Desarrollo de herramienta de web scraping para obtención y análisis espacial de datos del mercado inmobiliario: Caso de La Mesa, Cundinamarca

Karol Natalia Garcia Cotrino 20242094005

Jeyson Santiago Gomez Gomez 20242094006

Informe final de trabajo de grado presentado como requisito parcial para optar al título de Especialista en Sistema de Información Geográfica

Universidad de Distrital Francisco José de Caldas

Facultad de Ingeniería

Bogotá, 2025

Resumen

El presente trabajo desarrolla una herramienta basada en técnicas de web scraping para la recolección de datos del mercado inmobiliario, con el objetivo de generar insumos que apoyen el análisis espacial de las ofertas en el municipio de La Mesa, Cundinamarca. Esta iniciativa responde a la falta de actualización catastral en el municipio y a la necesidad de contar con información actualizada para la planificación territorial. Se aplicó un enfoque cuantitativo y exploratorio, implementando un flujo automatizado para la extracción, limpieza, normalización y análisis de datos obtenidos de portales inmobiliarios, específicamente de Finca Raíz. Los datos fueron procesados estadística y espacialmente mediante técnicas como la estimación de densidad de Kernel ponderada, empleando la métrica Haversine y algoritmos adecuados a la naturaleza geográfica de la información. Como resultado, se generó un mapa temático que representa la distribución espacial del valor de mercado de las ofertas, identificando zonas de mayor concentración y valorización. La herramienta propuesta se destaca por ser replicable, adaptable y útil para municipios con capacidades técnicas limitadas, al ofrecer una alternativa económica y eficiente para fortalecer la gestión territorial y el uso estratégico de datos abiertos.

Palabras clave: web scraping, análisis espacial, SIG, mercado inmobiliario, densidad de Kernel, La Mesa

Abstract

This project develops a tool based on web scraping techniques to collect real estate market data, aiming to generate inputs that support the spatial analysis of property listings in the municipality of La Mesa, Cundinamarca. The initiative addresses the lack of cadastral updates in the municipality and the need for current information to improve territorial planning. A quantitative and exploratory approach was applied, implementing an automated workflow for extracting, cleaning, normalizing, and analyzing data obtained from real estate websites, specifically from Finca Raíz. The data were processed using statistical and spatial techniques such as weighted Kernel density estimation, employing the Haversine metric and algorithms suitable for geographic information. As a result, a thematic map was generated to represent the spatial distribution of market property values, identifying areas with higher concentration and appreciation. The proposed tool stands out for being replicable, adaptable, and useful for municipalities with limited technical capacities, offering an efficient and low-cost alternative to strengthen territorial management and the strategic use of open data.

Keywords: web scraping, land value, spatial analysis, GIS, real estate market, Kernel density, La Mesa

Contenido

1. PLANTEAMIENTO DEL PROBLEMA DE INVESTIGACIÓN Y SU JUSTIFICACIÓN 7

1.1 DESCRIPCIÓN DEL ÁREA PROBLEMÁTICA 8
1.2 FORMULACIÓN DEL PROBLEMA 9
1.3 JUSTIFICACIÓN 10

2. OBJETIVOS 12

2.1 OBJETIVO GENERAL 12
2.2 OBJETIVOS ESPECÍFICOS 12

3. ANTECEDENTES 13

4. REFERENTE NORMATIVO Y LEGAL 17

5. REFERENTE TEÓRICO 19

6. METODOLOGÍA 21

6.1 ENFOQUE METODOLÓGICO 22
6.2 TIPO DE ESTUDIO 23
6.3 PROCEDIMIENTO 23

7. RESULTADOS 24

8. CONCLUSIONES 34

9. RECOMENDACIONES 36

10. REFERENCIAS 37

Lista de figuras

Figura 1 Mapa municipal de La Mesa Cundinamarca Fuente: Alcaldía de La Mesa.....21
Figura 2 Mapa político administrativo urbano del municipio de La Mesa Cundinamarca Fuente: Alcaldía de La Mesa.....22
Figura 3 Flujo de trabajo metodológico.....24
Figura 4 Registros extraídos de páginas inmobiliarias a partir del web scraping.....25
Figura 5 Resultado de limpieza y normalización de los datos.....26
Figura 6 Mapeo de ofertas inmobiliarias según su valor.....27
Figura 7 Caja de bigotes para identificación de valores atípicos.....27
Figura 8 Resumen estadístico del conjunto de datos.....29
Figura 9 Estadísticas Cuantiles y Estadísticas Descriptivas.....30
Figura 10 Histograma.....30
Figura 11 QQ plot.....31
Figura 12 Mapa de distribución Kernel.....33

Lista de tablas

Tabla 1 Atributos de interés 25
Tabla 2 Valores inmobiliarios atípicos 28

1. Planteamiento del problema de investigación y su justificación

En este capítulo se presenta el contexto general y específico que motiva la realización de la investigación. Se describe la situación actual del municipio de La Mesa, Cundinamarca en relación con su desactualización catastral, así como la necesidad de contar con insumos técnicos confiables y actualizados para la adecuada gestión del territorio. Se formula el problema de investigación con base en la identificación de vacíos técnicos e institucionales, y se plantea una pregunta orientadora que guía el estudio. Finalmente, se justifica la pertinencia de la investigación, destacando su valor como insumo para el análisis espacial de ofertas inmobiliarias.

1.1 Descripción del área problemática

El municipio de La Mesa Cundinamarca tuvo su ultima actualización catastral en el año 2015 elaborado por la Agencia Catastral de Cundinamarca su actual gestor catastral mediante un operador independiente. A partir de los datos abiertos del municipio se puede llegar a la conclusión que de la información desarrollada en este ejercicio no ha sido empleada para el beneficio de la entidad territorial, como el diagnóstico y formulación de su Plan Básico de Ordenamiento Territorial- PBOT (Concejo Municipal de La Mesa Cundinamarca, 2005) el cual su última modificación fue aprobada en el año 2005 y en su totalidad fue elaborado en el año 2000 (Concejo municipal La Mesa Cundinamarca, 2000); teniendo un vencimiento de 8 años según la ley 388 de 1997 la cual estipula que cada municipio tiene un máximo de 12 años de vigencia para su reformulación.

Para comprender la realidad territorial de un municipio, es imprescindible contar con insumos actualizados. Sin datos precisos, las decisiones fiscales y políticas pueden resultar equivocadas. En municipios donde el turismo es un motor clave —por su clima favorable y cercanía a la capital—, la dinámica inmobiliaria adquiere aún mayor relevancia. En estos lugares, conocer el valor del suelo urbano es esencial para una administración eficaz, pues, aunque el área urbana representa una porción reducida del territorio, concentra la mayor oferta y demanda inmobiliaria.

1.2 Formulación del problema

El municipio de La Mesa, ubicado en el departamento de Cundinamarca, forma parte del 73% de municipios en Colombia que no cuentan con una actualización catastral para 2025 (Trejos, 2025). Esta situación obstaculiza procesos fundamentales, como la actualización del impuesto predial y el enfoque multipropósito que el país busca implementar, lo cual es esencial para la formulación de políticas públicas en diversos sectores; sin embargo, la baja prioridad política otorgada a esta problemática está retrasando el desarrollo del municipio.

Además, el debilitamiento del componente técnico en las entidades territoriales —debido a la escasa capacitación del personal en tecnologías de información geográfica— ha generado la carencia de herramientas automatizadas para la recolección, integración y análisis de datos inmobiliarios. También existe desconocimiento sobre el uso potencial de datos abiertos del mercado inmobiliario, sumado a una creciente desconfianza institucional en la calidad y veracidad de dichos datos.

Como consecuencia, las autoridades locales enfrentan serias dificultades para tomar decisiones informadas en planificación territorial. Esto conduce a valoraciones del suelo inconsistentes y no representativas, lo que a su vez genera una cadena de efectos negativos como: expansión urbana desordenada, conflictos por el uso del suelo, débil regulación del mercado inmobiliario, desigualdad en la asignación de subsidios públicos y una mayor brecha tecnológica e institucional en la gestión del territorio

¿Cómo puede una herramienta que utiliza la técnica de web scraping para recolectar datos de portales inmobiliarios y contribuir en la generación de insumos referenciales para analizar el comportamiento de ofertas en el municipio de la Mesa Cundinamarca?

1.3 Justificación

El municipio de La Mesa, Cundinamarca, enfrenta un rezago significativo en su actualización catastral, cuya última versión data del año 2015. Esta situación limita de forma sustancial la capacidad institucional para gestionar eficazmente el territorio, formular políticas públicas actualizadas y aplicar una tributación equitativa. A ello se suma que su Plan Básico de Ordenamiento Territorial (PBOT) fue formulado en el año 2000 y actualizado en 2005, superando ampliamente el periodo de vigencia establecido por la Ley 388 de 1997, que estipula una reformulación cada 12 años. Esta obsolescencia normativa y técnica compromete la pertinencia de las decisiones de planificación y ordenamiento territorial.

