aconstruida	1348401	0.00000221 ***			
Excluyendo datos atípicos	Carros	16759180	0.08811 .	0,7478	0,7146	22,54
	Estrato	30450056	0.00022 ***			1,95E-10
	FittoCorvini	-24105809	0.04704 *			
	ZonaComun	20871685	0.00375 **			
	(Intercept)	-11156363	0.77109			
	Aconstruida	2203784	4.08e-14 ***			
Incluyendo datos atípicos	Ascensor	32531986	0.03199 *	0,8779	0,8634	60,41
	Carros	50160358	2.72e-05 ***			2,20E-16
	Estrato	31931861	0.00108 **			
	FittoCorvini	-20632382	0.11486			

En la tabla a continuación se relaciona el resumen de resultados de las modelaciones que incluyen la variable distancia, en las cuales se utilizaron la totalidad de los datos del muestreo.

Tabla 53. Resumen modelos distancia con todos los datos

Modelo	Distancia evaluada	Coeficientes		p-value < a	Multiple R-squared	Adjusted R-squared	F-statistic:	p-value < a
		(Intercept)	-31890285	0,468594				
		Aconstruida	2219611	5,09E-14				
		Ascensor	32762549	0,031089				
Modelo 2	Distancia a Pie (m)	Carros	49117014	4,19E-05	0,8807	0,8633	50,46	2,20E-16
		Distancia.a.Pie..m.	18122	0,332707				
		Estrato	33235341	0,000812				
		FittoCorvini	-20853125	0,111533				
		(Intercept)	-19944415	0,6425				
		Aconstruida	2221151	9,81E-14				
		Ascensor	32035251	0,0368				
Modelo 4	Distancia Euclidiana (Lineal)	Carros	49642590	4,09E-05	0,8786	0,8608	49,46	2,20E-16
		Distancia.Lineal.m	9824	0,6364				
		Estrato	32616764	0,0011				
		FittoCorvini	-21146739	0,1109				
		(Intercept)	27895760	0.7078				
		Aconstruida	2221342	7.59e-14				
		Ascensor	33753139	0.0289				
Modelo 6	Distancia en Carro	Carros	49402550	4.36e-05	0,8791	0,8614	49,67	2,20E-16
		Distancia.Car..m.	-11123	0.5398				
		Estrato	26620316	0.0401				
		FittoCorvini	-22099463	0.0998				

En la tabla a continuación se relaciona el resumen de resultados de las modelaciones que incluyen la variable distancia, en las cuales se utilizaron los datos que excluyen los atípicos.

Tabla 54. Resumen modelos distancia excluyendo datos atípicos

Modelo	Variable Distancia	Coefficientes		p-value < a	Multiple R-squared	Adjusted R-squared	F-statistic:	p-value < a
		(Intercept)	117808603	0,020662				
		Aconstruida	1347432	3,03E-06				
		Carros	16718293	0,093048				
Modelo 1	Distancia a Pie	Distancia.a.Pie..m.	4310	0,817753	0,7482	0,7074	18,32	9,63E-10
		Estrato	30749669	0,000279				
		FittoCorvini	-23539693	0,060158				
		ZonaComun	20628831	0,005073				
		(Intercept)	122822938.2	0,01679				
		Aconstruida	1348683.9	3,18E-06				
		Carros	16759758.6	0,092448				
Modelo 3	Distancia Euclidiana (Lineal)	Distancia.Lineal.m	315.4	0,988527	0,7478	0,7069	18,29	9,882E-07
		Estrato	30474372.6	0,000351				
		FittoCorvini	-24075505.9	0,053749				
		ZonaComun	20839032.6	0,006407				
		(Intercept)	142584873	0,04807				
		Aconstruida	1348196	2,94E-06				
		Carros	16253684	0,10537				
Modelo 5	Distancia en Carro	Distancia.Car..m.	-5713	0,72606	0,7487	0,7079	18,37	9,307E-07
		Estrato	28342145	0,00554				
		FittoCorvini	-25134858	0,04703				
		ZonaComun	20513152	0,00523				

13.6 Verificación de la incidencia de la distancia

Se verificó la incidencia de la distancia utilizando un apartamento tipo, basado en la mediana obtenida para los datos de cada variable, fijando dentro del modelo dichos valores y modificando solamente la variable distancia para verificar su incidencia. Para esta verificación de la incidencia solo se utilizaron los modelos 2, 4 y 6 en los cuales se utilizó la totalidad de los datos, depurados por incidencia como se explicó, utilizando los siguientes datos para las variables explicativas:

Tabla 55. Datos para verificación de la incidencia distancia

Dato	Valor
Área construida (m2)	105
Cantidad de Ascensores	1
Cantidad de carros	1
Estrato	5
Fitto Corvini	1,5

Al hacer uso del modelo 2, con las variables anteriormente definidas y variando la distancia a pie se obtienen los siguientes resultados:

Tabla 56. Verificación incidencia distancia modelo 2

Distancia a pie al Centro Comercial Viva Envigado (m)		Valor Total	Valor acumulado debido a la distancia	% respecto al valor del inmueble
50	\$	418.851.551		
100	\$	419.757.651	\$ 906.100	0,2%
200	\$	421.569.851	\$ 2.718.300	0,6%
300	\$	423.382.051	\$ 4.530.500	1,1%
400	\$	425.194.251	\$ 6.342.700	1,5%
500	\$	427.006.451	\$ 8.154.900	1,9%
600	\$	428.818.651	\$ 9.967.100	2,3%
700	\$	430.630.851	\$ 11.779.300	2,7%
800	\$	432.443.051	\$ 13.591.500	3,1%
900	\$	434.255.251	\$ 15.403.700	3,5%
1000	\$	436.067.451	\$ 17.215.900	3,9%
1100	\$	437.879.651	\$ 19.028.100	4,3%
1200	\$	439.691.851	\$ 20.840.300	4,7%
1300	\$	441.504.051	\$ 22.652.500	5,1%
1400	\$	443.316.251	\$ 24.464.700	5,5%
1500	\$	445.128.451	\$ 26.276.900	5,9%

De acuerdo con el modelo 4, en el cual se varió la distancia lineal al Centro Comercial Viva Envigado, los resultados de la incidencia de la distancia se presentan a continuación:

Tabla 57. Verificación incidencia distancia modelo 4

Distancia lineal al Comercial Viva Envigado (m)	Valor Total	% respecto al valor del inmueble	Distancia a pie al Centro Comercial Viva Envigado (m)
50	\$ 426.809.193		
100	\$ 427.300.393	\$ 491.200	0,1%
200	\$ 428.282.793	\$ 1.473.600	0,3%
300	\$ 429.265.193	\$ 2.456.000	0,6%
400	\$ 430.247.593	\$ 3.438.400	0,8%
500	\$ 431.229.993	\$ 4.420.800	1,0%
600	\$ 432.212.393	\$ 5.403.200	1,3%
700	\$ 433.194.793	\$ 6.385.600	1,5%
800	\$ 434.177.193	\$ 7.368.000	1,7%
900	\$ 435.159.593	\$ 8.350.400	1,9%
1000	\$ 436.141.993	\$ 9.332.800	2,1%
1100	\$ 437.124.393	\$ 10.315.200	2,4%
1200	\$ 438.106.793	\$ 11.297.600	2,6%
1300	\$ 439.089.193	\$ 12.280.000	2,8%
1400	\$ 440.071.593	\$ 13.262.400	3,0%
1500	\$ 441.053.993	\$ 14.244.800	3,2%

Respecto a la incidencia de la distancia en carro al centro comercial Viva Envigado, en los valores de los inmuebles, aplicando el modelo 6 se obtienen los siguientes resultados:

Tabla 58. Verificación incidencia distancia modelo 6

Distancia en carro al Comercial Viva Envigado (m)	Valor Total	% respecto al valor del inmueble	Distancia a pie al Centro Comercial Viva Envigado (m)
50	\$ 443.688.595		
100	\$ 443.132.445	-\$ 556.150	0,1%
200	\$ 442.020.145	-\$ 1.668.450	0,4%
300	\$ 440.907.845	-\$ 2.780.750	0,6%
400	\$ 439.795.545	-\$ 3.893.050	0,9%
500	\$ 438.683.245	-\$ 5.005.350	1,1%
600	\$ 437.570.945	-\$ 6.117.650	1,4%
700	\$ 436.458.645	-\$ 7.229.950	1,7%
800	\$ 435.346.345	-\$ 8.342.250	1,9%
900	\$ 434.234.045	-\$ 9.454.550	2,2%
1000	\$ 433.121.745	-\$ 10.566.850	2,4%
1100	\$ 432.009.445	-\$ 11.679.150	2,7%
1200	\$ 430.897.145	-\$ 12.791.450	3,0%
1300	\$ 429.784.845	-\$ 13.903.750	3,2%
1400	\$ 428.672.545	-\$ 15.016.050	3,5%
1500	\$ 427.560.245	-\$ 16.128.350	3,8%