Frente a este escenario, la presente investigación propone el desarrollo de una herramienta para la recolección de información de ofertas inmobiliarias fundamentada en el uso de la técnica de web scraping el cual permite obtener insumos para la realización de análisis geoestadísticos y de estadística espacial. Este enfoque cuantitativo permitió la recopilación, integración y procesamiento de datos numéricos del mercado inmobiliario publicados en plataformas digitales, los cuales fueron estructurados y analizados mediante procesos autónomos y personalizados. Este flujo de trabajo —desde la captura de los datos hasta la visualización cartográfica— evidencia una alternativa viable para suplir las limitaciones actuales en las capacidades técnicas del municipio, sin depender exclusivamente de grandes inversiones institucionales.

Lo más relevante de esta propuesta es la incorporación de información proveniente de fuentes secundarias que, aunque comúnmente no se consideran en procesos oficiales, resultan ser confiables, actualizadas y de rápida disponibilidad. A partir de esta información, se propone realizar una estimación de densidad de Kernel ponderada. Esta técnica, al centrarse inicialmente en una única variable (el valor de las ofertas), permitirá validar la utilidad de los datos recolectados.

Además, servirá como punto de partida para transitar de un enfoque meramente exploratorio hacia análisis más robustos de tipo correlacional o explicativo, integrando progresivamente variables adicionales.

Esta iniciativa no solo mejora la gestión fiscal y la planificación territorial, sino que también fomenta el uso estratégico de datos abiertos, fortalece la confianza institucional en el uso de información digital y contribuye al cierre de brechas tecnológicas en los gobiernos locales. Esta investigación responde a una necesidad estructural de información actualizada, a la vez que aporta una solución técnica viable, replicable y en sintonía con los principios del catastro multipropósito promovido a nivel nacional.

2. Objetivos

2.1 Objetivo general

Desarrollar una herramienta que utilice la técnica de web scraping para recolectar datos del mercado inmobiliario, con el fin de generar insumos referenciales que permitan analizar el comportamiento espacial de las ofertas en el municipio de La Mesa, Cundinamarca

2.2 Objetivos específicos

- Aplicar técnica de web scraping para la recopilación de ofertas inmobiliarias del municipio La mesa Cundinamarca teniendo en cuenta variables Tipo de inmueble, Descripción, Latitud, Longitud, Valor, Estrato, Fecha en la que se publica la oferta, Ultima fecha en la que se actualiza la oferta, Area Construida, Area privada, Vetustez, N° Habitaciones, N° Baños, N° Parqueaderos y Enlace_fuente.
- Limpiar y normalizar los datos de ofertas inmobiliarios capturados
- Proponer una herramienta para la creación de un mapa de partir de información recopilada, con la técnica de densidad de Kernel ponderada sobre la variable Valor.

3. Antecedentes

Tuta y Vela (2018) realizaron un estudio cuyo objetivo fue generar una superficie de valor del suelo rural en el municipio de Chía, a partir de ofertas del mercado georreferenciadas para identificar la correlación entre localización y valor del suelo, concluyendo que el análisis espacial permite evidenciar diferencias significativas en la dinámica inmobiliaria del territorio.

Salgado, Rivera y Bustos (2024) desarrollaron un modelo basado en zonas homogéneas físicas y geoeconómicas para calcular el valor unitario del suelo en Usaquén, Bogotá, integrando un aplicativo web geográfico que facilita la consulta pública, lo cual demuestra que la visualización espacial es clave para transmitir los resultados de análisis territoriales de manera didáctica.

Carranza, Piumetto y Lucca (2022) propusieron una metodología de avalúo masivo utilizando exclusivamente datos abiertos, como OpenStreetMap y modelos de aprendizaje automático, lo que evidenció la viabilidad de aplicar tecnologías accesibles en territorios con limitada capacidad fiscal, destacando la utilidad de estas fuentes para alimentar modelos espaciales automatizados.

Rojas y Ruiz (2020) describieron un proceso de avalúo comercial de inmuebles con herramientas digitales como Google Earth Pro, permitiendo integrar coordenadas, valores de mercado y normativa vigente, lo cual refleja el valor del uso de plataformas abiertas para análisis espaciales normativos.

Jiang y Ma (2025) diseñaron un modelo predictivo de vitalidad urbana mediante aprendizaje profundo, considerando variables como morfología urbana, flujos peatonales y densidad, lo que demuestra que la interpretación territorial puede ser enriquecida mediante datos complementarios tomados de fuentes dinámicas.

Liu y Tan (2024) estudiaron los patrones espacio-temporales de vitalidad urbana en Chengdu mediante modelos de regresión espacial y ponderada geográficamente, encontrando que variables como el NDVI, la densidad vial y la distancia al transporte público afectan la vitalidad, lo que reafirma que las condiciones urbanas pueden ser modeladas espacialmente con un enfoque predictivo.

Chaves y Ortiz (2025) implementaron un sistema automatizado para estimar mapas de valor de suelo urbano con imágenes satelitales y catastro, validando el modelo con predicciones ajustadas que permiten construir capas de información espacial interpretables.

Pérez y Mendoza (2025) aplicaron modelos tridimensionales en ArcGIS y visualización con Plotly para estudiar el uso del espacio urbano en Sídney, utilizando simulación 3D y autómatas celulares para representar dinámicas futuras, lo que demuestra cómo la representación espacial también puede ser volumétrica y dinámica.

Ramírez y Herrera (2025) analizaron estrategias de enfriamiento urbano con base en índices de temperatura y humedad en Wuhan, estructurando zonas de regulación térmica a partir de mapas que integran datos espaciales de cobertura vegetal, densidad edificatoria y uso industrial, lo cual valida el uso de capas de información espacial para planificar intervenciones urbanas.

Franco y Medina (2023) estudiaron los valores sociales del ecosistema en Kuala Lumpur mediante encuestas participativas y herramientas SIG, aplicando el modelo SolVES para representar espacialmente las preferencias ciudadanas respecto al uso del suelo, evidenciando que la percepción social puede ser georreferenciada y visualizada en mapas de conflicto o consenso territorial.

Castro y González (2025) propusieron un modelo de valoración del suelo mediante catastro armonizado con planificación territorial, integrando datos UAV y SIG para corregir la geometría de mapas regulatorios, lo que evidenció que una base cartográfica confiable es esencial para lograr coherencia en los análisis espaciales.

Díaz y Vargas (2025) desarrollaron un estudio bibliométrico sobre el desarrollo orientado al transporte, identificando las tipologías más investigadas y los países con mayor producción, demostrando que el análisis espacial en torno a la movilidad ha generado metodologías aplicables en contextos urbanos diversos.

Rincón y Zapata (2024) exploraron la relación entre vecindario, desigualdad y urbanismo en Canadá desde una perspectiva histórica, reconociendo cómo las características físicas y sociales del entorno inciden en oportunidades y acceso a la propiedad, lo que puede integrarse a modelos espaciales.

Sánchez y Castaño (2024) analizaron datos multisectoriales mediante modelos predictivos de aprendizaje automático para evaluar crecimiento urbano y sostenibilidad, aplicando algoritmos de regresión, clustering y redes neuronales sobre imágenes satelitales, censos y sensores urbanos, confirmando la capacidad de estas técnicas para representar fenómenos territoriales mediante mapas de calor, distribución o demanda.

Por último, Méndez y Duarte (2025) presentaron un modelo de predicción de vitalidad urbana en Nueva York con arquitectura UVPN, integrando elementos urbanos a distintas escalas como calles, densidad, red vial y puntos de interés, lo cual generó mapas detallados sobre actividad humana y uso del suelo, demostrando que la combinación de aprendizaje profundo y análisis espacial ofrece resultados precisos y útiles para la planificación.

En conjunto, estas investigaciones confirman que, independientemente de la fuente de datos (ya sean satelitales, censales, extraídas por scraping, recolectadas mediante encuestas, catastro o datos abiertos), el punto de convergencia metodológico es el uso de tecnologías emergentes para la obtención indirecta de datos, su posterior procesamiento en herramientas de análisis espacial permiten representar los resultados mediante servicios reproducibles, fundamentados y de fácil transmisión a la comunidad en general.

4. Referente normativo y legal

Ordenamiento territorial y uso del suelo:

- **Ley 388 de 1997** – *Ley de Ordenamiento Territorial*

  Esta es la norma fundamental que regula el ordenamiento territorial en Colombia. Establece que los municipios deben formular planes de ordenamiento territorial (POT) y determinar instrumentos como los avalúos, las zonas de uso del suelo y las estrategias de gestión del suelo. Reconoce explícitamente el **valor del suelo como elemento clave para la planificación**, y promueve instrumentos de **captura de valor** para financiar infraestructura urbana.

- **Ley 1454 de 2011** – *Ley Orgánica de Ordenamiento Territorial (LOOT)*

  Fue modificada y adicionada por la **Ley** 1962 de 2019, y establece el marco para la coordinación de la planificación territorial entre los niveles municipal, departamental y nacional. Refuerza el papel del ordenamiento del suelo como componente estratégico del desarrollo regional.

- **Decreto 1077 de 2015** – *Decreto Único Reglamentario del Sector Vivienda, Ciudad y Territorio*

  Compila normas reglamentarias sobre ordenamiento territorial, incluyendo disposiciones sobre planes parciales, avalúo del suelo, instrumentos de gestión y financiamiento del desarrollo urbano.

Catastro multipropósito y valoración del suelo:

- **Ley 1955 de 2019 (PND 2018–2022)** – *Plan Nacional de Desarrollo*

  Introduce el **Catastro Multipropósito** como política nacional, buscando actualizar y mejorar el catastro con enfoque técnico, jurídico y fiscal. Promueve el uso de tecnologías

abiertas, interoperabilidad y enfoque territorial, lo cual es directamente aplicable a métodos como el *web scraping* para levantar datos del suelo.

- **Resolución 662 de 2024** *–Por la cual se adopta el Modelo Extendido Catastro - Registro a partir del Modelo Núcleo LADM_COL para el desarrollo de la gestión catastral con enfoque multipropósito en el marco del Sistema de Administración del Territorio (SAT) y se modifica la Instancia de Coordinación para la gobernanza y la aprobación de nuevas versiones de este Modelo*

  Define los atributos y estructuras de datos requeridos para el catastro actualizado, incluyendo el **valor del suelo** como componente obligatorio. Apoya el uso de fuentes secundarias, imágenes satelitales, SIG y tecnologías abiertas.

Uso de tecnologías y datos abiertos:

- **Ley 1712 de 2014** – *Ley de Transparencia y del Derecho de Acceso a la Información Pública Nacional*

  Establece el derecho de acceso a la información pública, incluyendo bases de datos, lo cual respalda legalmente el uso de **datos abiertos para investigación** y desarrollo de productos basados en esta información.

- **Ley 1581 de 2012** – *Ley de Protección de Datos Personales*

  Regula el tratamiento de datos personales en Colombia. Si bien el *web scraping* puede implicar acceso a información pública, esta ley obliga a garantizar que no se recolecten ni usen datos personales sin autorización expresa, especialmente si provienen de fuentes privadas.

5. Referente teórico

El valor del suelo es una categoría fundamental en los estudios de ordenamiento territorial, planificación urbana y gestión del desarrollo sostenible, al constituirse como un indicador del comportamiento del mercado inmobiliario y de las dinámicas socioespaciales asociadas a la localización. Su definición parte del reconocimiento de que el suelo no es un bien homogéneo, sino un recurso finito cuyo valor se determina por múltiples factores, tanto físicos como sociales, jurídicos, económicos y ambientales. En términos generales, el valor del suelo se entiende como la estimación monetaria del terreno sin edificaciones, influida por variables como la accesibilidad, los usos del suelo permitidos, la oferta y demanda en el mercado, la infraestructura disponible, la normatividad urbanística, y la percepción de seguridad y calidad ambiental.

Desde una perspectiva económica, el valor del suelo está estrechamente relacionado con la renta del suelo y los principios de localización de Von Thünen y Alonso, quienes destacaron la importancia de la cercanía a centros de consumo y transporte como determinantes del precio. A nivel urbano, la teoría de los lugares centrales y los modelos de uso del suelo, como el de Burgess y Hoyt, han permitido comprender cómo la estructura espacial de las ciudades incide en la valorización del terreno. En consecuencia, los terrenos más próximos a nodos de conectividad y a centros de actividad económica tienden a registrar mayor valor, mientras que zonas periféricas o de difícil acceso presentan valores más bajos.

Por su parte, la valoración catastral y comercial del suelo obedece a metodologías específicas que integran criterios normativos y técnicos, los cuales buscan reflejar el valor real de mercado o el valor fiscal para fines tributarios y de ordenamiento. El avalúo masivo ha cobrado importancia en América Latina como una herramienta para actualizar catastros y mejorar la equidad fiscal, siendo el valor del suelo una de sus principales variables de análisis. No obstante, muchas

administraciones locales enfrentan limitaciones técnicas y financieras para implementar sistemas de valoración actualizados, lo cual genera rezagos significativos entre el valor registrado y el valor real del mercado.

En ese contexto, el uso de tecnologías como los sistemas de información geográfica (SIG), la teledetección, el aprendizaje automático y recientemente el web scraping, han abierto nuevas posibilidades para la estimación del valor del suelo mediante modelos automatizados y el uso de datos abiertos o recolectados en línea. El web scraping, en particular, permite extraer datos de portales inmobiliarios, catálogos de oferta y demanda, y plataformas digitales, aportando una fuente alternativa y actualizada de información sobre precios, ubicación y características del suelo. Estas herramientas, combinadas con modelos de predicción espacial, posibilitan la generación de estimaciones del valor del suelo a partir de información heterogénea y en tiempo casi real.

En municipios intermedios como La Mesa, Cundinamarca, donde convergen procesos de expansión urbana, presión sobre el recurso suelo y carencias en sistemas de información territorial, la estimación precisa y actualizada una estimación del valor del suelo a partir de las dinámicas inmobiliarias representa una necesidad estratégica para una gestión del territorio más eficiente, equitativa y sostenible.

6. Metodología

7.1. Zona de estudio

La investigación fue realizada en la cabecera municipal del municipio de La Mesa ubicado en el Departamento de Cundinamarca en las coordenadas 4°37′49″N 74°27′45″O a una altura de 1200 m.s.n.m, el cual tiene como límite al noroeste y oeste el municipio Quipile, al norte el municipio Cachipay, al noreste por Zipacón, al este y sureste el municipio Tena, al sur el municipio El Colegio y al suroeste con el municipio Anapoima.

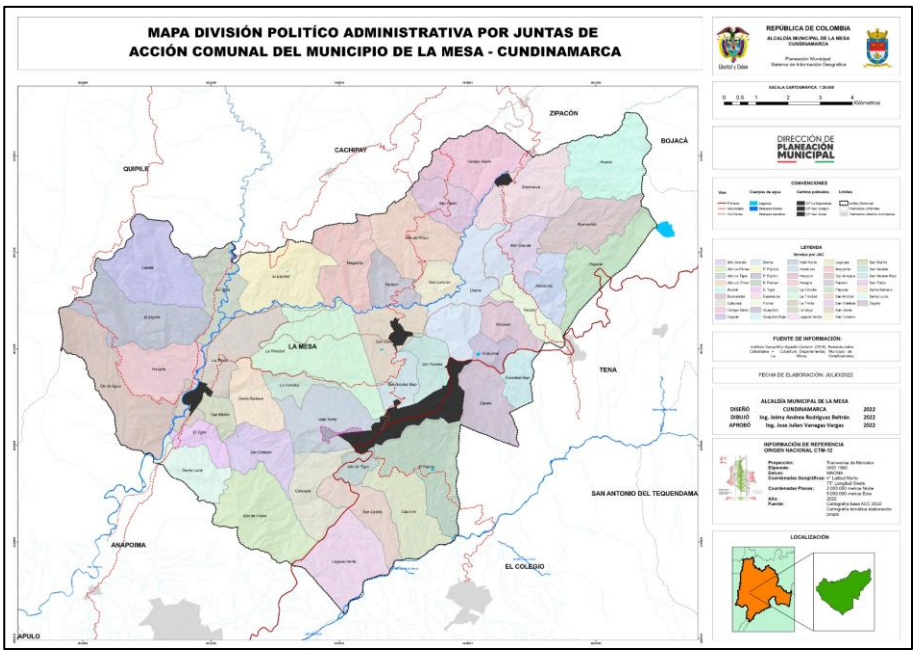


Figura 1 Mapa municipal de La Mesa Cundinamarca Fuente: Alcaldía de La Mesa

El casco urbano cuenta con un área de 2.75 km2 (Martínez, 2020), el municipio cuenta con una precipitación anual de 1000 a 1500 mm y una temperatura media anual que se encuentra entre los 18 y 24 °C (Palacios, Torres, Villarraga, & Meza, 2017).