14 CONCLUSIONES Y RECOMENDACIONES

Las modelaciones realizadas con la totalidad de los datos, depuradas luego de verificar los datos que influyen en los resultados y detectados a través de las pruebas de influencia, tienen una mejor capacidad explicativa al ser comparados con los modelos en los que los datos atípicos, en el análisis de normalidad de los valores del negocio, no fueron tenidos en cuenta en la modelación.

Al realizar la selección de las variables explicativas de los valores comerciales de los inmuebles residenciales en multifamiliares ubicados entre las calles 26 sur y calle 40sur, y entre las carreras 43A Avenida el Poblado y la Carrera 48 Avenida Las Vegas, el modelo que mejor representa el fenómeno se representa a continuación:

$$\text{Valor Inmueble} = -11156363 + 2203784 * \text{Área Construida} + 32531986 * \text{Ascensor} + 50160358 * \text{Cantidad Parquaderos} + 31931861 * \text{Estrato} - 20632382 * \text{Estado de Conservación}$$

En dicho modelo, que no incluye las variables de distancia, las variables explicativas: área construida del inmueble, cantidad de parqueaderos que tiene el inmueble y el estrato donde se encuentra son las que mayor peso tienen en la explicación del fenómeno. El modelo en este caso explica en un 86.34% el valor de los inmuebles, de acuerdo con el coeficiente de determinación ajustado R^2 de 0.8634, dejando el 13.66% del valor a otras causas desconocidas, y siendo validado con un p-crítico del modelo muy inferior al nivel de significancia (alfa) de 0.05. De igual forma el resultado del estadístico F infiere que el modelo tiene una alta capacidad explicativa, dadas las variables utilizadas.

En el caso de la modelación que incluye, además de las variables definidas anteriormente, la distancia a pie al Centro Comercial Viva Envigado, se obtiene un coeficiente de determinación ajustado de 0.8633 que indica que las variables utilizadas explican en 86.33% el valor del inmueble, con una alta capacidad de explicación del valor de los inmuebles de acuerdo con el resultado obtenido del estadístico F de 50.46, y una p-crítica del modelo de 2,20E-16 muy inferior a 0.05. Se concluye entonces que al incluir la variable distancia a pie al Centro Comercial Viva Envigado en el modelo, la disminución en el coeficiente de determinación es del 0.01% y por lo tanto la variable distancia a pie

tiene una incidencia marginal en el modelo. Se observa que la variable distancia a pie al Centro Comercial Viva Envigado tiene una relación directa respecto al valor del inmueble, y acorde con el modelo, por cada metro que deba recorrer a pie un peatón desde el inmueble al Centro Comercial Viva Envigado, el valor del mismo se incrementa en \$18122, y por lo tanto, por cada 100m de distancia a pie al Centro Comercial Envigado hay un incremento en el valor de la vivienda multifamiliar de 0.4%.

Cuando en el modelo principal se incluye la variable distancia lineal entre el inmueble y el Centro Comercial Viva Envigado, se obtiene un coeficiente de determinación ajustado de 0.8608 que indica que las variables utilizadas explican en 86.08% el valor del inmueble, con una alta capacidad de explicación dado el valor obtenido del estadístico F de 49.46, y una P-crítica del modelo de $2,20E-16$ muy inferior a 0.05. Se concluye entonces que al incluir la variable distancia lineal al Centro Comercial Viva Envigado en el modelo, la disminución en el coeficiente de determinación es del 0.26% y por lo tanto la variable distancia lineal tiene una baja incidencia en la explicación dada por el modelo. En este caso la variable distancia lineal al Centro Comercial Viva Envigado tiene una relación directa respecto al valor del inmueble, y acorde con el modelo, por cada metro que esté alejado el inmueble del Centro Comercial Viva Envigado, medido linealmente, el valor del inmueble se incrementa en \$9824, lo que quiere decir que por cada 100m hay un incremento en el valor del mismo del 0.22%.