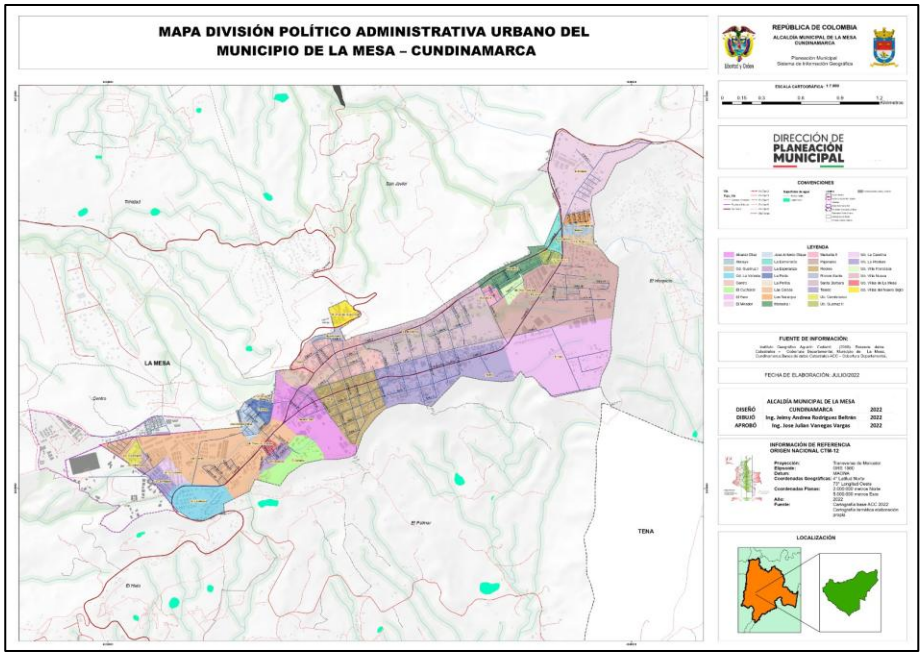


Figura 2 Mapa político administrativo urbano del municipio de La Mesa Cundinamarca Fuente: Alcaldía de La Mesa

6.1 Enfoque metodológico

La presente investigación adopta un enfoque cuantitativo, dado que su objetivo es desarrollar una herramienta basada en la técnica de web scraping para la recolección de ofertas inmobiliarias del municipio de La Mesa, Cundinamarca, publicadas en portales web, con el fin de generar insumos referenciales útiles para el análisis espacial de las ofertas.

Este enfoque fue elegido por su capacidad para recopilar y analizar información cuantificable, extraída de portales web de bienes raíces, de forma sistemática, objetiva y replicable. En este sentido, se prioriza el uso de datos numéricos, tales como el valor de las ofertas, ubicación geográfica, estrato y otras variables asociadas, las cuales permiten caracterizar el comportamiento de las ofertas inmobiliarias

La investigación se estructuró bajo un proceso secuencial, deductivo y probatorio, característico del enfoque cuantitativo. Se diseñó y ejecutó un flujo automatizado de recolección de datos mediante web scraping, seguido de su estructuración, análisis y representación espacial. Esta

estrategia metodológica permitió transformar datos no estructurados en información georreferenciada lista para el análisis geoestadístico y estadístico espacial, para el informe

6.2 Tipo de estudio

La investigación se concibió como un estudio de tipo exploratorio, dado que plantea el desarrollo de una herramienta para la recolección de ofertas inmobiliarias obtenidas mediante la técnica de web scraping. Esta técnica, aunque en crecimiento, aún no ha sido ampliamente aplicada en el análisis espacial, donde tradicionalmente se emplean otras fuentes de información.

Asimismo, el estudio presenta un componente descriptivo, ya que se tiene como propósito realizar una prueba geoestadística mediante una estimación de densidad de Kernel ponderado para ver la empleabilidad de los datos recopilados.

6.3 Procedimiento

El procedimiento inicio con la implementación de la técnica de web scraping, mediante la cual se recopilaron ofertas inmobiliarias disponibles en el portal web Finca Raíz donde se extrajeron los atributos de interés los cuales fueron los siguientes: tipo de inmueble, descripción, latitud, longitud, valor , estrato, fecha en que se publica la oferta, última fecha en que se actualiza la oferta, Area construida, Area privada, Vetustez, N° habitaciones, N° baños, N° parqueaderos y el enlace fuente.

Posteriormente, los datos recopilados fueron almacenados en un archivo Excel para su limpieza y normalización, después de ello se realizó un análisis estadístico exploratorio de la variable Valor, con la finalidad de observar el comportamiento general e identificar factores como la presencia de valores atípicos.

Una vez depurados, los datos fueron georreferenciados, dando paso al análisis espacial mediante la técnica de estimación de densidad de Kernel ponderada, utilizando de variable de peso

el atributo Valor. Finalmente, el resultado de la estimación se ajustó al polígono de la cabecera municipal de La Mesa, Cundinamarca, obteniendo así, un mapa de densidad espacial del valor de las ofertas inmobiliarias. El procedimiento se ilustra en el esquema Figura 3.

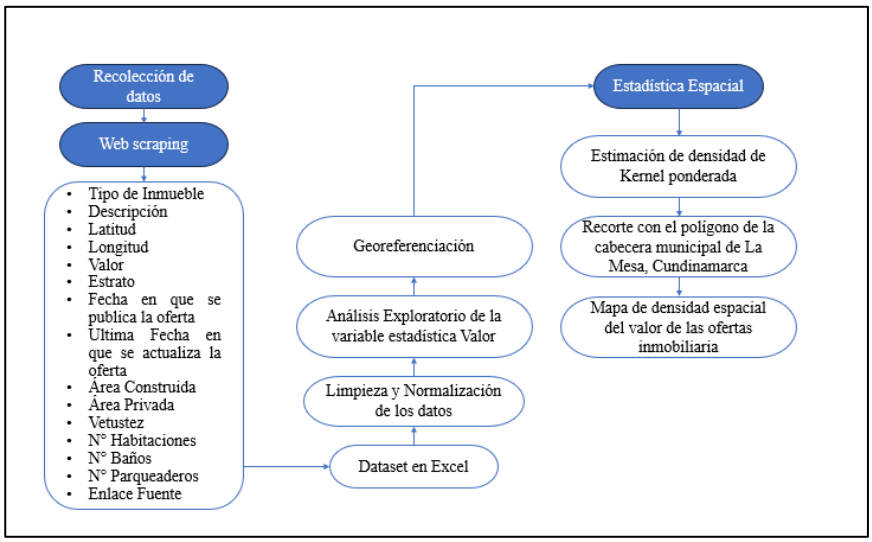


Figura 3 Flujo de trabajo metodológico
Fuente: Elaboración propia

7. Resultados

Mediante la aplicación de la técnica web scraping, utilizando el lenguaje de programación Python y la librería Selenium sobre el portal web Finca Raíz para inmuebles en venta en el municipio de La Mesa, el script desarrollado extrajo en total 365 de las 366 ofertas que se encontraban en el Portal Web, por lo que se tiene un porcentaje de recolección cerca del 99.7%.

La herramienta generó una salida de los datos en un dataset en Excel con las ofertas Inmobiliarias, cada registro incluye los atributos de interés los cuales se ilustran en la tabla 1.

Atributos
Tipo de Inmueble
Descripción
Latitud
Longitud
Valor
Estrato
Fecha en que se publica la oferta
Última fecha en que se actualiza la oferta
Área Construida
Área Privada
Vetustez
N° Habitaciones
N° Baños
N° Parqueaderos
Enlace Fuente

Tabla 1 Atributos de interés

Por otro lado, en la Figura 4 se presentan algunos registros extraídos, lo que permite visualizar cómo se encuentra estructurada la información dentro del archivo Excel.

Tipo de Inmueble:	Descripción:	Dirección:	Latitud:	Longitud:	Precio:	Estrato:	URL de la Imagen:	Fecha en que se tomo la oferta:	Fecha en que se publica la oferta:	Ultima Fecha en que se actualiza la oferta:	Area construida	Area privada	Vetustez	N° Habitaciones	N° Baños	N° Parqueaderos
Lote	invierte en	Vereda san	4.603551	-74.4808	5.8E+08	2	https://cdr	########	2025-01-14	2025-05-05	5,130.00 n	5,130.00 m2				
Casa	¡casa en ve	La Mesa, C	4.632147	-74.463	5.35E+08	4	https://cdr	########	2025-05-12	2025-05-12	140.00 m2	130.00 m2		3	4	4
Lote	se vende lo	La Mesa, C	4.632147	-74.463	1.5E+08	3	https://d3s	########	2024-04-12	2025-03-13	500.00 m2	500.00 m2	menor a 1 año			
Apartamen	¡vive por lo	Conjunto R	4.625902	-74.4717	7.5E+08	4	https://cdr	########	2025-04-20	2025-04-24	200.00 m2	158.00 m2	1 a 8 años	3	3	1
Finca	disfruta y r	CASA DOIN	4.671864	-74.4372	9E+08	2	https://cdr	########	2025-01-17	2025-03-17	280.00 m2	13,890.00	1 a 8 años	6	4	11
Casa	vive la tran	La Mesa, C	4.632147	-74.463	6.5E+08	4	https://cdr	########	2025-03-06	2025-03-06	1,500.00 n	1,500.00 n	1 a 8 años	4	4	4
Apartamen	adquiere u	La Mesa, La	4.632869	-74.4635	2.65E+08	3	https://cdr	########	2025-02-21	2025-04-03	75.00 m2	70.00 m2	1 a 8 años	3	2	1
Apartamen	? ¡vive con	La Mesa, C	4.632147	-74.463	2.7E+08	4	https://cdr	########	2025-04-23	2025-04-30	67.30 m2	60.00 m2	1 a 8 años	2	2	1
Finca	venta finca	Parque Ma	4.648719	-74.4407	6.4E+08	3	https://cdr	########	2025-02-25	2025-03-05	7,880.00 n	7,880.00 m2				
Casa	en venta ca	Vereda La T	4.640889	-74.4789	8.5E+08	3	https://cdr	########	2025-03-27	2025-05-15	250.00 m2	1,726.00 n	1 a 8 años	3	4	3
Apartamen	"¡apartam	La Mesa cu	4.625328	-74.4755	2.15E+08	3	https://cdr	########	2025-05-08	2025-05-12	49.00 m2	49.00 m2	9 a 15 años	3	2	1
Apartamen	se vende h	La Mesa, C	4.632147	-74.463	2.55E+08	3	https://cdr	########	2025-03-25	2025-03-31	80.00 m2	74.00 m2	1 a 8 años	3	2	1

Figura 4 Registros extraídos de páginas inmobiliarias a partir del web scraping

Al analizar algunos de los registros extraídos, se evidenció una serie de inconsistencias en los datos, lo cual dificultó el uso y análisis de los mismos. Por esta razón, fue necesario realizar un proceso de limpieza y normalización mediante expresiones regulares y con el uso de las librerías Pandas y Geopandas.

Como resultado de este proceso se pasó de tener un dataset conformado por 365 ofertas inmobiliarias a solamente 72, las cuales cumplen con las condiciones necesarias para el uso de la información en la zona de estudio.

Tipo de inmueble:	Descripción:	Dirección:	Latitud:	Longitud:	Precio:	Estrato:	URL de la imagen:	Fecha en que se tomo la oferta:	Fecha en que se publica la oferta:	Ultima Fecha en que se actualiza la oferta:	Area construida	Area privada	Vetustez	N° Habitaciones	N° Baños	N° Parqueaderos	Enlace_fuente
Casa	con un rea c	KR 19 # 5 - 1	4.631028	-74.461588	#########	3	https://d3s5	#########	2024-05-05	2024-08-15	1100	1100	4.5	11	10	5	https://www
Finca	finca compu	Cra. 21a 9-3	4.6336857	-74.464924	#########	6	https://cdn2	#########	2023-09-19	2025-04-23	169	169	30	3	1	0	https://www
Casa	hermosa cas	Calle 7 #20-2	4.6318857	-74.462833	#########	5	https://cdn2	#########	2025-05-20	2025-05-20	650	650	4.5	5	8	5	https://www
Casa	casa en vent	Calle 5 #4 -7	4.62932	-74.46625	#########	4	https://cdn2	#########	2025-05-27	2025-05-27	388	388	12	5	5	4	https://www
Casa	descubre est	Carrera 1 #1	4.6439495	-74.445905	#########	4	https://cdn2	#########	2025-03-25	2025-05-05	219	170	1	3	4	0	https://www
Finca	placentera y	-	4.6231059	-74.471202	#########	4	https://cdn2	#########	2024-09-06	2025-01-04	340	340	23	4	4	6	https://www
Local	gran oportur	Cl. 8 #13-50,	4.6345106	-74.457333	#########	3	https://cdn2	#########	2025-02-25	2025-02-26	325	325	3	0	0	0	https://www
Casa	imperesiona	CL 5 # 16 - 7(	4.6309839	-74.45954	#########	3	https://cdn2	#########	2024-12-09	2025-04-29	196	196	4.5	5	3	5	https://www
Casa	hermosa cas	Carrera 4 #4	4.6334498	-74.449294	#########	4	https://cdn2	#########	2024-10-24	2024-10-24	400	510	1	5	7	3	https://www
Casa	vendo excele	CALLE 7 6- 1	4.6402921	-74.446357	#########	3	https://cdn2	#########	2025-04-18	2025-04-18	190	600	30	4	3	4	https://www
Finca	finca ubicad	KR 28 # 5 - 2	4.6238861	-74.470768	#########	4	https://d3s5	#########	2024-04-21	2025-04-04	200	200	12	3	3	7	https://www
Casa	? area 500 m	Calle 3 # 2-3	4.6315425	-74.465489	#########	3	https://cdn2	#########	2025-02-23	2025-05-30	500	500	4.5	3	3	4	https://www
Casa	? ¡descubre t	El Hato, La M	4.6230357	-74.475365	#########	4	https://cdn2	#########	2025-02-14	2025-02-14	127	127	12	3	3	4	https://www
Apartamento	¡vive por lo a	Conjunto Re	4.625902	-74.471736	#########	4	https://cdn2	#########	2025-04-20	2025-04-24	200	158	4.5	3	3	1	https://www
Casa	? en venta: c	Pomarrosos	4.6431536	-74.448194	#########	4	https://cdn2	#########	2025-04-26	2025-05-04	210	210	2	4	3	3	https://www
Casa	¡t refugio id	Calle 8 #24 J	4.629741	-74.4674	#########	2	https://cdn2	#########	2025-04-29	2025-05-22	227	227	12	4	5	1	https://www
Casa	c digo 1460	Calle 8 # 2	4.6327388	-74.458439	#########	4	https://cdn2	#########	2025-04-11	2025-05-12	149	125	4.5	3	3	4	https://www
Casa	no sabes dor	CL 4 # 14 - 3(	4.6307782	-74.457178	#########	3	https://cdn2	#########	2024-12-30	2025-05-21	312	312	12	8	6	1	https://www
Apartamento	gran oportur	AVENIDA 3	4.6262151	-74.472107	#########	3	https://cdn2	#########	2024-08-30	2025-04-24	128	128	12	4	3	0	https://www

Figura 5 Resultado de limpieza y normalización de los datos

En la Figura 5 se visualiza la implementación de manera efectiva de las actividades de limpieza y normalización de los datos, como la eliminación de datos nulos, la transformación de los datos en el atributo Vetustez, la transformación de los atributos Area Construida y Area privada en variables netamente numéricas, así como la delimitación espacial de las ofertas inmobiliarias localizadas en la Cabecera Municipal de La Mesa, entre otros.

Al terminar la depuración de la información, se realizó un análisis estadístico exploratorio de la variable de interés que para este caso de estudio es el Valor. Por lo cual se realiza el mapeo de los datos lo cual se visualiza en la siguiente Figura.

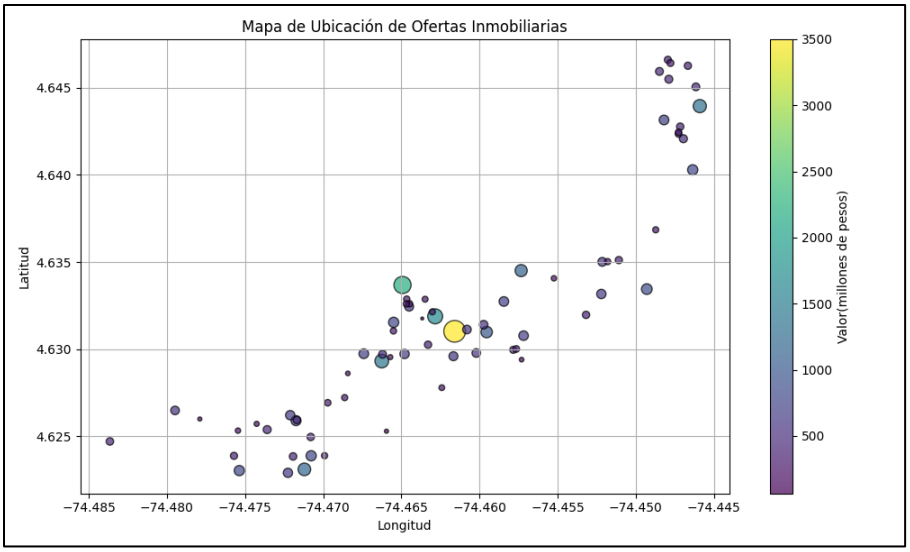


Figura 6 Mapeo de ofertas inmobiliarias según su valor

A partir del mapeo de la variable se visualizó su distribución espacial a través de puntos cuyo tamaño y color representan la magnitud de los valores. Se observa una concentración de valores altos en la zona central del área de estudio, mientras que en las zonas periféricas la distribución es más dispersa, sin una tendencia espacial clara y con predominancia de valores bajos a medios.

Durante este proceso se identificaron 5 valores atípicos (tabla 2) mediante la construcción de una herramienta que genera una gráfica de caja de bigotes, los datos se observan en la Figura 7.