Para el caso de la distancia en carro desde el inmueble al Centro Comercial Viva Envigado, se observa que ésta tiene un comportamiento inverso respecto al valor del inmueble, lo que quiere decir que por cada metro adicional que se deba recorrer en carro hasta el Centro Comercial Viva Envigado, hay una disminución de \$11123 en el valor del inmueble, lo que quiere decir que por cada 100m de distancia hay una disminución del 0.25% en el valor del inmueble. En dicha modelación se obtiene un coeficiente de determinación ajustado de 0.8614 que indica que las variables utilizadas explican en 86.14% el valor del inmueble, con una alta capacidad de explicación dado el valor obtenido del estadístico F de 49.67, y una p-crítica del modelo de $2,20E-16$ muy inferior a 0.05. Se concluye entonces que al incluir la variable distancia en carro al Centro Comercial Viva Envigado en el modelo, la disminución en el coeficiente de determinación es del 0.2%, y

por lo tanto la variable distancia en carro al Centro Comercial Viva Envigado tiene una incidencia marginal en la explicación del modelo.

De acuerdo con los resultados obtenidos, la distancia a pie al Centro Comercial Viva Envigado es la que presenta una mayor incidencia en el valor del inmueble, comparada los modelos que incluyen la distancia lineal y la distancia en carro al mismo. Dicho modelo se presenta a continuación:

$$\text{Valor Inmueble} = -31890285 + 2219611 * \text{Área Construida} + 32762549 * \text{Ascensor} + 49117014 * \text{Cantidad Parqueaderos} + 18122 * \text{Distancia a pie} + 33235341 * \text{Estrato} - 20853125 * \text{Estado de Conservación}$$

Si bien las variables distancia al Centro Comercial Viva Envigado en todos los casos no mejoraron la explicación del valor del inmueble, en cuanto a que no aumentaron el resultado del coeficiente de determinación R² ajustado y el resultado del estadístico F, se puede observar en las modelaciones con todas las variables que éstas tienen una mayor incidencia al ser comparada con variables como cantidad de cuartos útiles, nivel de piso donde se encuentra ubicado el inmueble, si al inmueble le da el poniente o no, o el sistema de portería del edificio multifamiliar, al ser comparados a través del valor de p-crítica individual, hecho observado en la modelaciones paso a paso.

Una vez obtenidos los resultados y analizando el caso particular de Viva Envigado con respecto a otros centros comerciales del Área Metropolitana del Valle de Aburrá, podemos notar que Viva Envigado está rodeado por el polígono industrial de Las Vegas y cuenta con dos grandes barreras artificiales que son la Avenida Las Vegas y la Avenida Regional. Dichas barreras lo separan ampliamente de las zonas residenciales a su alrededor. Es posible que esta sea la razón para que la distancia tenga una incidencia marginal en el valor comercial de los inmuebles residenciales en multifamiliares, tema que podría ser analizado en estudios posteriores.

Esta situación no se da en otros centros comerciales del Área Metropolitana del Valle de Aburrá como lo son Unicentro (Barrio Conquistadores), Santafé (Barrio Balsos II), Oviedo (Sector Milla de Oro), San Diego (Barrio San Diego), Los Molinos (Barrio Belén Los Alpes y La Palma). En estos casos, al menos por uno de los lados de los Centros Comerciales existe contacto directo con la zona residencial.

Valdría la pena, entonces, estudiar otros casos de incidencia de la distancia a otros Centros Comerciales en los valores de los inmuebles residenciales que tengan estas características descritas.

Sería importante además, incluir en estudios posteriores, un análisis sobre las preferencias de los compradores en esta zona, toda vez que variables como la distancia en carro al Centro Comercial Viva Envigado, no son fácilmente perceptibles para un comprador, razón por la cual es posible que la distancia en carro esté presentando una incidencia inversa en el valor de los inmuebles residenciales. Las decisiones de los compradores no necesariamente son tan racionales como con este estudio.