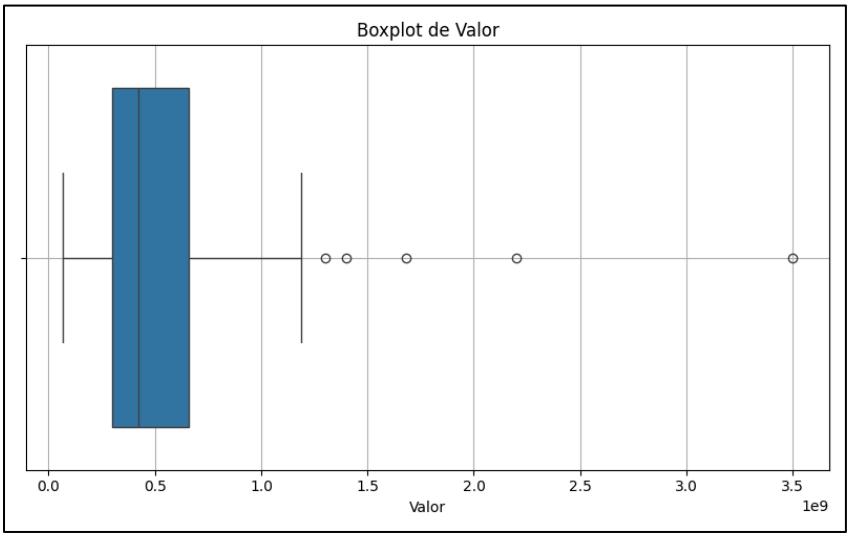


Figura 7 Caja de bigotes para identificación de valores atípicos

Valor
$ 3,500,000,000
$ 2,200,000,000
$ 1,680,000,000
$ 1,400,000,000
$ 1,300,000,000

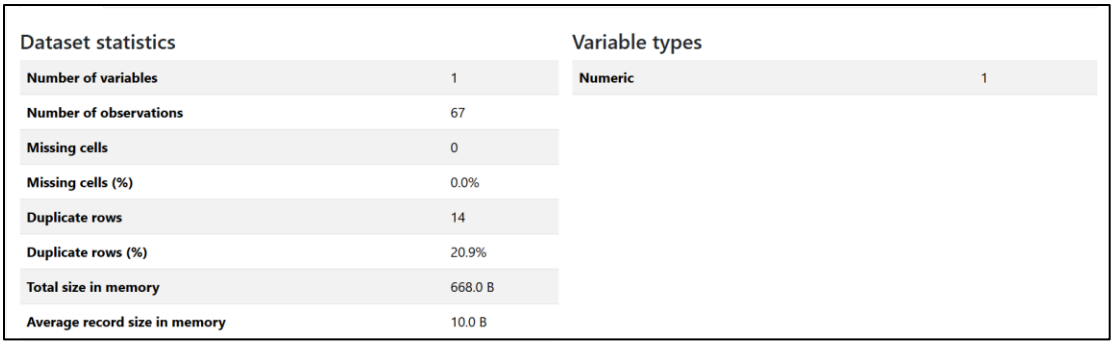
Tabla 2 Valores inmobiliarios atípicos

Para el análisis de estos datos atípicos se consultó el atributo enlace fuente, a partir de ello se evidenció que corresponden a inmuebles de uso comercial, como hoteles y alojamientos turísticos, los cuales cuentan con otros tipos de características que tienden a que ese tipo de ofertas inmobiliarias tengan un precio en el mercado considerablemente más alto, estos datos sesgaban en gran medida la variable de interés por lo que se excluyen del conjunto de datos analizados.

Para la construcción de las demás estadísticas descriptivas, excluyendo los valores atípicos, se desarrolló una herramienta empleando la librería ydata_profiling, a partir de la cual se generó como salida un informe en formato HTML la cual está divida en las siguientes secciones:

- Descripción del conjunto de Datos: La figura 8 presenta un resumen estadístico del conjunto de datos utilizado en el análisis. Se observa que el dataset contiene una sola variable de tipo numérico, con un total de 67 observaciones. No se reportaron valores faltantes, lo que indica que el conjunto de datos está completo en cuanto a registros.

  Sin embargo, se identificaron 14 filas duplicadas, lo que corresponde al 20.9 % del total de observaciones. Esta proporción de duplicados no afectó el conjunto de datos ya que este valor comercial se puede presentar en diferentes ofertas. Por lo tanto, es importante recalcar que los datos duplicados que no aportaban ningún tipo de valor ya fueron depurados en el proceso de limpieza y normalización.



Dataset statistics	
Number of variables	1
Number of observations	67
Missing cells	0
Missing cells (%)	0.0%
Duplicate rows	14
Duplicate rows (%)	20.9%
Total size in memory	668.0 B
Average record size in memory	10.0 B

Variable types	
Numeric	1

Figura 8 Resumen estadístico del conjunto de datos

- Estadísticas Cuantiles y Estadísticas Descriptivas: Figura 9 muestra un conjunto de estadísticas descriptivas y cuantiles, donde observamos que la variable oscila desde los 65 millones de pesos hasta los 1190 millones de pesos, lo que da lugar a un rango de 1125 millones de pesos. Esta amplia dispersión sugiere una heterogeneidad considerable entre las ofertas, lo cual podría estar asociado a diferencias significativas en aspectos como la tipología de los inmuebles o la destinación conomica de cada una de las ofertas inmobiliarias.

  Los cuantiles permitieron identificar esta diversidad: el 25% de las ofertas no supera los 295 millones (Q1), mientras que el 75% se encuentra por debajo de los 620 millones (Q3). La mediana, en 402 millones, indicando un valor central que aún refleja una notable dispersión, reforzada por un IQR de 325 millones. Además, los percentiles extremos muestran que solo un 5% supera los 830.5 millones, lo que evidenció la existencia de un segmento exclusivo.

  Por otro lado, las estadísticas descriptivas mostraron un valor promedio aproximado a los 464 millones de pesos, mientras que la mediana se situó en los 402 millones de pesos, donde se pudo evidenciar una distribución asimétrica positiva, esto se refuerza con el coeficiente de asimetría de 0.87 y el histograma que se visualiza en la figura 10,

estos valores y distribución en el histograma indicaron una clara concentración de Precios bajos.

La mediana absoluta de las desviaciones (MAD) fue de 148 millones, y el coeficiente de variación (0.50) señalaron una variabilidad moderada respecto al valor medio. Por su parte, la curtosis de 0.74 revelo una distribución más plana que la normal, con colas ligeras y menor presencia de valores extremos. En conjunto, estos resultados reflejaron un mercado inmobiliario heterogéneo, pero sin anomalías severas lo que indicó que la exclusión de valores atípicos redujo en gran medida la heterogeneidad, lo cual facilitó el análisis y representación de los datos.

Quantile statistics

Minimum	65000000
5-th percentile	1.6×10^8
Q1	2.95×10^8
median	4.02×10^8
Q3	6.2×10^8
95-th percentile	8.305×10^8
Maximum	1.19×10^9
Range	1.125×10^9
Interquartile range (IQR)	3.25×10^8

Descriptive statistics

Standard deviation	2.3353691×10^8
Coefficient of variation (CV)	0.50258546
Kurtosis	0.74515979
Mean	4.6467104×10^8
Median Absolute Deviation (MAD)	1.48×10^8
Skewness	0.87533353
Sum	3.113296×10^{10}
Variance	5.4539488×10^{16}
Monotonicity	Decreasing

Figura 9 Estadísticas Cuantiles y Estadísticas Descriptivas

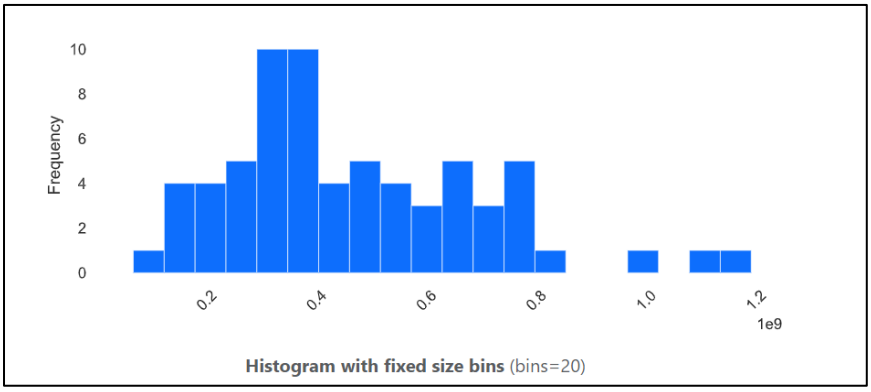


Figura 10 Histograma

Posteriormente, se validó la normalidad de los datos con una gráfica QQ plot para el conjunto de valores atípicos y sin valores atípicos, como se visualiza en la figura 11.

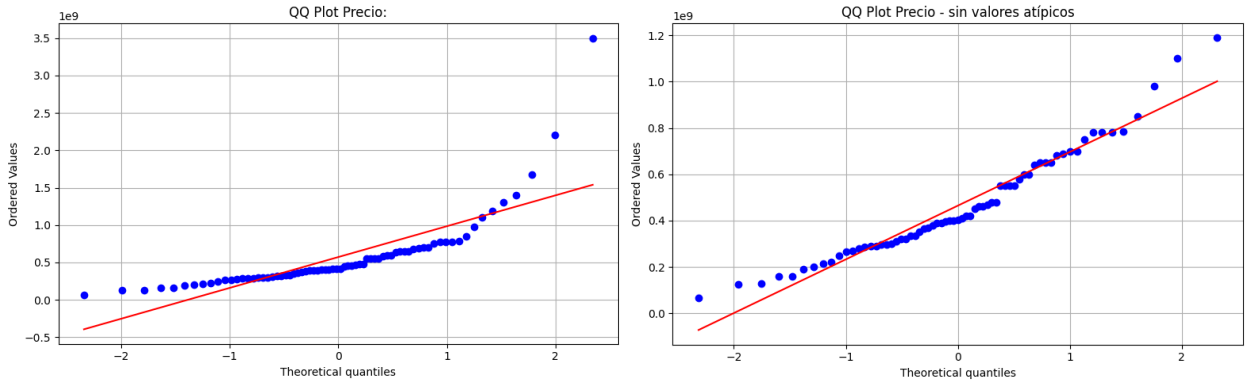


Figura 11 QQ plot

Al observar las gráficas anteriores, se observó que la eliminación de valores atípicos redujo la presencia de colas pesadas en los extremos de las líneas. Sin embargo, aún se observó un patrón curvo en la parte central de la línea de referencia lo que indica una desviación respecto a la normalidad. Aunque las colas ahora son más ligeras, persisten, lo cual indica que los datos no siguen una distribución normal.

Posteriormente, se realizó una prueba de normalidad mediante la prueba de Lilliefors bajo un nivel de significancia del 5%. La prueba arrojó un estadístico D con un valor de 0.12809, el cual se comparó con el valor crítico aproximado de 0.096 correspondiente a un tamaño poblacional de 67 observaciones,

Dado que D > 0.096, se rechazó la hipótesis nula de normalidad. Además, el *p* - valor es menor al nivel de significancia, lo que reforzó la no normalidad de los datos, por lo que se concluyó que los datos no siguen una distribución normal, con un 95% de confianza.

Lilliefors (Kolmogorov-Smirnov) normality test
D = 0.12809, p-value = 0.008225

Posterior al análisis estadístico exploratorio, se desarrolló una herramienta basada en el método de estimación de densidad de Kernel ponderado, utilizando la librería sklearn. Para el cálculo de distancias se empleó una métrica no euclidiana, dado que los datos corresponden a distancias angulares. En este caso, se utilizó la métrica Haversine , la cual calcula la distancia entre dos puntos sobre la superficie de una esfera. Esta fórmula considera que la primera coordenada corresponde a la latitud y la segunda a la longitud, ambas expresadas en radianes (Scikit learn, 2025).

El uso de esta métrica es debido a que la Tierra puede aproximarse como una esfera, lo que permite a la fórmula de Haversine estimar con buena precisión la distancia entre dos ubicaciones geográficas, con un error promedio menor al 1%.

Dado que la mayoría de los algoritmos de librería sklearn están diseñados para métricas euclidianas, se utilizó el algoritmo Ball Tree, que sí permite integrar la métrica Haversine. Para la estimación de la densidad, se aplicó un Kernel gaussiano como función de suavizado.

Finalmente, para determinar el parámetro de ancho de banda ideal en el proceso de estimación, se empleó el método GridSearchCV, el cual nos indicó que el parámetro de ancha de banda ideal es de 0.0008° grados, la figura 12 muestra el resultado de la estimación de la densidad espacial utilizando este método sobre los datos recopilados. El mapa está delimitado según el polígono de la cabecera municipal de La Mesa y permite visualizar la distribución espacial del valor de mercado de las ofertas inmobiliarias estimado en el territorio.

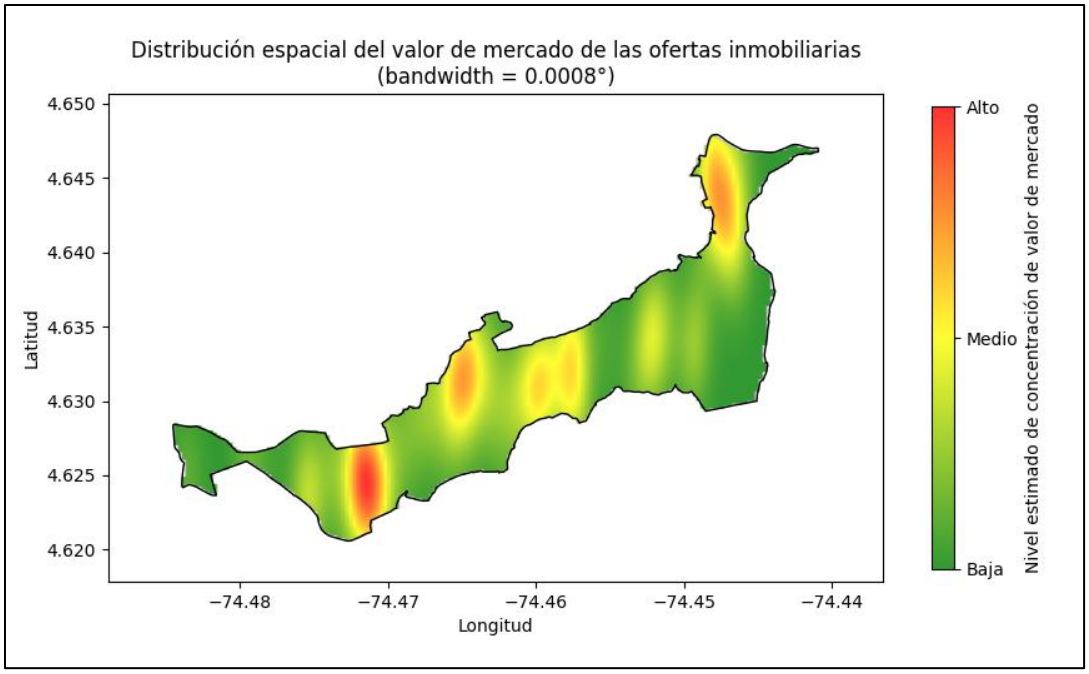


Figura 12 Mapa de distribución Kernel

En el mapa generado se puede visualizar el nivel estimado de concentración del valor de mercado de las ofertas inmobiliarias en la cabecera municipal. La leyenda clasifica las concentraciones en tres niveles: bajo (verde), medio (amarillo) y alto (rojo). Se observa que las zonas con mayor concentración de valores altos tienden a ubicarse en sectores específicos del territorio, lo que puede estar relacionado con factores como el acceso a vías principales, cercanía a servicios o presencia de desarrollos urbanísticos recientes.

Además, la distribución no es homogénea, lo que evidencia una posible segmentación espacial del mercado inmobiliario. Estas zonas de alta concentración podrían indicar focos de valorización o mayor demanda, mientras que las áreas con valores bajos podrían estar asociadas a menor interés comercial o restricciones en el uso del suelo.

8. Conclusiones

La herramienta desarrollada a partir de la técnica de web scraping recolecta los datos haciendo un rastreo página por página extrayendo los atributos planteados, almacenándolos en un dataset de Excel facilitando la estructuración de columna y fila planteada durante el desarrollo de la metodología. Durante el escaneo que realizó, se pudo identificar casos especiales que dificultaban la extracción debido a que había presencia de codigo ASCII por esto fue fundamental en el script hacer una expresión regular que analizara y remplazara estos caracteres.

Mediante el uso de las librerías Pandas, Geopandas y re se desarrollaron acciones para la limpieza y normalización de los datos, en cuanto a limpieza se depuraron aquellas ofertas inconsistentes o que estuvieran fuera de la zona de estudio (zonas rurales o fuera del municipio), mientras que para la normalización se transforman variables, entre estas la vetustez debido a que se encuentran datos numéricos, en rango o como cadena de texto y puede llegar a ser un obstáculo para su manejo.

A partir de la información recolectada y procesada se realizó un análisis estadístico espacial de distribución de Kernel ponderada, evidenciando la usabilidad de estos datos como insumos referenciales, como resultado final se obtiene el mapa de distribución espacial del valor del mercado de ofertas inmobiliarias. En este mapa se puede visualizar en donde se encuentran las mayores concentraciones de valor según las ofertas inmobiliarias a partir de una simbología de tres categorías: Bajo, Medio y Alto.

Aunque este trabajo se enfoca principalmente en un análisis cuantitativo, es importante mencionar que también se incorporan variables cualitativas relevantes, como descripción de la oferta o el tipo de inmueble. Estas últimas enriquecen la información al ofrecer una perspectiva más completa y

contextualizada del mercado inmobiliario, logrando una visión integral que combina la precisión numérica con la profundidad cualitativa.

Teniendo en cuenta lo desarrollado a lo largo del proyecto, se evidenció que la herramienta desarrollada es adaptable y escalable. Es adaptable, ya que puede emplearse no solo en el campo inmobiliario, sino también en otras temáticas. Aunque en este caso se definieron ciertos atributos, la herramienta puede ajustarse a cualquier conjunto de atributos que el usuario requiera. Asimismo, es escalable, pues si bien en el informe se presentó el proceso aplicado a una única página inmobiliaria, su implementación es posible en múltiples páginas con solo conocer la estructura HTML de la página web.