15 REFERENCIAS

- Addae-Dapaah, K., & Lan, Y. S. (2010). Shopping Centres and the Price of Proximate Residential Properties. *Paper presented at PRRES2010*.
- Amaya, L. (2018). Criterio de Akaike para la selección de modelos con transformaciones. Bogotá: Universidad Santo Tomás.
- Appraisal Institute. (22 de Abril de 2009). *Appraisal Institute*. Recuperado el 22 de Septiembre de 2019, de Report Explores Appraisers' Role in Housing Bubble: <https://www.appraisalinstitute.org/ReportExploresAppraisersRoleinHousingBubble/>
- Aydin, R., Crawford, E., & Smith, B. (2011). Commercial development spillover effects upon residential values. *Southwestern Economic Review*.
- Aznar Bellver, J., & Guijarro Martínez, F. (2012). *Nuevos métodos de valoración. Valoración multicriterio* (Segunda ed.). Valencia: Editorial Universidad Politecnica de Valencia. Recuperado el 22 de Septiembre de 2019
- Bernal, C. (2010). *Metodología de la investigación*. Bogotá: Pearson.
- Borrero Ochoa, O. A. (2018). *Economía urbana y plusvalías del suelo*. Bogotá: Bhandar Editores Ltda.
- Castaño, J., Laverde, M., Morales, M. Á., & Yaruro, A. M. (Junio de 2013). Índice de Precios de la Vivienda Nueva para Bogotá: Metodología de Precios Hedónicos. *Temas de Estabilidad Financiera*(78). Obtenido de http://www.banrep.gov.co/docum/Lectura_finanzas/pdf/tef_78.pdf
- Colliers International . (2019). *Reporte Mercado Retail Medellín Colombia Abril- Junio 2019*. Bogotá: Colliers International.
- Colliers International. (2019). *Reporte Mercado Retail Medellín - Colombia*. Medellín: Colliers International.

- Dane. (2019). <https://www.dane.gov.co>. Recuperado el 2019, de <https://www.dane.gov.co/index.php/servicios-al-ciudadano/servicios-informacion/estratificacion-socioeconomica>
- Eaton, J. (14 de Abril de 2009). *The Center for Public Integrity*. Recuperado el 22 de Septiembre de 2019, de *The Appraisal Bubble*: <https://publicintegrity.org/environment/the-appraisal-bubble/>
- Franco Polanco, M. A. (2018). Comportamiento en el valor comercial de inmuebles Residenciales sometidos a regimen de propiedad horizontal en el sector cercano a unos centros comerciales y un almacen gran superficie en el municipio de Soacha, Cundinamarca.
- García, M. P. (2007). *Introducción a la valoración inmobiliaria*. Barcelona: Departament de Construccions Arquitectòniques I. ETSAB.
- Guevara, D. (10 de Febrero de 2008). *El Espectador*. *Diez años de la crisis de 2008: cuando la economía trabaja para las finanzas*. Recuperado el 22 de Septiembre de 2019, de <https://www.elespectador.com/economia/diez-anos-de-la-crisis-de-2008-cuando-la-economia-trabaja-para-las-finanzas-articulo-738343>
- La Nota Económica. (5 de Julio de 2019). *La Nota Económica*. Obtenido de <http://lanotaeconomica.com.co/negocios/la-evolucion-de-los-centros-comerciales.html>
- La República. (5 de Octubre de 2018). *El centro comercial Viva Envigado abre sus puertas con 394 locales y establecimientos*. Recuperado el 9 de Julio de 2019, de <https://www.larepublica.co/empresas/el-centro-comercial-viva-envigado-abre-sus-puertas-con-394-establecimientos-2778530>
- La República. (29 de Marzo de 2019). *Plaza Fabricato el nuevo centro comercial ya es una realidad en el Valle de Aburrá*. Obtenido de <https://www.larepublica.co/empresas/plaza-fabricato-es-una-realidad-2845128>
- Lever D., G. (s.f.). <http://www.ucipfg.com/Repositorio/MAES/PED/Semana4/PreciosHedonicos.pdf>.

Recuperado el 19 de Septiembre de 2019, de <http://www.ucipfg.com/Repositorio/MAES/PED/Semana4/PreciosHedonicos.pdf>.

Matthews, J. W. (Abril de 2007). Retail Proximity and Residential Values or Do Nearby Stores Really Run Down Property Values? *Journal of the American Planning Association – revise and resubmit*. Obtenido de <http://ssrn.com/abstract=989049>

Ministerio de Vivienda, Ciudad y Territorio. (s.f.). *Ministerio de Vivienda Ciudad y Territorio*. Obtenido de <http://www.minvivienda.gov.co/POTPresentacionesGuias/Componente%20Urbano.pdf>

Municipio de Envigado. (12 de abril de 2011). Acuerdo 010 2011. *POT Municipio de Envigado 2011-2023*. Envigado, Antioquia, Colombia: La Patria.