La aplicación de tecnologías de información geográfica y técnicas automatizadas de recolección de datos representa un avance significativo para municipios con baja capacidad técnica. Esta herramienta puede ser un primer paso hacia la digitalización de procesos en entidades territoriales, contribuyendo a reducir la brecha tecnológica e institucional que actualmente limita la gestión eficiente del territorio.

El proyecto demuestra que los datos disponibles en portales inmobiliarios, aunque no oficiales, pueden convertirse en insumos complementarios útiles para el análisis del mercado de suelo urbano. La integración de datos abiertos con herramientas de análisis espacial fortalece el conocimiento territorial y permite abordar fenómenos como la expansión urbana desordenada o la desigualdad en la asignación de recursos públicos.

Esta información puede ser utilizada por las autoridades locales para tomar decisiones más informadas en materia de planificación territorial, superando parcialmente las limitaciones derivadas de la ausencia de actualización catastral.

9. Recomendaciones

Para el desarrollo de esta técnica es indispensable tener conocimiento del lenguaje de marcado de hipertexto junto con conocimiento de lenguaje de estilo CSS, con el fin de poder identificar en donde se localizan los atributos de interés en la estructura de la página web, a demás se recomienda tener conocimiento básico de lenguaje de programación que pueda hacer uso de librerías como Selenium por ejemplo, Python o Java.

Este estudio se limita a la recopilación y análisis estadístico espacial básico, estas técnicas se pueden emplear en futuras investigaciones donde se podrían profundizar en análisis cualitativos, explorando percepciones de compradores, opiniones de expertos o el impacto de factores socioeconómicos subjetivos en el valor de las propiedades, algo que no fue el objetivo principal.

Esta herramienta tiene un gran potencial para resolver las necesidades de municipios con bajo presupuesto tecnológico, se sugiere avanzar hacia un sistema más robusto que incluya una interfaz gráfica o dashboard para facilitar la visualización y análisis de la información recopilada, así como el almacenamiento sistemático de los datos en bases de datos o archivos organizados. Además, se propone ampliar los módulos de análisis mediante técnicas como series temporales, co-kriging, regresiones espaciales, filtros interactivos o clustering espacial, que permitan explorar dinámicas más complejas del mercado inmobiliario u otras temáticas.

10.Referencias

Bala, A. K., & Kumar, A. (2025). Research Progress Knowledge Mapping on Transit Oriented Development (TOD) Typology. *Engineering, Technology and Applied Science Research*, *15*(2), 21847-21853. Scopus. https://doi.org/10.48084/etasr.10097

Carranza, J. P., Piumetto, M. A., Lucca, C. M., & Da Silva, E. (2022). Mass appraisal as affordable public policy: Open data and machine learning for mapping urban land values. *Land Use Policy*, *119*, 106211. https://doi.org/10.1016/j.landusepol.2022.106211

Elvestad, H. E., & Holsen, T. (2024). Valuation practices in urban land readjustment cases in Norway. *Land Use Policy*, *145*, 107242. https://doi.org/10.1016/j.landusepol.2024.107242

Ertunç, E., Muchová, Z., Tomić, H., & Janus, J. (2022). Legal, Procedural and Social Aspects of Land Valuation in Land Consolidation: A Comparative Study for Selected Central and Eastern Europe Countries and Turkey. *Land*, *11*(5). Scopus. https://doi.org/10.3390/land11050636

Harris, R. (2025). *The rise of the neighbourhood in Canada, 1880s-2020s* (p. 353). Scopus. https://doi.org/10.3138/9781487520441

Hsu, Y.-Y., & Han, H. (2025). Toward volumetric urbanism: Analysing the spatial-temporal dynamics of 3D floor space use in the built environment. *Environment and Planning B: Urban Analytics and City Science*, *52*(5), 1181-1196. Scopus. https://doi.org/10.1177/23998083241286592

Ji, Z., Yin, L., & Wang, J. (2026). Spatiotemporal Pattern Analysis Of Land Use In Jiangsu Province Based On Long-term Time Series Remote Sensing Images. *Journal of Applied Science and Engineering*, *29*(1), 1-12. Scopus. https://doi.org/10.6180/jase.202601_29(1).0001

Jiang, F., & Ma, J. (2025). Predicting Urban Vitality at Regional Scales: A Deep Learning Approach to Modelling Population Density and Pedestrian Flows. *Smart Cities*, *8*(2). Scopus. https://doi.org/10.3390/smartcities8020058

Mei, Y., Sijtsma, F., Ballas, D., Vegter, D., & Ongena, Y. (2023). *"The Most Deafening Silence You Can Hear": The International Comparison of Landscape Value Mapping on Linguistic Expressions of Islands: Vol. 14112 LNCS* (p. 14). Scopus. https://doi.org/10.1007/978-3-031-37129-5_1

Rojas Cárdenas, J. D., & Ruiz Ocampo, J. M. (2020). *Apoyo y soporte técnico en labores ingenieriles para la ejecución del proyecto "Determinación de valores del suelo y comerciales de inmuebles propiedad de Empresa de Telecomunicaciones" llevado a cabo por la firma Salmer Ingeniería S.A.S*. http://hdl.handle.net/11349/25283

Salgado Miranda, J. H., Rivera Jiménez, F. A., & Bustos Urrego, W. Y. (2024). *Aplicativo web geográfico del estudio del valor del suelo de la localidad de Usaquén en la ciudad de Bogotá*. http://hdl.handle.net/11349/36306

Swain, S. S., Raghavendra Rao, R. V., Mohan Reddy, C. R., Kaur, H., Meenakshi, & Boopathi, S. (2025). Predictive analytics for urban development using ML technique. En *Urban Mobility and Challenges of Intelligent Transportation Systems* (pp. 383-402). Scopus. https://doi.org/10.4018/979-8-3693-7984-4.ch021

Tuta Rincón, Á. R., & Vela Castro, L. V. (2018). *Análisis Espacial del Mercado del Suelo Rural del Municipio de Chía en el Año 2017 para la Determinación de Superficies de Valor*. http://hdl.handle.net/11349/7755

Yan, J., Tong, Z., Wang, Y., Feng, S., Su, Y., Song, Y., Wen, Q., & Yin, C. (2025). From measurements to regulations: An actionable approach for sustainable urban cooling via heat-

resilient urban planning. *Sustainable Cities and Society*, *125*. Scopus. https://doi.org/10.1016/j.scs.2025.106361

Hernández, R., Fernández, C. y Baptista, P. (2010). *Metodología de la investigación*. (5ª ed.) McGrawHill.

Ministerio de Ciencia, Tecnología e Innovación (2021). *Anexo 1 Convocatoria nacional para el reconocimiento y medición de grupos de investigación, desarrollo tecnológico o de innovación y para el reconocimiento de investigadores del Sistema Nacional de Ciencia, Tecnología e Innovación – 2021.* https://minciencias.gov.co/sites/default/files/upload/convocatoria/anexo_1_-_documento_conceptual_2021.pdf

Concejo Municipal de La Mesa Cundinamarca. (2005). Acuerdo 003, Colombia OT.

Concejo municipal La Mesa Cundinamarca. (22 de Mayo de 2000). Acuerdo N° 005, Colombia OT. Obtenido de https://serviciosgeovisor.igac.gov.co:8080/Geovisor/descargas?cmd=download&token=eyJhbGciOiJIUzUxMiJ9.eyJzdWIiOiIxMDAwMDciLCJleHAiOjE3NDk0NDMyMjMsImp0aSI6ImRvY3VtZW50by0xMjcyMyJ9.3iq9dW-8d0EPFrnNxVhP4JugLndGEQaWozSBvQIOEvLzizFYOreAI3zoJVyTY7t4OPaZs2jUjzBd

Martínez, D. (2020). DISEÑO DE BOCATOMA Y LÍNEA DE CONDUCCIÓN PARA GARANTIZAR EL CAUDAL DEL SISTEMA DE ACUEDUCTO DE LA CABECERA MUNICIPAL DE LA MESA, CUNDINAMARCA. Obtenido de https://repository.umng.edu.co/server/api/core/bitstreams/50f64024-ac3a-4c03-a459-78d3ef4bec55/content

Palacios, A., Torres, J., Villarraga, L., & Meza, M. (2017). Valoración del paisaje y evaluación del potencial interpretativo como herramienta para el turismo sostenible en el Ecoparque Las Monjas (La Mesa, Cundinamarca). Obtenido de http://www.scielo.org.co/pdf/rcdg/v26n2/0121-215X-rcdg-26-02-00177.pdf

Scikit learn. (2025). Scikit Learn. Obtenido de https://scikit-learn.org/stable/modules/generated/sklearn.metrics.pairwise.haversine_distances.html

Trejos, A. (17 de 05 de 2025). En Colombia llevamos 16 años sin actualización catastral en la mayoría de municipios: Director IGAC. Caracol Radio, págs. https://caracol.com.co/2025/05/17/en-colombia-llevamos-16-anos-sin-actualizacion-catastral-en-la-mayoria-de-municipios-director-igac/.