Municipio de Envigado. (2017). <https://www.envigado.gov.co>. Recuperado el 9 de Julio de 2019, de <https://www.envigado.gov.co>: <https://www.envigado.gov.co/planeacion/paginas/contenido/planes-zonales-de-planeacion-y-participacion-/planes-zonales>

Munoz-Raskin, R. (2009). Walking accessibility to bus rapid transit: Does it affect property values? The case of Bogotá, Colombia. *Transport Policy*, 72-84.

Muñoz Echavarría, J. (2016). *Los centros comerciales en la reconfiguración territorial de la ciudad de Medellín y los nuevos imaginarios Urbanos 1990 - 2011*. Universidad Nacional de Colombia, Medellín.

Osorio, J. A. (2008). *La burbuja inmobiliaria en Estados Unidos*. Banco Central de Reserva de El Salvador, Departamento de Investigación Económica y Financiera.

Ospinas & Cia S.A. (2017). Centros Comerciales Antioquia. *Centros Comerciales Guía 2017*, 206-207.

Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T., & French, N. (2018). Real estate appraisal: A review of valuation methods. *Journal of Property Investment & Finance*, 383 - 401. doi:<https://doi.org/10.1108/14635780310483656>

- Perdomo, J. (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.
- Perdomo, J. A. (2011). A methodological proposal to estimate changes of residential property value: case study developed in Bogotá. *Applied Economics Letters*. doi:10.1080/13504851.2011.554360
- Pope, D., & Pope, J. (2014). When Walmart Comes to Town: Always Low Housing Prices? Always?
- Quintana, L., & Mendoza, M. (2016). *Econometría Aplicada utilizando R*. México: Universidad Nacional Autónoma de México.
- Quintero García, L. K., & Rivera Grajales, L. J. (2012). Efecto de la distancia al acceso de transporte masivo megabús sobre el valor de la propiedad en la zona sub-centro del barrio cuba. *Grafías Disciplinarias de la UCP*, 73-78.
- Rey Carmona, F. J. (2014). *Alternativas y determinantes en valoración de inmuebles urbanos*. Universidad de Córdoba, Departamento de estadística, econometría, IO, organización de empresas y economía aplicada. Córdoba: Servicio de Publicaciones de la Universidad de Córdoba.
- Rodríguez, D., & Targa, F. (Septiembre de 2004). Value of accessibility to Bogotá's bus rapid transit system. *Transport Reviews: A Transnational Transdisciplinary Journal*, 24(5), 587-610. doi:10.1080/0144164042000195081
- Sale, M. (19 de Mayo de 2015). The Impact of a Shopping Centre on the Value of Adjacent Residential Properties. *Economic Research Southern Africa. Working paper 518*.
- Sandroni, P. (s.f.). *Renovación Urbana y Gentrificación: evaluación de algunos impactos demográficos, sociales y económicos de la operación urbana Faría Lima en la ciudad de San Pablo*.
- Semana. (13 de Octubre de 2018). <https://www.semana.com>. Recuperado el 24 de Octubre de 2019, de <https://www.semana.com>:

<https://www.semana.com/economia/articulo/centros-comerciales-en-colombia-si-hay-cliente-para-tantos/586746>

Suárez L., V. (10 de Junio de 2019). *America Retail*. Obtenido de <https://www.america-retail.com/colombia/colombia-esto-se-invierte-en-centros-comerciales-al-ano/>

Tekel, A., & Akbarishahabi, L. (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*.

Trojaneck, R., Gluszek, M., & Tanas, J. (2017). The effect of urban green spaces on house prices in Warsaw. *International Journal of Strategic Property Management*.

Vargas, F., & Caicedo, M. (2013). Efecto sobre el valor de la vivienda por su proximidad a un centro comercial en Bogotá.

Votsis, A. (Noviembre de 2016). Planning for green infrastructure: The spatial effects of parks, forests, and fields on Helsinki's apartment prices. *Ecological Economics*, 279–289.

Zhou, J., Zhang, L., Hui, E. C., & Wen, H. (2019). The effects of a shopping mall on housing prices: A Case study in Hangzhou. *International Journal of Strategic Property Management*, 23. doi:<https://doi.org/10.3846/ijspm.2019.6360>